

Copyright
by
Amelia Jean Druckenbrod
2014

**The Report Committee for Amelia Jean Druckenbrod
Certifies that this is the approved version of the following report:**

**Reframing “Quality” in Quality Rating and Improvement Systems:
A Critical Analysis**

**APPROVED BY
SUPERVISING COMMITTEE:**

Supervisor:

Christopher Brown, Supervisor

Noah De Lissovoy

**Reframing “Quality” in Quality Rating and Improvement Systems:
A Critical Analysis**

by

Amelia Jean Druckenbrod, B.A

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

December 2014

Abstract

Reframing “Quality” in Quality Rating and Improvement Systems: A Critical Analysis

Amelia Jean Druckenbrod, M.A.

The University of Texas at Austin, 2014

Supervisor: Christopher Brown

Abstract: Quality rating and improvement systems (QRIS) are a state-level policy designed to assess and improve quality in early childhood education and care through rating systems and financial incentives. Current research suggests that QRIS are not meeting their stated goals of increasing access to quality care and improving child outcomes. This report investigates concepts of quality in QRIS by critically analyzing their use of standardized quality measurement scales. It uses postmodern perspectives and cultural relativism theory to argue for an alternate conceptualization of quality that incorporates community context and multiple perspectives. Finally, this report proposes alternative policies that could be used to promote ongoing conversations about quality within a community context.

Table of Contents

| | |
|---|----|
| I. Introduction | 1 |
| I.I The Increasing Popularity of QRIS Policy..... | 4 |
| I.II Standardized Assessment in QRIS..... | 5 |
| I.III Problems with QRIS' Framing of Quality | 6 |
| I.IV Structure of the Paper | 7 |
| II. Background Information on Quality Rating and Improvement Systems..... | 9 |
| III. Goals of Quality Rating and Improvement Systems..... | 12 |
| III.I QRIS Programs and Improving Access to Quality Care for Families...12 | |
| III.II The Ability of Early Childhood Programs to Access QRIS..... | 15 |
| III.III The Importance of Issues of Access in a High-Stakes Context | 16 |
| III.IV Current Concepts of Quality and Quality Measurement Related to Issues of Access | 17 |
| III.V QRIS Programs and Improving Children's School Performance | 18 |
| III.VI Quality Measurements in QRIS and Their Relation to Child Outcomes | 20 |
| IV. Concepts of Quality in the Early Childhood Environment Rating Scale- Revised..... | 25 |
| IV.I The Development of the ECERS-R..... | 25 |
| IV.II Current Uses of the ECERS-R in Research and Policy | 26 |
| IV.III Assessing Quality Using the ECERS-R Instrument | 28 |
| IV.IV The ECERS-R and Its Assertions about Quality..... | 29 |
| V. Critiques of the Early Childhood Environment Rating Scale-Revised..... | 32 |
| V.I Critiques on the Reliability of the ECERS-R | 32 |
| V.II Critiques of Validity Issues Involving the ECERS-R | 34 |
| V.III Critiques of Other Content in the ECERS-R | 37 |
| V.IV Critiques from Teachers in QRIS that use the ECERS-R Assessment | 39 |

| | |
|--|----|
| V.V Critique of the ECERS-R and its Use in QRIS Policy | 40 |
| VI. Framing Current Concepts of Quality and Quality Measurement..... | 43 |
| VI.I Current Conceptualization of Quality in Early Childhood Discourses and Research..... | 44 |
| VI.II Disrupting Universal and Objective Claims about Early Childhood Quality..... | 45 |
| VI.III Implications of the Current Conceptualization of Quality in High-Stakes Policy | 48 |
| VI.IV An Alternative Model of Quality in Early Childhood Education and Care..... | 51 |
| VII. Alternative Approaches to Promoting Quality through Early Childhood Policy | 54 |
| VII.I Empowering Parent Choice | 55 |
| VII.II Defining “Quality” as Ongoing Conversation Responsive to Community Context..... | 58 |
| VII.III QRIS Policy Recommendations..... | 60 |
| VIII. Conclusion | 62 |
| VIII.I Opportunities for Further Research | 63 |
| IX. Appendices | 67 |
| IX.I Appendix A | 68 |
| IX.II Appendix B..... | 70 |
| XI.III Appendix C..... | 74 |
| References..... | 75 |

I. Introduction

Research has shown a strong relationship between children's performance in school and their early childhood experiences, including participation in early childhood education and care (ECEC) programs (Shonkoff & Phillips, 2000). In a summary and analysis of the literature, Love, Schochet, & Meckstroth (1996) found that children appeared to benefit both socially and cognitively from being enrolled in quality child care. Love et al. (1996) write that the studies they reviewed define quality care as “developmentally appropriate programs, [where] caregivers encourage children to be actively engaged in a variety of activities; have frequent, positive interactions with children . . . use positive guidance techniques, and encourage appropriate independence” (p. 5). Such findings demonstrate that exposure to quality early childhood programs can improve a child's cognitive skills and ensure that he or she is ready for school, while negative interactions or high-stress environments during the early years can impair brain development and lead to negative outcomes for children (Shonkoff & Phillips, 2000). Despite the benefits of quality care, the majority of children are in low- or mediocre-quality care, with low-income children the most likely to be enrolled in low-quality care (Cost, Quality and Outcomes Study, 1995; Love et al., 1996). Based on these findings, economists and policymakers have argued that increasing access to quality ECEC is a cost-effective and efficient measure to increase school readiness and improve child outcomes (e.g. Heckman, Pinto, & Savelyev, 2012).

One of the ways policymakers and early education advocates have sought to improve the quality of care in ECEC is through Quality Rating and Improvement

Systems (QRIS). QRIS are a systematic, “comprehensive approach to understanding, assessing, and improving early care and education programs at the state level” (Hestenes, Kintner-Duffy, Wang, Paro, Mims, Crosby, Scott-Little, & Cassidy, 2014, p. 2). Within a QRIS, states typically set structural requirements for programs and use standardized measurements to define and assess quality (Connors & Morris, 2014; Jeon, Buettner, & Hur, 2014; Le, Schaack, & Setodji, 2014; Tarrant & Huerta, 2014). Program eligibility is determined by each state’s QRIS policy (National Center on Child Care Quality Improvement). In most states, a variety of different early childhood programs can participate in QRIS, including private child care facilities, family day care centers, Head Start programs, and public pre-kindergarten classrooms (Tarrant & Huerta, 2014). Based on the standards and assessment score, early childhood programs are given a quality rating within a standardized, statewide rating system (Connors & Morris, 2014; Jeon et al., 2014; Le et al., 2014; Sabol & Pianta, 2014; Tarrant & Huerta, 2014). These ratings are advertised to parents and tied to financial incentives such as subsidy reimbursements, improvement grants, and teacher bonuses (Hofer, 2010; Le et al., 2014; Stoney, 2012; Tarrant & Huerta, 2014). State agencies have used a variety of sources to fund these QRIS programs, including: federal grants, lottery or tax revenue, Head Start or Temporary Aid for Needy Family (TANF) funding, and private philanthropic donations (National Center on Child Care Quality Improvement). States hope by using a rating system, parents will seek early childhood programs with a higher rating, and the programs themselves will be motivated to improve their quality in order to stay competitive within the childcare market (Jeon et al., 2014; Sabol & Pianta, 2014; Tarrant

& Huerta, 2014). Both measures aim to use neoliberal ideals of choice and market competition to increase the quality of early childhood care programs (Nagasawa, Peters & Swadener, 2014).

Advocates of quality rating and improvement systems point to two main benefits of the programs: increasing the quality of available childcare for young children and families and improving children's outcomes and school performance (Hestenes et al., 2014; QRIS National Learning Network, 2009). These goals are typically used causally; it is assumed that improving the quality of care will improve children's outcomes in the future (Hestenes et al., 2014). However, researchers have argued that QRIS have the potential to increase the cost of early care — decreasing access — and existing research on QRIS and child outcomes have not conclusively proven that these systems improve child outcomes (Child Trends, 2010; Duncan, Jenkins, Auger, Burchinal, Domina, Bitler, 2014).

This paper will critically examine QRIS policy by analyzing current concepts of quality in QRIS policy and argue for a reconceptualization of quality that incorporates multiple perspectives and community context. It will do so by focusing in on the Early Childhood Environment Rating Scale-Revised (ECERS-R). Analyzing the ECERS-R offers the chance to examine how quality is being conceptualized within the majority of these QRIS across the United States as well as explore why these assessments are potentially problematic in policy. Arguments in this paper suggest that a more nuanced view of quality in early childhood research and policy could be used to create alternative QRIS that are more responsive to the needs of young children and families.

1.1 The Increasing Popularity of QRIS Policy

QRIS have been frequently recommended in early childhood research and policy (see: Blau, 2007; Rigby, Ryan, & Brooks-Gunn, 2007; Scarr, Eisenberg, & Deater-Deckard, 1994), although many states have not enacted QRIS programs until recently (Tout, Starr, Soli, Moodie, Kirby, & Boller, 2010). These recommendations were based on research by economists that found the current system of mandatory child care regulations constricted markets and did not improve quality levels (Blau, 2003, 2007). Many researchers and economists argued that a system of ratings and financial incentives, similar to programs developed in Oklahoma and North Carolina, would use market forces to motivate providers to improve quality levels (Blau, 2007; Rigby et al., 2007; Scarr et al., 1994). Oklahoma is the first state credited with developing a QRIS — it implement a rating system in 1998 — followed by North Carolina in 1999 (Tout et al., 2010). Despite these recommendations, by 2005 only nine states had implemented QRIS, two had pilot QRIS programs, and one state had a regional QRIS (Tout et al., 2010).

QRIS systems gained importance after being a priority in both the 2010 and 2013 Race to the Top - Early Learning Challenge grant application (Stoney, 2012; Tarrant & Huerta, 2014). In order to receive the grant, states were required to implement a tiered quality rating and improvement system — or a QRIS with a standardized assessment and rating system for early childhood programs. Currently, thirty-nine states have implemented QRIS, eight states are in the process of planning or piloting QRIS, and two states have regional QRIS programs (National Center on Child Care Quality

Improvement). Although QRIS are primarily a recent phenomenon, the amount federal money tied to these systems has led to a policy situation where every state except Missouri is either piloting or implementing their own version of a QRIS (National Center on Child Care Quality Improvement).

I.II Standardized Assessments in QRIS

The specific policies of a QRIS vary by states; however, almost every state uses a standardized assessment as an accountability measure and ties financial incentives to the assessment score (Connors & Morris, 2014; Jeon et al., 2014; Le et al., 2014; Sabol & Pianta, 2014; Tarrant & Huerta, 2014). The majority of states use the Environmental Rating Scales (ERS) to assess programs, a group of scales comprised of: the Early Childhood Environmental Rating Scale-Revised (ECERS-R), for classrooms serving children ages two and a half to five; the Infant-Toddler Environmental Rating Scale-Revised (ITERS-R), for classrooms serving children birth through two and a half; the School-Age Care Environmental Rating Scale-Revised (SACERS-R), for group-care programs serving children five through twelve years old; and the Family Day Care Environmental Rating Scale-Revised (FDCERS-R), for care programs in a provider's home (Le et al., 2014; National Center on Child Care Quality Improvement). Of these scales, the ECERS-R for preschool classrooms is the most widely used in early childhood research and the most discussed in QRIS policy (Fenech, 2011; Gordon, Fujimoto, Kaestner, Korenman, & Abner, 2013; Le et al., 2013; Perlman, Zellman, & Le, 2004).

Although QRIS typically have structural standards in addition to the assessment, qualitative research with teachers participating in QRIS found that teachers placed a very high emphasis on the instrument used to assess them (Tarrant & Huerta, 2014). Teachers believed that this instrument represented the views of quality espoused by the QRIS and often changed their classrooms and teaching to perform better on the assessment (Tarrant & Huerta, 2014). This research suggests that the instrument used to observe and assess programs is particularly important when analyzing concepts of quality that are conveyed in QRIS. As the ECERS-R is the most widely used scale in QRIS policy (Duncan et al., 2014; Gordon et al., 2013; National Center on Child Care Quality Improvement), this paper will examine the ECERS-R instrument as a discourse on quality within QRIS policy.

I.III Problems with QRIS' Framing of Quality

Concepts of quality in early childhood policy position “quality” as a universal, objective reality (Dahlberg, Moss & Pence, 1999). If quality is objective, it can be then discovered, thoroughly defined, and represented through measurement instruments (Dahlberg et al., 1998; Lubeck, 1998; Tobin, 2005). If quality is universal, quality measurements and standards can then be systematically applied to all settings regardless of community context (Lubeck, 1998; Tobin, 2005). These two discourses contribute to a dominant concept of “quality” that is rarely questioned or analyzed (Dahlberg, Moss & Pence, 1999). Peter Moss (1994) writes:

‘Quality’ is an international buzz word, not only in early childhood services but in connection with every kind of product and service. Yet in its mantra-like

repetition, the word is in danger of being rendered meaningless. It attracts widespread support — for who could not want ‘good quality’? — unless and until we have to say what we actually mean, at which point it becomes far more elusive (p. 1).

Quality is repeatedly identified as the end goal within early childhood policy, yet as Moss writes, the word has become so repeated that its meaning is assumed. QRIS are based on the assumption that quality has been accurately defined and encapsulated in a standardized instrument, and that scoring higher on this assessment will improve the quality of care and result in meaningful outcomes for all children in the program (Hestenes et al., 2014). Yet there is little research or reflection on the concepts of quality in the standardized instruments and whether these concepts align with broader goals of QRIS (Hestenes et al., 2014). Due to the large role that quality plays in QRIS programs — both in terms of a rating score and financial incentives — these basic assumptions must be critically examined and compared with other conceptualizations of “quality” in early childhood care.

I.IV Structure of the Paper

This paper will use a critical framework to deconstruct assumptions about quality in QRIS, particularly in light of the policy’s goals of improving child outcomes and increasing access to quality care. First, this paper will present additional background on state’s quality improvement systems and their stated goals, followed by a discussion of potential reasons why the current structure of QRIS do not meet these goals. Next, this paper will explore the ECERS-R scale as an investigation into the way quality is typically

defined and measured in QRIS policy. This paper will analyze critical research on the scale to suggest potential problems with using the scale to measure quality in policy. The paper will frame larger discourses on quality in ECEC policy to disrupt the dominant idea that quality is universal and adequately measured in standardized instruments, presenting an alternative concept that allows for flexible, context-specific definitions of quality. Finally, this paper will examine a small minority of states who are not using pre-selected standardized assessments in their QRIS to argue for alternative policy approaches to conceptualizing and improving quality. Throughout this analysis, this paper will show that QRIS and their high-stakes use of standardized assessments promote a very limited and rigid concept of quality that does not advance early childhood education and care. Instead, this paper proposes solutions that incorporate stakeholder perceptions and community context into policy that better meets the needs of parents and young children.

II. Background Information on Quality Rating and Improvement Systems

Quality rating and improvement systems (QRIS) are a state-level policy approach to improving quality in early childhood education and care. Currently thirty-nine states have implemented QRIS programs, three states are piloting QRIS programs, five states are in the process of planning QRIS, and two states have regional QRIS program (National Center on Child Care Quality Improvement). Hestenes et al. (2014) identified five essential aspects of QRIS: 1) quality standards, 2) accountability measures to assess quality standards, 3) technical support for programs to improve quality, 4) financial incentives for participation, and 5) efforts to educate families in child care programs.

Aside from these common elements, QRIS vary greatly between states (Connors & Morris, 2014). Each state has a different QRIS program and states use different standardized assessments, different scoring methods, and often have different tiers of achievement (National Center on Child Care Quality Improvement). For example, Arizona's Quality First rates early childhood programs from one to five stars, while Illinois designates programs as having a "Licensed Circle of Quality," "Bronze Circle of Quality," "Silver Circle of Quality," or "Gold Circle of Quality" (National Center on Child Care Quality Improvement). Almost all states have a tiered rating system, such as the aforementioned examples, that represents levels of quality and is used to determine financial incentives (Connors & Morris, 2014; Jeon et al., 2014; Let et al., 2014; Sabol & Pianta, 2014; Tarrant & Huerta, 2014). These incentives also vary based on the state: improvement grants can range from \$250 to \$5,000 across QRIS programs (Mitchell, 2012). Teacher bonuses for participating in the QRIS program range from \$200-\$1,000 in

Maryland to \$600-\$4,000 in Pennsylvania, with the exact amount determined by their program's quality rating (Mitchell, 2012). Although the majority of QRIS rely on a rating system and financial incentives to improve quality and improve school performance, these measures vary both in their structure and size of financial incentives.

QRIS programs share a common goal to improve quality; however, specific definitions and standards of quality vary based on the standardized assessment used to assess quality. Of the thirty-nine states that currently have QRIS program, five states do not use any kind of program assessment or do not specify which assessment must be used (National Center on Child Care Quality Improvement). Four states use a program assessment that is different from the Environment Rating Scales; two of these states developed an original program assessment for their QRIS. Eighteen states use the ERS scales (ECERS-R, ITERS-R, FDCERS-R, and SACERS-R) as the only program assessments in the QRIS. Twelve states use the ERS in conjunction with other scales. Massachusetts' QRIS program uses the most assessments — eight other assessment instruments in addition to the ERS (National Center on Child Care Quality Improvement). Although states have implemented a variety of QRIS programs, the ERS are commonly used across most programs. Over three-quarters of states with QRIS use the ERS in some capacity and almost one half of states use the ERS as the only assessment (National Center on Child Care Quality Improvement).

Participation in QRIS is still voluntary in the majority of states (Jeon et al., 2014). In many states, a program must apply or opt in to participate in the QRIS program (National Center on Child Care Quality Improvement). A few states, such as Oklahoma

and North Carolina, automatically place all licensed centers on the first tier of quality within the QRIS and give programs the option of working towards higher ratings (National Center on Child Care Quality Improvement). No states mandate that all early childhood programs are evaluated or must move beyond the first level (Connors & Morris, 2014), although states encourage participation through grants, tax credits, tiered subsidy reimbursements, and other financial incentives (Hofer, 2010; Le et al., 2014; Stoney, 2012; Tarrant & Huerta, 2014).

III. Goals of Quality Rating and Improvement Systems

Advocates of quality rating and improvement systems point to two main benefits of the programs: increasing the quality of available childcare for young children and families and improving children’s social-emotional and cognitive outcomes (Hestenes et al., 2014; QRIS National Learning Network, 2009). Policymakers use research that links ECEC experiences and school performance, as discussed in the introduction, to support the assumption that improving the quality of care will improve children’s outcomes in the future (Hestenes et al., 2014; Kaurez & Thorman, 2011). However, policy must move beyond a focus on measuring quality to analyze the concepts of quality assessed by the quality measurements. Hestenes et al. (2014) wrote: “Unfortunately, recent discussions of measurement efforts related to QRIS demonstrate that there is often a lack of alignment among the goals of QRIS assessment efforts and constructs addressed in the quality measures” (p. 5). This section will analyze two central goals of QRIS in order to illustrate how current policy structures and quality assessments do not align with these stated goals.

III.1 QRIS Programs and Improving Access to Quality Care for Families

One of the main goals of QRIS programs is to increase the level of quality in child care programs, so that children and families will have more opportunities to enroll in high-quality care programs (QRIS National Learning Network, 2009). Although this goal is common across QRIS, states specify different objectives and aims within this goal. For example, the stated vision of Delaware’s Stars for Early Success is that: “All of

Delaware’s children and families will have access to quality early care and education from birth through school-age (QRIS National Learning Network, 2009, p. 1). Maine’s Quality for ME program states that its goals are to “Recognize child care programs that provide quality care, [e]ncourage providers to increase their level of quality, and [p]rovide parents with identifiable standards of quality (QRIS National Learning Network, 2009, p. 2).” Maryland’s QRIS aims to “Promote a high level of program quality through application of standards for program accreditation and environment rating scales” (QRIS National Learning Network, 2009, p. 2). These quotes illustrate that even within QRIS program, states have focused on different issues related to improving quality. Some states, such as Delaware and Kentucky, explicitly state that all young children should have access to high-quality care as a result of their QRIS (QRIS National Learning Network, 2009). Other states direct focus on raising standards within existing ECEC and encouraging parents to choose these programs. Despite stating different targets, all states hope that QRIS programs will raise levels of quality and benefit families and young children.

Although this is an important and admirable goal, the current structure of QRIS programs do not adequately address the fact that families have unequal resources and access when seeking childcare. A central tenant of QRIS programs is parent education, based on the principle that if parents are made aware of high-quality programs they will shift demand to these programs (Jeon et al., 2014; Sabol & Pianta, 2014; Tarrant & Huerta, 2014). Policies built on parent choice assume that all parents have the resources to demand quality, ignoring the fact that low-income families have the smallest amount

of resources and therefore the least amount of voice in the market (Shlay, Trab, Weinrab, & Harmon, 2005). QRIS programs also fail to address the barriers parents face when seeking care, factors such as location, hours, and cost. Although all parents are must balance practical considerations when seeking care, research shows that low-income parents, mothers that work full-time, and mothers who report higher stress levels are the groups most constrained by practical factors (Peyton, Jacobs, O'Brien, & Roy, 2001). This is largely because these families have fewer access to resources — such as time, money, or transportation — that constrain a parent's ability to access high-quality care (Kim & Fram, 2009; Peyton et al., 2001). QRIS assume that increasing the number of quality early childhood programs will “*raise all boats* through competition” (Nagasawa et al., 2014, p. 284). However, if these quality programs are not affordable, in an accessible location, or open when parents need care, then parents do not realistically have a choice to enroll their child in high quality care (Leslie, Ettenson, & Cumsille, 2000; Shlay et al., 2005).

Additionally, there is evidence that families have differing access to highly-rated early childhood programs depending on their racial and class status. Lower (2010) compared the quality “star” ratings of early childhood programs in North Carolina to the demographic information of surrounding neighborhoods. She found that centers with high star ratings, indicating that a program provided high-quality care, were more concentrated in predominantly Caucasian and high-income communities (Lower, 2010). There were much fewer centers with high star ratings in communities that were predominantly African-American, had higher rates of poverty, and had higher rates of

welfare recipients (Lower, 2010). Lower’s study suggests that families who are the most constrained by practical factors, such as location, have the least amount of access to centers with high-quality QRIS ratings (Kim & Fram, 2009; Lower, 2010; Peyton et al., 2001). Unless QRIS target programs that serve low-income children or address issues related to access — including practical considerations such as location, hours, and cost — there is little evidence that these programs will benefit low-income children and families.

III.II The Ability of Early Childhood Programs to Access QRIS

Aside from parents’ ability to access highly rated programs, there are also concerns about the ability of early childhood programs to access QRIS (Nagasawa et al., 2014; Tarrant & Huerta, 2014). Many states require that programs apply to participate (National Center on Child Care Quality Improvement), and there is often little to no data on how programs are selected (Nagasawa et al., 2014). Tarrant & Huerta (2014) write that early childhood programs in Colorado must pay Qualistar, a private organization, to be rated. Urban centers are typically eligible for state or local improvement grants that cover the cost, but this cost is typically too high for rural early childhood programs in Colorado (Tarrant & Huerta, 2014). If early childhood programs do not have access to QRIS programs, they potentially lose out on improvement grants, higher child care subsidies, technical assistance, and other benefits of participating in a QRIS. Although there is currently not enough data to analyze which early childhood programs have better access to participate in QRIS programs, this could have important implications as to which families can participate in highly rated early childhood programs.

III.III The Importance of Issues of Access in a High-Stakes Context

Within quality rating systems, there is evidence that families may be limited in their ability to access highly rated programs and that programs may face limitations to participating in QRIS (Lower, 2010; Tarrant & Huerta, 2014). Issues of program access are critical within states that have QRIS, to ensure families and care providers have equal opportunity to benefit from financial incentives. If programs with higher star ratings are consistently concentrated in affluent and Caucasian communities, similar to the findings in Lower's (2010) research in North Carolina, then higher incentive grants, tax credits, and other forms of financial assistance are going into early childhood programs in affluent and Caucasian communities. Additionally, there is currently a shortage of qualified early childhood education teachers in many states (U.S Department of Education, Office of Postsecondary Education, 2014). As teacher bonuses and wage incentives are tied to a program's rating, qualified teachers are incentivized to seek out highly rated programs, which are more likely to be concentrated in affluent communities (Lower, 2010). Instead of benefiting low-income children and improving student achievement, QRIS programs risk diverting funds and qualified teachers to high-income communities. Although limited, current evidence suggests that QRIS programs may be further perpetuating existing inequalities in care through a tiered rating system that does not allow all families to access highly rated care.

III.IV Current Concepts of Quality and Quality Measurement Related to Issues of Access

Current discourses on quality and quality measurement in QRIS do not account for issues of access, nor do they increase access to quality care. Quality assessment such as the ECERS-R do not measure parents' perceptions of quality, nor do they measure whether early childhood programs meet the needs of parents (Shlay et al., 2005). Shlay et al. (2005) point out that parents may have differing views of quality, valuing program attributes such as flexible hours as an important part of seeking care. Additionally, many standards on assessments such as the ECERS-R focus on the materials available in the classroom (Fenech, 2011; Le et al., 2014; Mathers, Linskey, Seddon, & Sylva, 2007; Perlman et al., 2004; Sakai, Whitebook, Wishard, & Howes, 2003; Sylva, Siraj-Blatchford, Taggart, Sammons, Melhuish, Elliot, & Totsika, 2006; Tarrant & Huerta, 2014). For example, in order to receive a score of "seven" on the "Dramatic Play" item, a classroom must have a variety of prop boxes that are rotated throughout the year (Harms, Clifford, & Cryer, 2005, p. 48). Although QRIS often give early childhood programs additional funds to purchase materials (Tarrant & Huerta, 2014), it stands to reason that centers who are able to charge higher fees are in a better position to purchase more materials. Early childhood programs may also be pressured to purchase more materials, raising the cost of attendance for families. There is currently no research on these effects within a QRIS system, so it is unclear whether centers are increasing rates to improve performance on standardized assessments. However, Layzer & Goodson (2006) argue

that policies aimed at increasing quality also have the secondary effects of increasing cost and decreasing access, writing:

Most activities aimed at increasing child care quality have the concomitant effect of raising the cost of that care. When parents pay for care, the result may be to place some types of child care beyond their reach. When public funds pay for child care, fewer children can be served at a fixed level of resources (p. 573).

QRIS could have a similar outcome, by incentivizing early childhood programs to purchase more materials for a higher score. Early childhood programs in states like Colorado, who must pay to be evaluated by the QRIS, may also face additional costs for participation (Tarrant & Huerta, 2014). States implementing QRIS must assess whether the system increases the cost of child care (Child Trends, 2010). If advocates of QRIS programs want to increase the number of children in high-quality care, they also need to include provisions or instruments that measure the accessibility of quality care.

Policymakers should also focus on increasing the amount of subsidies available to young children and families who are currently unable to afford quality care.

III.V QRIS Programs and Improving Children's School Performance

A second goal of QRIS is to improve the quality of early childhood programs in order to better the outcomes of children in these programs (Hestenes et al., 2014; Kaurez & Thorman, 2011). “Child outcomes” is typically not defined in policy briefs or literature on QRIS (see Child Trends, 2010; Kaurez & Thorman, 2011; QRIS National Learning Network, 2009; Stoney, 2006; Tout et al., 2010 for examples) although stated goals have included improving school readiness and — in the case of Maryland’s QRIS goals — “promot[ing] a well-qualified workforce” (QRIS National Learning Network, 2009, p. 1).

Proponents of QRIS often point to the “achievement gap” between low-income and minority children compared to high-income and Caucasian children as compelling reason to improve the quality of care (Kaurez & Thorman, 2011). By increasing the quality of care and allowing “families, ECE funders, and policymakers . . . to use ratings to rationalize and direct their early care and education choices and investments”, proponents argue the outcomes of low-income and minority children will also improve (Kaurez & Thorman, 2011, p. 7).

Despite these assertions, there is a scarcity of evidence that proves a higher rated center within a QRIS significantly improves child outcomes (Child Trends, 2010; Duncan et al., 2014). Several studies have attempted to validate QRIS and found mixed results on the relationship between QRIS ratings and child outcomes. Thornburg, Mayfield, Hawks, & Fuger (2009) analyzed a pilot program in Missouri found significant positive associations between a QRIS rating and child outcomes in vocabulary skills, social and behavioral measures, and improved self-regulation and self-control abilities. In spite of these findings, Missouri has not implemented a state-wide QRIS program and is currently the only state not planning, piloting, or implementing some version of a QRIS policy (National Center on Child Care Quality Improvement). Zellman, Perlman, Le, & Setodji (2008) evaluated Colorado’s Qualistar program and found there were few significant relationships between a program’s rating and the developmental outcomes of the children in the program. Hestenes et al. (2014) measured children’s social emotional outcomes using standardized instruments and analyzed the data against their program’s star rating in North Carolina. They found children who attended four- and five-star centers had a

significant difference in outcomes over children who attended one-, two-, and three-star center (Hestenes et al., 2014). However, Hestenes et al. (2014) found no significant differences in outcomes between children who attended four- and -five centers, nor between children who attended one-, two-, and three-star centers. This data suggests that North Carolina's five-star ratings do not represent distinct tiers of quality (Hestenes et al., 2014). Sabol & Pianta (2014) found that children in three- and four-star centers in Virginia's QRIS had higher relative gains in math and literacy scores compared to children in two-star centers. However, these differences between scores were not present after the summer, leading Sabol & Pianta (2014) to conclude that centers with higher quality ratings have short-lived relative gains. Taken together, this research suggests that QRIS scores are an inconsistent predictor of child outcomes. The findings from these studies also raise questions about one of the central goals of QRIS: Using funding and financial incentives to improve quality scores will result in higher child outcomes for children in care.

III.VI Quality Measurements in QRIS and Their Relation to Child Outcomes

The methods used to define and assess quality in QRIS could explain why higher ratings are not consistently correlated with improved child outcomes. Research on the quality measurements used in QRIS, particularly the ECERS-R, have not consistently found a correlation between score and child outcomes (Duncan et al., 2014). Burchinal, Kainz and Cai (2011) analyzed the relationship between the ECERS-R score and the outcomes of low-income children across long-term data sets, such as the Cost, Quality

and Outcomes Study (1995), and found a significant but modest correlation. Sabol & Pianta (2014) found in their Early Childhood Longitudinal Study that the ECERS-R score was not a consistent predictor of children's outcomes. Bryant et al. (2003) also analyzed the ECERS-R score against teacher's reporting of children's social skills and found that after controlling for gender, ethnicity, and poverty that there was not a significant relationship between the two factors. The majority of states use the Environmental Rating Scales, including the ECERS-R, to define and assess quality in QRIS (Duncan et al., 2014; Gordon et al., 2013; National Center on Child Care Quality Improvement). Yet, research suggests that scores on the ECERS-R are not a significant predictor of children's cognitive or social-emotional outcomes on standardized achievement tests (Duncan et al., 2014).

In order to meet the stated goal of improving child outcomes, QRIS systems should reconceptualize quality so that it does not rely on scoring incrementally better on a quality assessment. Layzer & Goodson (2006) argue that the link between small environmental changes in a classroom and child outcomes has frequently been overstated. Research on quality and child outcomes typically compares child outcomes in low-quality care with child outcomes in high-quality care; this research is not generalizable toward taking a medium-quality classroom and making it higher quality (Layzer & Goodson, 2006). Therefore scoring a point or two higher on the ECERS-R could increase on a center's star rating while having no significant impact on a child's outcomes.

Additionally, QRIS rely on global quality measurements that may provide an incomplete picture of the processes that contribute to child outcomes. A global quality

scale is named such because it measures both structural and process elements (Cassidy, Hestenes, Hansen, Hegde, Shim, & Hestenes, 2010; Sakai et al., 2003). Structural indicators are environmental factors that are believed to provide the foundation for quality in early childhood, such as teacher credentials, group size, or teacher-child ratios, while process indicators are considered to be the experiences a child has in classroom, including interactions with teachers, peers, and learning activities (Cassidy et al., 2010). Early childhood care policy has typically focused on regulating structural quality factors, mainly because these inputs are able to be quantifiable and easily observed (Scarr et al., 1994). Quality measurement scales such as the ECERS-R also include more emphasis on structural quality factors — such as the materials in the classroom — over the experiences of young children (Cassidy et al., 2010). Despite this, process quality factors are more significantly correlated with child outcomes than structural quality factors (Cassidy et al., 2010; Clifford, Rezska, & Rossbach, 2010; Le et al., 2014). Global quality assessments assume that by combining structural and process quality elements, they are able to provide a complete picture of quality factors in a classroom and are a more reliable predictor of children’s outcomes. However, Cassidy et al. (2010) point out that while we assume global quality is comprised of structural and process elements, this has not been empirically proven through research or theory. Child learning and outcomes are related to a variety of different interactions and experiences in the early childhood classroom, a scale limited to measuring process and structural elements may not be adequately capturing the entire classroom environment.

Finally, concepts of quality in QRIS and quality measurements may be inadequate predictors of child outcomes because they measure the overall classroom and not individual experiences. Layzer & Goodson (2006) point out that the ECERS-R is a blunt measure that attempts to measure the quality of the overall classroom environment, which this does not easily translate into predicting individual child outcomes (Layzer & Goodson, 2006). Even within the same classroom environment, children have very different experiences within the same classroom (Layzer & Goodson, 2006). For example, one child in a classroom could receive extensive attention from the teacher and as a result have better language and literacy outcomes. Other children in the classroom could receive little to no attention from the teacher, have lower outcomes, and this discrepancy would not be noted on the scales (Layzer & Goodson, 2006). Despite different experiences and outcomes, both children would be cared for in the classroom and their outcomes would be correlated with the same ECERS-R score. QRIS and their quality measurements focus on the overall attributes of the classroom, but do not measure the experiences of individual children in care that affect school performance and child outcomes.

These critiques and examples illustrate why concepts of quality and quality measurements within QRIS provide a limited definition of quality that does not promote access to quality care or improve children's cognitive and social-emotional outcomes. This next section of this paper will provide background on the ECERS-R scale in order to give an example of how QRIS assessments measure and define quality. The following section will analyze critical research and discuss how a research scale such as the

ECERS-R is not appropriate for a high-stakes measure of quality. Through this investigation, the paper will examine dominant constructs of quality in early childhood policy and argue for an alternative, flexible concept of quality that would better increase access to care for families and improve child outcomes.

IV. Concepts of Quality in the Early Childhood Environment Rating Scale- Revised

The ECERS-R is a standardized assessment tool that measures the quality of a classroom instructing children ages two-and-half through five years old. It is currently the most used quality assessment in early childhood research and QRIS policy (Duncan et al., 2014; Fenech, 2011; Gordon et al, 2013; Hofer, 2010; National Center of Child Care Quality Improvement; Perlman et al., 2004) This section will provide background on the ECERS-R and its development, the assertions it makes about quality, and its scoring methods in order to provide a concrete example of how quality is typically defined and measured within QRIS.

IV.I The Development of the ECERS-R

The original ECERS, the Early Childhood Environmental Rating Scale, was developed in 1980 by Thelma Harms and Richard Clifford (Harms & Clifford, 1983). Both are early childhood education researchers with the Franklin Porter Graham (FPG) Child Development Institute, an organization associated with the University of North Carolina-Chapel Hill. According to a newsletter by the FPG Child Development Institute (2003): “The scale was based on a checklist of items for improving the quality of environments in early childhood classrooms that Harms had compiled during her nearly twenty years of teaching and observation” (pg. 9). Harms & Clifford (1983) also wrote that the scale was developed through research, observing quality early childhood programs, and “extensive input from classroom practitioners and supervision staff” (p.

262). In order to test the ECERS for validity, the authors also assembled a panel of seven early childhood experts to review the scales and rate them based on importance (Harms & Clifford, 1983).

In 1998, the ECERS was revised in order to “incorporate the advances in our own understanding of how to measure quality” (Harms et al., 2005, p. 1). During this process, Deborah Cryer was added as third author to the scale and several new items were included. Revisions to the scale focused on two “emerging” issues in early childhood education and care: the inclusion of children with special needs and “sensitivity to cultural diversity” (Harms et al., 2005, pg. 1). These revisions were written based on feedback from three focus groups with experts on special needs inclusion and cultural diversity in early childhood education (Harms et al., 2005). Additionally, the authors also based revisions on feedback sessions and questionnaires with early childhood researchers and programs that used the ECERS (Harms et al., 2005). After the revision process, the ECERS-R was published into 1998 and has remained unrevised until today.

IV.II Current Uses of the ECERS-R in Research and Policy

The ECERS and ECERS-R are primarily used for early childhood research and for program evaluation and improvement (Clifford et al., 2010). Harms & Clifford (1983) wrote that before the ECERS was developed in 1980, assessing early childhood program quality in research was a challenge and that researchers needed an efficient and practical instrument to use. In their definition, assessing the quality of an early childhood environment included noting the physical environment experienced by adults and

children as well as the interactions between program staff and children (Harms & Clifford, 1983). In addition to its current use as a quality assessment in QRIS, the ECERS-R has also been used as a self-assessment and improvement tool for early childhood programs and as an alternative to standardized testing students in public pre-kindergarten (see Hooks, Scott-Little, Marshall, & Brown, 2006; Warash, Markstrom, & Lucci, 2005). The ECERS-R is also used internationally; the scale has been translated into many different languages and the authors have documented its use in Canada, Germany, Italy, Sweden, Russia, Iceland, Portugal, England, Spain, Austria, Singapore, Korea, Hungary and Greece (Clifford et al., 2010).

Since its development in 1980 and revision in 1998, the ECERS and ECERS-R have played an important role in early childhood research and policy. When the ECERS was created, it quickly became the field standard for early childhood quality research (Gordon et al., 2013). The ECERS-R is considered both comprehensive and durable, two of the reasons that it is one of the most used measures of classroom quality (Perlman et al., 2004). Fenech (2011) reviewed research about quality in early childhood education and care from 1980 to 2008 and found that the ECERS or ECERS-R instrument was the most frequently used measure of classroom quality; in her sample 47.5% of all research articles used the scale to measure quality. The ECERS-R is also considered an anchor scale, and quality measurement instruments that correlate to the scale are more likely to be used in early childhood research and policy (Hofer, 2010).

IV.III Assessing Quality Using the ECERS-R Instrument

The ECERS-R measurement scale specifies that it should be completed by a neutral observer during a three-hour or longer classroom observation (Harms et al., 2005). A neutral observer is typically a researcher or assessor outside of the program that has been trained on the use of the instrument, although a program director or staff member may complete the instrument in the case of a self-assessment (Harms et al., 2005). The observer uses the ECERS-R to rate classroom on forty-three items organized into seven subscales: Space and Furnishings, Personal Care Routines, Language-Reasoning, Activities, Interaction, Program Structure and Parents and Staff (Harms et al., 2005). (For a comprehensive chart of all forty-three quality criteria, see Appendix A.) Clifford et al. (2010) write that these subscales were designed to “guide the observer to practically meaningful areas of interest in early childhood classrooms” (p. 3). Each item on the scale is scored from one to seven, with one representing “inadequate” and seven being a score of “excellent” (Harms et al., 2005). For each item, quality standards are provided for the ratings of one, three, five, and seven. Each number has several standards that need to be met in order for a classroom to receive this score; these standards are related to the item but not mutually dependent (Harms et al., 2005).

In order for a classroom to receive a number as a score, it must meet all of the standards under that number (Harms et al., 2005). For example, on item nine, “Greeting/departing”, there are three standards listed to achieve a score of five: “5.1 Each child is greeted individually (Ex. staff say ‘hello’ and use child’s name; use child’s primary language spoken at home to say ‘hello’). 5.2 Pleasant departure (Ex. children not

rushed, hugs and good-byes for everyone). 5.3 Parents greeted warmly by staff” (Harms et al., 2005, p. 22). A classroom must meet all three standards in order to score a five on this item, if one standard is checked as “no” then the classroom automatically receives a lower score and any standards met under higher scores are not counted in the scoring process (see Appendix B for item nine). It is also possible for classrooms to receive an even score on the assessment; a classroom that met at least half, but not all, of the standards for a score of five would be scored a four on the item (Harms et al., 2005). After the observation, there is a brief interview period for the observer to ask specific, pre-determined questions to the classroom teachers about standards they were not able to observe during the assessment; for example, whether teachers receive yearly performance reviews (Harms et al., 2005).

IV.IV The ECERS-R and Its Assertions about Quality

The authors of the ECERS-R make several claims about the nature of quality in early childhood and its ability to measure quality. First, as discussed in the last section, the scale is considered a global measure of quality because it measures both structural elements and process elements, meaning that it captures the total quality level in the classroom (Cassidy et al., 2010). Secondly, the scale focuses on “aspects of preschool environment that are important as viewed by child development experts” (Harms & Clifford, 1983, p. 4). This assumes that early childhood and child development researchers are the foremost experts on quality in ECEC. It also promotes dominant early childhood education discourses on Developmentally Appropriate Practice (DAP) as the

best way for children to learn in an early childhood setting (Layzer & Goodson, 2006; Sakai, 2003) For example, DAP contends that children learn best when playing independently with self-selected materials; this theory is so central to the ECERS-R scales that a classroom can receive a lower score on the instrument if teachers select materials for their students (Layzer & Goodson, 2006). The ECERS-R is believed to provide a comprehensive measure of quality that promotes developmentally appropriate practices as identified through child development research (Clifford et al., 2010; Harms & Clifford, 1983; Layzer & Goodson, 2006; Sakai et al., 2003).

The authors of the ECERS-R also assume that the scale's measurements are universally valid across all cultures and early childhood programs. When discussing the international use of the ECERS-R scale, Clifford et al. (2010) write: "In spite of the cultural differences between these [countries], each adheres to a core set of child development goals and early childhood practices that align with those assessed by the Environment Rating Scales" (p. 6). This belief in universal early childhood "best practices" promotes the idea that an observer can measure quality in a classroom without any idea of community context and with a standardized instrument (Lubeck, 1998).

These assumptions contribute to the widespread use of the ECERS-R in U.S. and international research and policy. The ECERS-R claims that a three-hour observation with this instrument can provide a numerical score that is a valid measure of quality in the early childhood classroom (Harms et al., 2005). The next section will analyze critical research on the reliability, validity, and content of the ECERS-R instrument, to point out potential implications with using the scale as a quality measurement in high-stakes

policy. As the most widely used assessment in QRIS policy (Duncan et al., 2014; Gordon et al., 2013; National Center of Child Care Quality Improvement), the ECERS-R has a large impact on the way quality is framed and assessed with current early childhood policy.

V. Critiques of the Early Childhood Environment Rating Scale-Revised

Quality rating and improvement systems typically use quality measurements to evaluate early childhood programs in a high-stakes context, meaning that teacher bonuses, improvement grants, state tax credits, and other financial incentives are largely influenced by a program's score on these standardized assessments (Hofer, 2010; Le et al., 2014; Stoney, 2012; Tarrant & Huerta, 2014). As the ERS and ECERS-R are the most widely used in QRIS policy, researchers have begun to critically examine the scale. Gordon et al. (2013) point out that for an assessment as widely used as the ECERS-R, it is surprising that so little is known about the scale's validity. Paget (2001) notes of the ECERS-R: "Certainly, a measurement tool that contributes so significantly to the quality of programs for young children deserves more sustained empirical support" (as cited in Perlman et al., 2004, p. 399). This section will synthesize researchers' critiques of the scale's reliability, validity, and content and then discuss potential implications of these critiques in light of the scale's use in QRIS policy.

V.I Critiques on the Reliability of the ECERS-R

The ECERS-R claims that it provides a consistently valid measure of quality based on a three-hour observation. However, research has found that incidental factors can have a significant impact on the assessment score. Hofer (2010) administered the ECERS-R assessment to early childhood classrooms and found that scores varied significantly based on the length of the observation period, time of day of the assessment, or month the classroom is assessed — even within the same classroom. Longer

observations corresponded to a lower score on the ECERS-R, while observations performed at the end of the year correlated with higher scores on the ECERS-R. Hofer (2010) points out that unless these assessment schedules are standardized across QRIS assessments, early childhood programs could be unfairly penalized by receiving a lower rating.

Researcher also argue that the ECERS-R's scoring method is not evidence-based and does not accurately measure the differences in quality between classrooms (Gordon et al., 2013; Hofer, 2010; Layzer & Goodson, 2006). As discussed in the previous section, all standards must be present from an item in order for it to receive that score. Within the assessment, scoring progresses (from one to seven) until a quality standard is not met, at which point the scoring stops and no other quality indicators are counted (Harms et al., 2005). Hofer (2010) points out that under this scoring technique, classrooms could receive the same score and yet have very different levels of quality regarding an item. For example, two different classrooms could miss one quality standard under a score of three and both receive a score of two on the item. This score does not capture the fact that one classroom may meet higher quality standards associated with a score of five or seven and the other does not (Hofter, 2010). Under the scoring rules of the assessment, both classrooms are considered to be a comparable level of quality (Gordon et al., 2013; Hofer, 2010). This scoring method is not supported by any data and does accurately demonstrate the quality levels of classrooms that may meet some quality standards but not others (Gordon et al., 2013; Hofer, 2010). The method, called a "stop-rule" or "stop-

scoring” method by researchers, is not able to provide a complete and nuanced picture of quality in an early childhood program (Gordon et al., 2013; Hofer, 2010)

V.II Critiques of Validity Issues Involving the ECERS-R

Researchers have also questioned the validity of the ECERS-R, pointing out that a three hour observation with the assessment is not a valid representation of life in the classroom or its level of global quality (Layzer & Goodson, 2006; Wiltz & Klein, 2001). Layzer & Goodson (2006) points out that observation is completed during a brief period in just one day of a preschool classroom; this time period cannot accurately represent all activity that occurs every day in the classroom, nor can it capture the individual experiences of each child in the classroom.

Wiltz & Klein (2001) used the ECERS-R in their research project to sort classrooms into categories of high or low quality. At the end of their research, they conclude:

While the purpose of this study was not to validate the ECERS or any other assessment measure, it was discovered that none of these measures captured the essence of the daily events of child care. These measures just provided a starting point for the research. The actual dynamic of the complicated contextual setting of child care demanded more study, more time, and more finely-grained research tools. (Wiltz & Klein, 2001, pg. 24)

Wiltz & Klein’s (2001) analysis demonstrates that while the ECERS-R is used to sort and categorize classrooms in early childhood research, additional classroom observation provides a more nuanced and complete picture of the classroom that cannot be adequately captured in a quality measurement scale. Other researchers using the ECERS-R have argued that despite the scale’s reputation as a measure of “global quality,” it is best used

in conjunction with other quality measure scales to provide a more complete picture of quality (Sakai et al., 2003; Sylva et al., 2006).

Another validity concern is whether the tool is sensitive to multiculturalism and diversity issues. So while the ECERS-R was revised to include quality standards about multiculturalism and diversity, there is evidence that the assessment is not a valid measurement of cultural sensitivity in a classroom (Sakai et al., 2003; Sylva et al., 2006). For instance, item number twenty-eight on the scale, entitled “Promoting acceptance of diversity”, focuses on cultural diversity in the classroom (see Appendix B for item twenty-eight). Standards under this item concentrate on materials in the classroom and whether or not they represent people of different races, cultures, ages, abilities, and genders (Harms, Clifford, & Cryer, 2005, pg. 56). However, there is evidence that simply using diverse books, posters, and dolls is not enough to adequately teach young children about cultural sensitivity. Aboud & Levy (2000) found in their research analysis that exposing children to materials with cultural and racial diversity did not have any effect on their attitudes and beliefs about other cultures. Instead, teachers needed to facilitate meaningful discussions with children in order to challenge their ideas and expose them to differing viewpoints (Aboud & Levy, 2000).

Derman-Sparks & Ramsey (2011) also point out that by merely including materials in the classroom, the class “can easily slip into a ‘tourist’ curriculum that is unconnected to the daily life experiences and underlying values of children in the program and people in the larger community” (pg. 115). When examining the standards under the “Promoting acceptance of diversity” item, it appears that they do not require

classrooms to move beyond a “tourist” curriculum approach, suggesting improvements such as having “ethnic clothing,” serving “ethnic foods” as a regular part of meals and snacks, and encouraging parents to “share family customs with children” (Harms, Clifford, & Cryer, 2005, pg. 56). By emphasizing differences or the exotic nature of other cultures, teachers risk “Othering” diverse cultures and races and confirming children’s bias (Derman-Sparks & Ramsey, 2011).

Sakai et al. (2003) also echo concerns that the ECERS-R revision does not do enough to address cultural diversity. In their research with the scale, teachers who spoke the same home language as the children and included their culture in the classroom often scored very low on item twenty-eight, “Promoting acceptance of diversity” (Sakai et al., 2003). Meanwhile, teachers who did not speak the home language of the children in the classroom or practice culturally responsive pedagogy were able to receive a high score on the item. Sakai et al. (2003) concluded that future revisions of the ECERS-R should work to better include children’s culture in the classroom.

While the ECERS-R has been revised to address cultural diversity in the classroom, researchers have shown how the scale focuses primarily on diversity of materials and not on teachers and their cultural sensitivity in the classroom. As the ECERS-R is increasingly being used to define and measure quality, this has important implications for the way cultural diversity is framed and taught in early childhood classrooms throughout the United States. It is important that quality standards move beyond a “tourist approach” to including culturally responsive curricula and pedagogy.

V.III Critiques of Other Content in the ECERS-R

In addition to the above critiques about the scales' validity and reliability, other researchers have questioned the pedagogical assumptions and content of the ECERS-R. One main critique is that the ECERS-R primarily focuses on the classroom environment and materials and does not adequately emphasize teacher interaction or instruction (Fenech, 2011; Le et al., 2014; Mathers et al., 2007; Perlman et al., 2004; Sakai et al., 2003; Sylva et al., 2006; Tarrant & Huerta, 2014). Many of the items — for example item number twenty-four, “Dramatic play” (see Appendix B for item twenty-four) — have standards that deal solely with the materials in the center and do not require teacher interaction in the center in order to receive a high score (Harms et al., 2005, pg. 48; Mathers et al., 2007). This concept of quality reinforces the idea of the child as an individual explorer and learner, but does not provide guidance on how teachers could scaffold, guide, or contribute to learning within centers (Mathers et al., 2007).

For instance, teachers in Colorado participating in a QRIS echoed this sentiment, expressing frustration because they felt that the ECERS-R was primarily evaluating their classroom environment and not their teaching abilities (Tarrant & Huerta, 2014). Layzer & Goodson (2006) believe that these critiques about the ECERS-R stem from a change in ideas about Developmentally Appropriate Practice (DAP); since the revision of the scales in 1998, DAP has shifted to place more emphasis on teacher-directed learning. Sakai et al. (2003) argue that the scale focuses on concrete practices, such as materials and the environment, over teacher interactions and child experiences because those practices are much easier to standardize and measure.

The ECERS-R's focus on materials and the environment has important implications for its use in QRIS policy, particularly because ratings on the scale are tied to grants, tax credits, and teacher bonuses (Hofer, 2010; Le et al., 2014; Stoney, 2012; Tarrant & Huerta, 2014). Researchers worry that in order to improve their scores on the assessment, early childhood staff will focus on improving their classroom materials over improving their teaching and interactions with children (Sakai et al., 2003; Tarrant & Huerta, 2014). There is currently very little research on what kinds of changes teachers to their classroom after participating in a QRIS (see Tarrant & Huerta, 2014). However, the research that has been done provides some evidence to support these concerns.

For example, Hooks et al. (2006) surveyed teachers of public pre-kindergarten and kindergarten classrooms in South Carolina that were being evaluated by the ECERS-R as an accountability measure. Although fifty percent of teachers said they made many changes as a result of the assessment, the most common changes reported were: rearranging centers, improving health and safety practices, adding books and materials, and revising schedules for more center time (Hooks et al., 2006). Tarrant & Huerta (2014) interviewed teachers participating in Qualistar, Colorado's QRIS, and found that teachers who received high quality scores reported largely symbolic compliance with the ECERS-R that focused on environmental changes. One teacher complained that she had to spend a month preparing her classroom for just one day of evaluation and another described the experience as "jump[ing] through hoops" (Tarrant & Huerta, 2014, p. 7). If teachers focus on environmental changes to improve their score, they could inflate their scores without improving the quality of their teaching or interactions with children (Sakai

et al., 2003; Tarrant & Huerta, 2014). Researchers worry about this because under many QRIS programs, a high score on the ECERS-R could reflect excellent classroom materials and mask poor teacher-child interactions (Sakai et al., 2003). In reality, teacher interactions provide an essential foundation to early learning and responsive care (Hamre, 2014).

V.IV Critiques from Teachers in QRIS that use the ECERS-R Assessment

Teachers participating in QRIS have other criticisms of the ECERS-R that have not previously been addressed in research, including the fact that scheduling requirements were too rigid and clashed with their understanding of DAP (see Tarrant & Huerta, 2014). The ECERS-R has eight items with a standard that requires that centers and materials be available for “a substantial portion of the day” (Harms et al., 2005, pg. 52). Essentially, in order to receive a score of a five or above on these eight items, free play in centers or with learning materials must constitute a “substantial portion of the day” — or one third of the school day (Harms et al., 2005, p. 7). Tarrant & Huerta (2014) reported that this one requirement constituted such a large part of a teacher’s score that some teachers used timers during the evaluation in order to make sure that they followed their schedule precisely. This rigid conceptualization of scheduling frustrated many teachers, who wanted the flexibility to adapt their schedule to the needs and behaviors of the students (Tarrant & Huerta, 2014). One public pre-kindergarten teacher added:

I would say the risk of defining something is that you tend to put it into a box . . . As you all know you work with kids and you can have the greatest plan and it can go out the window in five minutes. You have to be able to flow where your kids are and what works in one room might not work in another room. What works

with one child might not with another. (as quoted in Tarrant & Huerta, 2014, pg. 6)

This teacher felt that such a strict schedule became “counterproductive” to student learning in her classroom, as it did not allow her to react to the environment and tailor her teaching to her students’ needs (Tarrant & Huerta, 2014, p. 6). Another teacher echoed this sentiment; explaining that under Qualistar she was less able to adapt the schedule to the children’s desires. For example, when asked to read an extra book during circle time, she felt pressured to move onto the next activity instead of “tak[ing] into account what our kids want” (Tarrant & Huera, 2014, p. 5). Although some teachers appreciated the scheduling requirements in the ECERS-R and used it expand their classroom’s time for free play, others only complied with it during the assessment process in order to receive a higher score (Tarrant & Huerta, 2014). This critique of the ECERS-R scheduling process has not been addressed in early childhood research, but was a dominant theme of interviews with teachers participating in Qualistar (Tarrant & Huerta, 2014). Additional research with early childhood staff participating in QRIS programs could provide other important insights about the ECERS-R and its implications for use in classrooms.

V.V Critique of the ECERS-R and its Use in QRIS Policy

Quality rating and improvement systems were designed to define, assess, and promote quality in early childhood care programs (Hestenes et al., 2014). Within a QRIS, the standardized instrument used to assess programs becomes extremely important; it provides a concept of quality that programs must follow in order to receive a high rating (Perlman et al., 2004; Tarrant & Huerta, 2014). The amount of program funding tied to

QRIS ratings — in the form of subsidy reimbursements, improvement grants, tax credits, and teacher bonuses — also acts as a strong incentive for early childhood programs to comply with the ECERS-R assessment (Hofer, 2010; Le et al., 2014; Stoney, 2012; Tarrant & Huerta, 2014). Because the ECERS-R plays such an important role in many QRIS systems, it is important that the scale is rigorously analyzed and critically examined. Early childhood researchers have questioned the scale's reliability as a consistent measurement of quality (Gordon et al., 2013; Hofer, 2010). Researchers have also challenged the scale's validity, shedding doubt on its ability to accurately measure global quality in an early childhood classroom (Perlman et al., 2004; Sakai et al., 2003, Wiltz & Klein, 2001). Finally, there are arguments that standards in the ECERS-R focus too heavily on classroom materials and do not give enough weight to instruction and teacher interaction (Fenech, 2011; Le et al., 2014; Mathers et al., 2007; Perlman et al., 2004; Sakai et al., 2003; Sylva et al., 2006; Tarrant & Huerta, 2014). Although there is limited research on teacher's response to QRIS and the ECERS-R, the reported research data does suggest that many teachers focus their improvements on materials and schedule changes (Hooks et al., 2006; Tarrant & Huerta, 2014).

This research points to potential shortcomings with using the ECERS-R assessment as a quality measurement within QRIS programs. Available data suggests that QRIS policies may not be improving access to quality ECEC or improving child outcomes (Child Trends, 2010; Duncan et al., 2014; Layzer & Goodson, 2006). This analysis of the ECERS-R, the most widely used assessment in QRIS, illustrates how quality is typically defined and measured within these programs (Duncan et al., 2014;

Gordon et al., 2013; National Center of Child Care Quality Improvement). It also demonstrates potential problems with the assessment that could have a large impact on the implementation of QRIS policies. This next section will deconstruct dominant discourses on quality, including those that contribute to quality measurements such as the ECERS-R, in order to advocate for a reconceptualization of quality that could lead to more responsive early childhood policy.

VI. Framing Current Concepts of Quality and Quality Measurement

When researcher Arthur Emlen (1999) began surveying parents on their opinions of child care quality for the Oregon Child Care Research Partnership, he reported encountering resistance from both colleagues and other researchers. According to Emlen (1999), one early childhood professional told him exasperatedly: “We already know what quality of care is! Why ask parents?” (p. 2). This statement reflects problems with the way that “quality” has typically been conceptualized and discussed in early childhood education and care. Because researchers feel they “know what quality of care is,” they do not critically examine dominant discourses or research other stakeholder’s opinions on quality (Emlen, 1999, p. 2; Lubeck, 1998; Tobin, 2005).

This section will argue that the majority of early childhood research uses a positivist framework that conceptualizes quality as an objective, universal experience (Dahlberg et al., 1999; Fenech, 2011; Lubeck, 1998; Tobin, 2005). This section will then deconstruct ideas about “best practice” and a uniform concept of quality using arguments by Joseph Tobin (2005), discussing how these discourses disregard the perspective of key stakeholders. Finally, this section will use three works to promote an alternative concept of quality that could be used in early childhood policy, a concept that uses multiple stakeholder perspectives to create conversations about quality specific to a program’s community context.

VI.I Current Conceptualization of Quality in Early Childhood Discourses and Research

Early childhood research often uses a positivist framework and positions quality as an objective, universal concept defined by early childhood researchers (Ceglowski, 2004; Ceglowski & Bacigalupa, 2002; Dahlberg et al., 1999; Fenech, 2011). Positivism is the belief that there exists a universal, objective reality, and that through technically and scientifically rigorous methods it is possible to know more about this reality (Dahlberg et al., 1999; Schwandt, 1996). Within early childhood quality discourses, this positivist framework manifests itself in the belief that an objective quality is “out there” and can be discovered and defined by early childhood researchers (Dahlberg et al., 1999, p. 92).

Dahlberg et al. (1999) write that developmental psychology has had a significant impact on early childhood education, contributing to a widespread positivist framework in the field. Just as developmental psychology makes universal claims about child development, early childhood education research often contends there is a “best practice” when caring for and teaching young children, based on empirically proven methods (Lubeck, 1998; Tobin, 2005). These claims are the basis for developmentally appropriate practice (DAP) and become entrenched in discourses about quality and codified into quality measurements such as the ECERS-R (Layzer & Goodson, 2006). Lubeck (1998) writes: “[The DAP] guidelines present research results as reflecting accurate, up-to-the-minute results, pointing us irrevocably to one conclusion. The language used makes it appear that these results are incontrovertible” (pg. 287). Through their positivist framing,

quality discourses in early childhood education are seen both as the universal “truth” and as the only correct method to educate young children.

However, contrary to its positivist framing, developmental psychology and neurological research does not directly translate into empirically proven best educational practices. As Tobin (2005) writes:

the claim that quality standards follow in some simple and direct and value-free way from scientific discovery is spurious . . . Knowing how children develop does not automatically suggest any particular best practice, any particular student/teacher ratio, any particular approach to dealing with misbehavior, or any particular strategy for serving children of recent immigrants” (p. 426).

Lubeck (1998) and Tobin (2005) reject the idea that ECEC best practices is based on a rational application of developmental psychology, instead arguing that conceptualizing quality largely depends on one’s beliefs about young children. The next subsection will use examples of early childhood education in other cultures to disrupt the idea that quality is objective and universal, instead arguing that measuring quality in ECEC is a philosophical act based on personal and cultural beliefs about childhood (Dahlberg et al., 1999; Tobin, 2005).

VI.II Disrupting Universal and Objective Claims about Early Childhood Quality

Early childhood researchers and quality measurements assume quality can be objectively measured and that an incrementally higher score on the measurement represents better care and outcomes for children (Hestenes et al., 2014). In reality, quality assessments rely on value judgements about young children; assessing an “appropriate” interaction or “safe” play area depends on researchers’ perspectives of how children

should be spoken to or the amount of risk children should be allowed to take (Tobin, 2005). These beliefs vary between parent, communities, and cultures; subscribing to dominant discourses about child development and early childhood education does not guarantee the best method to care for and educate young children. Tobin (2005) uses the examples of French and Japanese preschools to point out how quality standards are largely cultural and do not necessarily represent the best way to care for young children. Japanese preschools have ratios of one adult to thirty children, typically do not intervene when children fight on the playground, and allow older preschool children to help care for infants and toddlers (Tobin, 2005). Japanese preschools do not follow current U.S. theories about quality care and:

If our theories of quality in early childhood education are correct, the ratios, class size, non-intervention in fights and other factors of Japanese early childhood practices should produce children who leave preschool socially, linguistically, emotionally, and cognitively impaired. But by every appearance, Japanese children love their preschools, and they seem to move on to first-grade at least as ready for success as do American children (Tobin, 2005, p. 164).

Tobin's (2005) analysis points out that if DAP's assertions are correct, then there should be only method to best prepare young children for school. However, if Japanese preschools can "break almost every National Association of Early Childhood Education (NAEYC) quality standard" and still produce socially and cognitively competent children, then we must examine the possibility that there are other ways to provide quality and responsive care to young children (Tobin, 2005, p. 422). Similarly, French preschools have ratios of one teacher per twenty-five students, reject play-based learning in favor of teacher-directed instruction, and yet ninety-five percent of French children

still attend the voluntary state-sponsored preschools (Tobin, 2005). These two examples demonstrate that dominant concepts of quality and quality measurements in the United States are not absolute nor do they represent the only way to adequately care for young children.

Tobin (2005) uses the examples of French and Japanese preschoolers for argue for cultural relativism in the early childhood quality debate. Cultural relativism contends both that one culture's beliefs cannot be considered universal truths and that "the beliefs and practices of a culture cannot be meaningfully evaluated using the criteria of another culture" (Tobin, 2005, p. 425). This theory can be used to argue against international use of quality measurement scales, it can also be applied to standardized assessments within QRIS programs. Within the United States, there are multiple cultural and community-specific viewpoints on quality and young children (Tobin, 2005). Using the theory of cultural relativism, a standardized assessment developed in one community and cultural context is not a valid measurement across all cultures and community in the United States (Tobin, 2005). Dalhberg et al. (1999) also point out that because there are multiple cultural definitions of quality then quality measurement tools privilege one perspective over others (Dahlberg et al., 1999). Within QRIS policies, this means that financial incentives are tied to one culture's ideas on quality while ignoring others.

VI.III Implications of the Current Conceptualization of Quality in High-Stakes Policy

Current research on quality in early childhood care has important implications for quality measurement scales and early childhood policy. Research discourses often rely on a single concept of quality as the only way to improve care, isolating researcher views from outside critique and discourages research on additional elements that could contribute to quality care (Emlen, 1999; Fenech, 2011). In order to provide evidence for this claim, Fenech (2011) analyzed journal articles published between January 1980 and July 2008 that included the words “quality” and “childcare” or “early childhood education.” She found that 83.7% of the 338 articles analyzed used a positivist paradigm and 87.3% of the articles were quantitative research studies (Fenech, 2011). The articles were overwhelmingly from Western countries and 70.4% of the articles reviewed were conducted by researchers in the United States (Fenech, 2011). Most notable in Fenech’s (2011) research was the dominance of researcher’s perspectives of quality, which was included in almost 90% of the articles in her analysis. Even when journal articles included other stakeholders’ perceptions, over half of these articles used closed response systems (such as surveys) that limited participants to pre-determined questions and answers based on researcher perceptions of quality (Fenech, 2011). Fenech’s (2011) analysis of journal articles provides evidence of the fact that a Western, positivist framing of quality is dominating early childhood research, while parent, children, and teacher opinions of quality are largely restricted or excluded from research discourses. Fenech (2011) writes that what we consider to be quality in early childhood education and care

“may well be confined to what has been measured” in early childhood education research (p. 110).

When researchers have allowed parents, teachers, and children to speak freely about their perceptions of quality, they have expressed ideas in addition not what is currently included in our concepts of quality (Fenech, 2011). For example, research with low-income, African-American families found that parents valued ECEC settings where their children were exposed to economically and racially diverse classrooms (Shlay et al., 2005). These views could be used to create early childhood policy that is more responsive to the needs of families, instead of policy that promotes a single view of quality espoused by a standardized quality measurement (Emlen, 2010; Shlay et al., 2005).

Teachers are not only excluded from the research process and creation of quality measurements, they also do not get a voice in the assessment itself. Quality measurement scales claim that both the instrument and the observer are neutral and objective assessors (Lubeck, 1998). In order to maintain the impression of impartiality, observers have very limited, scripted interactions with teachers (Harms et al., 2005). For example, the ECERS-R guidelines instruct administrators during the observation to “Maintain a pleasant but neutral facial expression. Do not interact with the children unless you see something dangerous that must be handled immediately. Do not talk to or interrupt the staff” (Harms et al., 2005, pg. 5). This pretense of impartiality largely ignores the judgment calls that an observer must make on whether an entity is “insufficient,” “appropriate,” or “ample” (as described in Harms et al., 2005). (Although the scale

provides notes to clarify many of these terms, as displayed in Appendix B, these loose guidelines still rest ultimately on the judgment of the assessor to figure out what is appropriate for young children.) These instructions also restrict the ability of teachers to clarify their teaching, defend their choices, or have any voice in the assessment process. This process of excluding teacher input is endemic throughout early childhood quality measurements. Lubeck (1998) writes:

The current organizational structure tends to shut down conversations. Supervisors, coordinators and evaluators commonly enter classrooms, fill out rating scales or observation protocols, then briefly indicate what was done right and what needs to be improved — all of this without any input from the teachers themselves (p. 3).

Early childhood teachers typically face lower career esteem and less professional status than other teachers (Fenech, 2011). Their opinions are largely excluded from research on early childhood care quality, the creation of assessment scales, and during the observation itself, potentially reinforcing the idea that early childhood are not professionals (Fenech, 2011). Perlman et al. (2004) noted that high-stakes accountability measures incentivize elementary teachers to “teach to the test” at the expense of implementing other curricula. If scores on the ECERS-R are tied to high-stakes measures such as teacher bonuses, there is no reason to think QRIS will have a different effect (Perlman et al., 2004). High-stakes regulation rewards teachers who conform to dominant discourses and have the potential to exclude or eradicate teachers with alternative points of view (Fenech, 2006; Tarrant & Huerta, 2014). These concerns are particularly troubling considering that teachers do not have a voice in quality discourses or during the observation itself.

VI.IV An Alternative Model of Quality in Early Childhood Education and Care

Through positivist framing and measurement scales, the majority of QRIS imply that quality is universal, objective and measurable. However, this report aims to disrupt dominant views and reframe “quality” as a subjective concept depending upon context and positionality. This subsection will use three reconceptualizations of quality by early childhood theorists to propose alternative concepts that could guide early childhood policy and research.

Ceglowski (2004) and Ceglowski & Bacigalupa (2002) use the work of Lilian Katz (1993) to propose an alternative model of quality to inform policy (see Appendix C for model). This model incorporates four different stakeholders with differing but valid perceptions of quality: researchers have a top-down perspective that concentrates on program attributes and global quality measures; children have a bottom-up perspective that encompasses their individualized experiences in the classroom; parents have an outside-in perspectives that focuses on how programs respond to parent needs, and teachers have an inside-out perspective that incorporates their relationships with children, family, and other staff (Ceglowski, 2004). According to Ceglowski (2004), Ceglowski & Bacigalupa (2002), and Lilian Katz (1993), the researchers’ perspective has dominated early childhood research on quality, thereby being the basis of early childhood policy. The authors propose that early childhood researchers use Katz’s (1993) theoretical model on quality to guide research that incorporates all four stakeholder perspective, therefore creating more balanced and responsive policy (Ceglowski, 2004; Ceglowski & Bacigalupa, 2002; Lilian Katz, 1993).

Dahlberg et al. (1999) use postmodern perspectives to trouble the idea that quality represents a single, universal reality and instead argue that quality is socially constructed. In their book, *Beyond Quality in Early Childhood Education and Care: Postmodern Perspectives*, the authors argue for a theoretical shift away from quality as an objective, finalized concept to “the adoption of a process of questioning, dialogue, reflection and meaning making” (Dahlberg et al., 1999, p. 16). In their proposal, through “meaning making” early childhood educators would deconstruct their current conceptualizations of young children and early childhood institutions and engage in continual reflection and dialogue about their practices. Dahlberg et al. (1999) believe that through this deconstruction and reflection, early childhood educators will have the opportunity to move towards a richer conceptualization of children that includes children as co-constructors of knowledge in the classroom. It is this reconceptualization of children, and not technical measurements of quality, that Dahlberg et al. (1999) argue will improve early childhood education and care.

Tobin (2005) proposes a similar model of conversation to guide conceptualizations of quality; however, his work stresses the importance of community context and cultural relevance within the conversations. As discussed in subsection VI.II, Tobin believes that the current system of quality standards in the United States imposes dominant cultural beliefs on heterogeneous communities. Instead, Tobin (2005) points to the dominance of constructivism in early childhood education, or the idea that knowledge should be constructed by the child instead of received from the teacher. He argues for a similar approach between parents and teachers that would construct concepts of “quality”

specific to the school and community. Tobin (2005) argues for rejecting the concept of standards — which propose a universal approach to quality — or at the very least including standards that necessitate “a process for involving parents in discussions of best practice, and ‘show[ing] evidence of adapting the standards to the needs and values of the local community’” (p. 434). Tobin (2005) contends that an altered concept of quality should include continued discussion and a commitment by early childhood educators to meet the needs of families within their cultural and community context.

These three reconceptualizations provide an alternate framing of quality that would incorporate ideas outside the dominant perspective. Standardized assessments such as the ECERS-R promote a specific notion of early childhood programs and quality that largely conforms to dominant Western and middle-class ideas (Fenech, 2006; Mathers et al., 2007). As a consequence, there are other ideas about quality that are not being included in the ECERS-R assessment or QRIS policy. Fenech (2011) writes: “What is pertinent here is that the ECERS/ECERS-R was designed in a specific context for a specific purpose. Accordingly, it will illuminate particular aspects of quality, while others remain unnoticed, unexplored, or hidden from view” (p. 8). Through using these assessments, QRIS also promote a limited definition of quality. The next section of this report proposes alternative policies that would reframe quality within QRIS to better align with these reconceptualizations of quality.

VII. Alternative Approaches to Promoting Quality through Early Childhood Policy

As the critical analysis in this paper demonstrates, QRIS programs primarily rely on standardized assessments such as the ECERS-R that promote a positivist conceptualization of quality. This concept of quality that may not improve access to quality care or child outcomes; it is also unlikely to meet the needs of all parents, children, or teachers (Tarrant & Huerta, 2014). Although the majority of states use a tiered rating and reimbursement system based on a standardized assessment, state-based QRIS do not always follow this format (National Center on Child Care Quality Improvement).

A small minority of states use different policy approaches to QRIS; these programs provide other models of how to define and improve quality in early childhood education. The previous section proposes several alternative concepts of quality using culturally relevant and postmodern perspectives. This section will highlight two proposals on improving quality — empowering parent choice and providing for ongoing conversations about quality within a community context — that align with these alternative framework. It will also explore how some states' QRIS are aligned with these proposals. This section will end with a policy recommendation for QRIS, one that could use state funding to improve access to and early childhood care without relying on a pre-chosen standardized assessment to define and assess quality. This policy recommendation aims to increase access to “quality” care — care where parents and children are respected

and valued — while seeking to avoid imposing a rigid and uniform concept of quality on all early childhood programs.

VII.I Empowering Parent Choice

Portland State University researcher Arthur Emlen (2010) argues that child care policy and regulations are hurting parents. State regulations focus on protecting parents and children from low-quality child care, because it is assumed that families may be unable to judge levels of quality (Blau, 2007; Mocan, 2005). However, there is evidence that stringent child care regulations do not increase the quality of care and have negative effects of the childcare market (Blau, 2003, 2007). Researchers have found that more stringent child care regulations result in higher costs for families and a decreased supply of center care (Blau, 2003, 2007; Hotz & Xiao, 2011; Phillips, Lande, & Goldberg, 1990; Rigby et al., 2007). These trends seem particularly true for child care centers in low-income markets (Hotz & Xiao, 2011). Emlen (2010) argues that in order to increase quality in early childhood education and care, states need to reduce barriers to care, and allow focus on policies that promote flexible care options for parents. Emlen's (2010) central argument is that parents are able to choose the best care for their children, and that by decreasing regulation and shifting focus to increasing child care flexibility and parent access, parents will be in a position to choose better care.

Quality rating and improvement systems are typically voluntary; therefore, they may not have the same effects on child care markets (Connors & Morris, 2014; Jeon et al., 2013). However, Emlen's (2011) arguments provide an interesting framework

through which to view parent choice and early childhood education and care policy. QRIS include goals on “parent education” but this education appears to mainly consist of telling parents about the rating system and encouraging parents to choose a child care center with high ratings (Jeon et al., 2014; Sabol & Pianta, 2014; Tarrant & Huerta, 2014). This assumes that the rating systems are meaningful for parents and that a highly rated center will be the best fit for all children and parents (Tarrant & Huerta, 2014). Instead, early childhood care policy could increase focus on parent access and educate parents on all available early childhood programs.

According to the National Center on Child Care Quality Improvement, Utah considers its QRIS program to be a “Quality Recognition and Information System.” Instead of rating early childhood programs, Utah developed a comprehensive website to provide information to parents seeking child care. The website, Care About Childcare, includes information on all programs registered with Utah’s Child Care Resource and Referral Agencies (National Center on Child Care Quality Improvement). Each program has a profile page and providers can choose to apply for different “Quality Indicators” that will be advertised to parents on their profile page (National Center on Child Care Quality Improvement). In addition, the website functions as an important resource for parents that are seeking care. Parents can search for care providers in their zip codes, specify a search radius, and limit the search to programs that have openings for children in a specific age group. Each provider profile page also has extensive information about the program, including center size and hours, licensing citations or complaints, and the educational levels of all center staff. Providers are allowed to add a description or

mission statement and pictures of the program. The profile even states which special needs the child care program is willing to accommodate, for parents of children with special needs. Utah's Care about Childcare website provides extensive information for parents and easy-to-use search options for parents seeking care. The website goes beyond a single rating system and includes information sensitive to parents' needs, including location, hours, and special-needs accommodation. Utah also provides optional quality indicators about the child care environment, program, and staff and lets parents interpret and use that information in their search. By focusing on increasing information, Utah's QRIS allows parents to learn about and access high-quality child care.

In addition to a comprehensive website for parents, QRIS could address other issues related to parent access. Emlen (2010) writes that our definition of quality care should expand to include early childhood programs that provide flexible scheduling for families (see also: Shlay et al., 2005). Parents who work irregular schedules often have trouble finding child care to accommodate their needs (Lower, 2010). Financial incentives could be used to compensate centers that provide flexible scheduling for families.

Additionally, the funding could be used to increase both the amount and number of child care subsidies for low-income families. While the Department of Health and Human Services requires that states "set subsidy rates high enough to ensure that families using subsidies have sufficient access to child care," it defines sufficient access at seventy-five percent of the current market rate (Oliveira, 2006, p. 5). In states that only reimburse three-quarters of the market rate, parents receiving child care subsidies must

pay the difference or find a provider willing to care for the children at a reduced cost. QRIS funding could be used to set child care subsidies at market rates, increasing the purchasing power of low-income families to find high-quality child care. All of these policies will improve parent access and increase the ability of parents to choose high-quality child care for their children.

VII.II Defining “Quality” as Ongoing Conversation Responsive to Community

Context

If QRIS include quality standards and improvement instruments, these standards should be responsive to the needs of the community and families in care (Tobin, 2005). QRIS programs typically use the ECERS-R and other Environmental Rating Scales, standardized assessments developed by three Caucasian early childhood researchers in North Carolina. It is unlikely that these assessments reflect the needs and values of all communities across the United States. Additionally, the ECERS-R and related scales are standardized instruments that do not allow for a flexible view of early childhood care quality. Tobin (2005) writes: “Quality in early childhood education should be a process rather than a product, an ongoing conversation rather than a document” (p. 434). Any assessment of early childhood program quality should allow for critical and continued reflection, so that programs view quality as a ongoing process to meet the needs of children and families in care.

Maine’s QRIS, Quality for ME, allows early childhood programs to select a standardized instrument to be used only as a self-assessment tool (National Center on

Child Care Quality Improvement). To encourage reflective practice, early childhood programs in Quality for ME keep a portfolio with a written improvement plan. Programs then choose an assessment tool (options include Accreditation Guidelines, Head Start Standards, the Environmental Rating Scales, and High Scope, among others) to complete a yearly self-assessment and use the results to update and reflect on their written improvement plan (National Center on Child Care Quality Improvement). Quality for ME allows early childhood programs to choose an assessment that aligns with their own philosophies about early childhood education and care. This is an important distinction from QRIS policy in the majority of states, in which states determine the assessment instrument used (National Center on Child Care Quality Improvement). Quality for ME also uses self-assessment instead of an outside assessor. This encourages reflection and conversation instead of the typical quality assessment process, which “tends to shut down conversations” (Lubeck, 1998, p. 3). Programs can examine their own views about children and respond to specific program and child needs through a chosen self-assessment. Quality for ME encourages quality improvement practices that are ongoing and rooted in the needs of the program and its community. Maine’s QRIS also allows for teacher and staff perspectives of quality to be an important component in the reflection and improvement process, instead of mandating that a specific quality measurement be used to assess early childhood programs.

VII.III QRIS Policy Recommendations

QRIS programs should not rely on a universal standardized assessment, such as the ECERS-R, to define and assess quality for rating systems and financial incentives. These standardized assessments do not increase parent access to quality centers, nor do they correlate to improved child outcomes. Instead, QRIS programs should focus on increasing access and providing technical assistance to early childhood care providers. Increased access will ensure that all families can participate in the available care options, and technical assistants can guide early childhood programs toward a vision of quality that is responsive to the needs of parents and young children. Additionally, QRIS have limited funding, so available programs and funding should target early childhood settings that serve low-income children, where responsive care has been shown to have the largest impact on child outcomes (Shlay et al., 2005).

QRIS should increase parent access by providing comprehensive informational systems about available care providers, such as Utah's website, Care about Childcare. They should use funding to increase child care subsidies for low-income families and link reimbursement levels to current market rates, so that low-income families have the purchasing power to access care that meets their needs. The definition of "quality" early childhood care should also be expanded to include factors that are important to parents. Additionally, states should provide technical assistance to early childhood programs, so that programs can design and implement individualized improvement plans to ensure that they are meeting the needs of parents, children, and the larger community.

These ideas represent a starting point for alternative approaches to the current system of tiered quality rating systems. QRIS typically use a standardized assessment to define and assess quality, yet may not meet their stated goals of increasing access to quality care or improving child outcomes (Child Trends, 2010; Duncan et al., 2014). By reconceptualizing quality in ECEC to include issues of parent access and ongoing conversations about the needs of young children in the learning community, alternative early childhood policy could promote more flexible and responsive early childhood programs.

VIII. Conclusion

This paper examined the structure and goals of quality rating and improvement systems, then argued that QRIS are failing to meet these goals because their quality assessments promote a limited and inadequate measurement of quality. Through critically analyzing the ECERS-R, the most common assessment in QRIS policy, this report demonstrated that there are critiques about the scale's reliability, validity, content, and implementation in QRIS programs. Doing so helps disrupt the dominant discourses on quality conveyed in the scales. The report then offered a reconceptualized understanding of quality to include stakeholder's perspectives and community context. Finally, this report proposed alternative models of ECEC improvement policy that could be used to reframe quality improvement and rating systems.

As QRIS become more widespread, there is evidence that researchers and state agencies are beginning to question use of standardized assessments, particularly the ECERS-R (e.g. Gordon et al., 2013; Hofer, 2010; Layzer & Goodson, 2006; Le et al., 2014; Perlman et al., 2004). Stoney (2012) reviewed states' Race to the Top grant applications and found several states proposed adding assessments in addition to using the Environmental Rating Scales. New Mexico proposed using the ERS as self-assessments to improve teaching instead of an accountability assessment. North Carolina, Delaware, and Kentucky all proposed in their grant application to pilot test a new quality assessment designed specifically for use in QRIS (Stoney, 2012). North Carolina, the second state to adapt QRIS in 1999 and the home state of the ERS, convened a QRIS Advisory Committee in 2012 to analyze its policy. The committee recommended that

North Carolina explore other instruments and: “Continue to use the Environment Rating Scales (ITERS, ECERS-R, SACERS, and FCCERS-R) as a measure of global quality in North Carolina’s QRIS system while further study and development of alternative measures continues” (North Carolina Quality Rating and Improvement System Advisory Committee, 2012, p. 22). Reviewing grant applications is an imperfect research method and does not conclusively prove that states will change their QRIS policy (Stoney, 2012). However, it does suggest that a handful of states are starting to question using the Environmental Rating Scales as their sole instrument to define and measure quality. This questioning represents an opportunity to promote alternative models of quality and reframe the idea of using a standardized measurement as an accountability measure to define and assess quality.

VIII. I Opportunities for Further Research

Additional research is needed to inform QRIS policy and advocate for alternative models to improve policy. In particular, research should analyze QRIS policy and its impact on teachers, parents, and children. Researchers should explore whether parents are using QRIS ratings in their search for quality care, particularly if these ratings are more meaningful for some groups than others. Early research on parent’s use of QRIS has found that many parents are unfamiliar with the policy (Elicker, Ruprecht, Langill, Lewsader, Anderson, & Brizzi, 2013). A statewide survey in Indiana found that only twelve percent of parents had heard of the QRIS program, Pathways to QUALITY (PTQ), and only nineteen percent of families enrolled in PTQ centers were familiar with

the program (Elicker et al., 2013). Other evidence on parents' use of QRIS rating scales has been speculative or anecdotal. In an article in *Education Week*, the director of the QRIS National Learning Network conjectured that parents were not using the rating systems because they felt they were unable to afford centers with high ratings (Blair, 2013). The FPG Child Development Institute (2003) published a newsletter on the success of the ERS, including an article profiling a highly educated couple from Duke University who used the rating system to guide their search for child care. These two sources suggest that families of differing income levels have contrasting reactions to states' rating systems, but the evidence is far from conclusive. Additional research would determine if rating early childhood programs is a helpful resource for parents, particularly parents who have decreased access to quality early childhood care.

There is also limited research on how teachers and early childhood program staff are navigating QRIS and the ECERS-R assessment (Tarrant & Huerta, 2014). Tarrant & Huerta (2014) found in interviews that many Colorado teachers reported largely symbolic adherence to the assessment. There was also a group of teachers that were critical of the assessment, and many teachers who were frustrated when ECERS-R standards clashed with state, Head Start, or other program standards (Tarrant & Huerta, 2014). One teacher even reported having two copies of her class schedule, and switching them out depending on which program was currently evaluating her (Tarrant & Huerta, 2014). She stated in interviews: "I have to keep both [the Colorado Preschool Program] happy because they fund us and I have to keep Qualistar happy because they fund us. So it'd be great if both programs could work together on how long a four year old should play" (teacher at non-

profit center, Tarrant, 2014, p. 7). Aside from Tarrant & Huerta's work (2014), there does not appear to be qualitative data on how teachers are reacting to QRIS programs and their use of the ECERS-R. More evidence could be beneficial to guide early childhood policy, particularly to ensure that QRIS are aligned with quality standards from other programs.

Tarrant & Huerta (2014) also suggests that QRIS could discourage or eliminate teachers with alternative definitions of quality. Early childhood researchers should explore this question and also analyze the effect of QRIS policy on programs with differing concepts of early childhood education. Alternative models, such as Montessori and Reggio Emilia, have philosophies that sometimes differ from the quality standards in the ECERS-R. Qualitative research with Montessori and Reggio Emilia teachers could illuminate how teachers with differing educational philosophies navigate the ECERS-R assessment and participating in QRIS programs. Other qualitative research with teachers could explore whether or not QRIS policy and the ECERS-R dissuades alternate definitions of quality, and how teachers manage these differences.

Finally, research on early childhood quality needs to promote a flexible and context-specific definition of quality that allows for differing cultural views and multiple stakeholders' perspectives. This research could be used to create alternative rating scales and measurements of quality. For example, Emlen (1999) interviewed parents in Oregon on their ideas about quality and incorporated these ideas into a quality measurement scale for the Oregon Child Care Research Partnership. These assessments should not be used to replace existing scales as definitive measurements of quality, but rather as self-assessments to be used by early childhood staff and state technical advisors to guide

reflection and conversations about early childhood quality. Early childhood research on quality has been dominated by positivist, Western ideals that privileges researcher definitions of quality (Fenech, 2011). It is time that researchers allowed for open-ended discussion with teachers, parents, and children, to advance current definitions of quality.

As a growing number of young children are cared for outside of their home (Hotz & Xiao, 2011), early childhood policy needs to ensure that all families have access to positive care that meets the needs of both children and parents. Quality rating and improvement systems attempt to use accountability measures and financial incentives to increase child care quality and children's outcomes, but many systems use a rigid assessment that imposes a specific definition of quality on teachers and families. It is also unclear that QRIS improve access to quality care or better child outcomes. QRIS need to promote better information and access for parents and ongoing conversations about quality that allow for multiple viewpoints. Policies in Utah and Maine provide models on how to develop child care information resources for parents and support individualized improvement plans for early childhood programs. In this way, policy can promote equal access to caring, responsive care for all young children and their families.

IX. Appendices

Appendix A

Subscales and Items in the ECERS-R

Space and Furnishing

1. Indoor Space
2. Furniture for routine care, play and learning
3. Furnishing for relaxation and comfort
4. Room arrangement for play
5. Space for privacy
6. Child-related display
7. Space for gross motor play
8. Gross motor equipment

Personal Care Routines

9. Greeting/departing
10. Meals/snacks
11. Nap/rest
12. Toileting/diapering
13. Health practices
14. Safety practices

Language-Reasoning

15. Books and pictures
16. Encouraging children to communicate
17. Using language to develop reasoning skills
18. Informal use of language

Activities

19. Fine motor
20. Art
21. Music/movement
22. Blocks
23. Sand/water
24. Dramatic Play
25. Nature/science
26. Math/number
27. Use of TV, video, and/or computers
28. Promoting acceptance of diversity

Interaction

29. Supervision of gross motor activities

30. General supervision of children (other than gross motor)
31. Discipline
32. Staff-child interactions
33. Interactions among children

Program Structure

34. Schedule
35. Free play
36. Group time
37. Provisions for children with disabilities

Parents and Staff

38. Provisions for parents
39. Provisions for personal needs of staff
40. Provisions for professional needs of staff
41. Staff interaction and cooperation
42. Supervision and evaluation of staff
43. Opportunities for professional growth

Appendix B

Referenced Items in the ECERS-R

| | Inadequate 1 | 2 | Minimal 3 | 4 | Good 5 | 6 | Excellent 7 |
|-------------------------------|---|---|--------------|--|-----------|---|----------------------|
| PERSONAL CARE ROUTINES | | | | | | | |
| 9. Greeting/departing* | | | | | | | |
| 1.1 | Greeting of children is often neglected. * | | | | | | |
| 1.2 | Departure not well organized. | | | | | | |
| 1.3 | Parents not allowed to bring children into the classroom. | | | | | | |
| 3.1 | | Most children greeted warmly (Ex. staff seem pleased to see children, smile, use pleasant tone of voice). * | | | | | |
| 3.2 | | Departure well organized (Ex. children's things ready to go). | | | | | |
| 3.3 | | Parents allowed to bring children into the classroom. | | | | | |
| 5.1 | | | | Each child is greeted individually (Ex. staff say "hello" and use child's name; use child's primary language spoken at home to say "hello"). * | | | |
| 5.2 | | | | Pleasant departure (Ex. children not rushed, hugs and good-byes for everyone). | | | |
| 5.3 | | | | Parents greeted warmly by staff. * | | | |
| | | | | | | | <i>NA permitted.</i> |
| 7.1 | | | | | | When they arrive, children are helped to become involved in activities, if needed. | |
| 7.2 | | | | | | Children busily involved until departure (Ex. no long waiting without activity; allowed to come to comfortable stopping point in play). | |
| 7.3 | | | | | | Staff use greeting and departure as information sharing time with parents. * | |
| | | | | | | | <i>NA permitted.</i> |

| | Inadequate 1 | 2 | Minimal 3 | 4 | Good 5 | 6 | Excellent 7 |
|---|-----------------|---|--------------|---|-----------|---|----------------|
| 24. Dramatic play* | | | | | | | |
| 1.1 No materials or equipment accessible for dress up or dramatic play. | | | | | | | |
| 3.1 Some dramatic play materials and furniture accessible, so children can act out family roles themselves (Ex. dress-up clothes, housekeeping props, dolls). | | | | | | | |
| 3.2 Materials are accessible for at least 1 hour daily.* | | | | | | | |
| 3.3 Separate storage for dramatic play materials. | | | | | | | |
| 5.1 Many dramatic play materials accessible, including dress-up clothes.* | | | | | | | |
| 5.2 Materials accessible for a substantial portion of the day.* | | | | | | | |
| 5.3 Props for at least two different themes accessible daily (Ex. housekeeping and work).* | | | | | | | |
| 5.4 Dramatic play area clearly defined, with space to play and organized storage.* | | | | | | | |
| 7.1 Materials rotated for a variety of themes (Ex. prop boxes for work, fantasy, and leisure themes). | | | | | | | |
| 7.2 Props provided to represent diversity (Ex. props representing various cultures; equipment used by people with disabilities).* | | | | | | | |
| 7.3 Props provided for active dramatic play outdoors.* | | | | | | | |
| 7.4 Pictures, stories, and trips used to enrich dramatic play. | | | | | | | |

Notes for Clarification

Item 24. Dramatic play is pretending or making believe. This type of play occurs when children act out roles themselves and when they manipulate figures such as small toy people in a dollhouse. Thus, activities used to teach children to follow specific sequences to properly complete household chores, such as table washing or silver polishing activities, are not contained to meet the requirements of this item. Children must be free to use the materials in their own way, as part of their own make-believe play, to get credit for this item.

Dramatic play is enhanced by props that encourage a variety of themes including *housekeeping* (e.g., dolls, child-sized furniture, dress-up, kitchen utensils); *different kinds of work* (e.g., office, construction, farm, store, fire-fighting, transportation); *fantasy* (e.g., animals, dinosaurs, storybook characters); and *leisure* (e.g., camping sports).

3.2. To give credit, the materials must be *accessible* for at least 1 hour daily in programs operating 8 hours or more. Less time is required for programs operating less than 8 hours a day, with the amount of time calculated proportionally (see "Explanation of Terms Used Throughout the Scale" on p. 7 for time required for shorter programs).

5.1. "Many" dramatic play materials means that three or more children can use the materials at one time, without undue competition, and the materials are plentiful enough to encourage more complex play. Dress-up clothes are required as part of the "many" materials, but many examples of dress-up clothes are not required. Hats, purses, and shoes count as dress-up clothes. However, since children are developing gender-role identity during the preschool years, they require concrete examples of dress-ups that are associated with being men or women. Thus, two or three gender-specific examples of dress-up items are required (such as ties, hard hats, or shoes to represent men's clothes; purses or flowery hats for women's). More generic clothing, such as sweatshirts or running shoes, can also be provided, but these do not count as gender-specific dress-ups.

5.2. Consider materials both indoors and outdoors when calculating accessibility for a substantial portion of the day. Dress-up clothes, required in 5.1, are not required for outdoor dramatic play because they might be dangerous. However, props outside must be complete enough to permit meaningful pretend play. For example, an outdoor house must have furniture and other props, doll strollers must have dolls, kitchen furniture must have things to use in a kitchen, child-sized riding cars should have a gas pump or things to transport, cars in the sandbox should have a toy garage or people.

5.3. Consider small toys that children can pretend with, both indoors and outdoors, when scoring this indicator (e.g., small dolls, trucks, animals). (For further discussion about dramatic play themes see *All About the ECEERS-R*, pp. 239–241.)

5.4. Organized storage means that materials of the same type (e.g., dolls, dress-ups, cooking props, food props) are generally stored together (e.g., in containers or in furniture). Storage does not have to be perfectly neat.

7.2. Consider dolls of different races, cultures, ages, and abilities as props for this indicator, as well as dress-up clothes, play foods, and cooking utensils representing different cultures.

7.3. The intent of this indicator is that children are provided a large enough space so that their dramatic play can be very active and noisy without disrupting other activities. A large indoor space such as a gymnasium or multi-purpose room may be substituted for the outdoor space. Structures (such as small houses, cars, or boats) and props for camping, cooking, work, transportation, or dress-up clothes may be available to the children.

Questions

7.1. Are there any other dramatic play props children can use? Please describe them.

7.3. Are props for dramatic play ever used outside or in a larger indoor space?

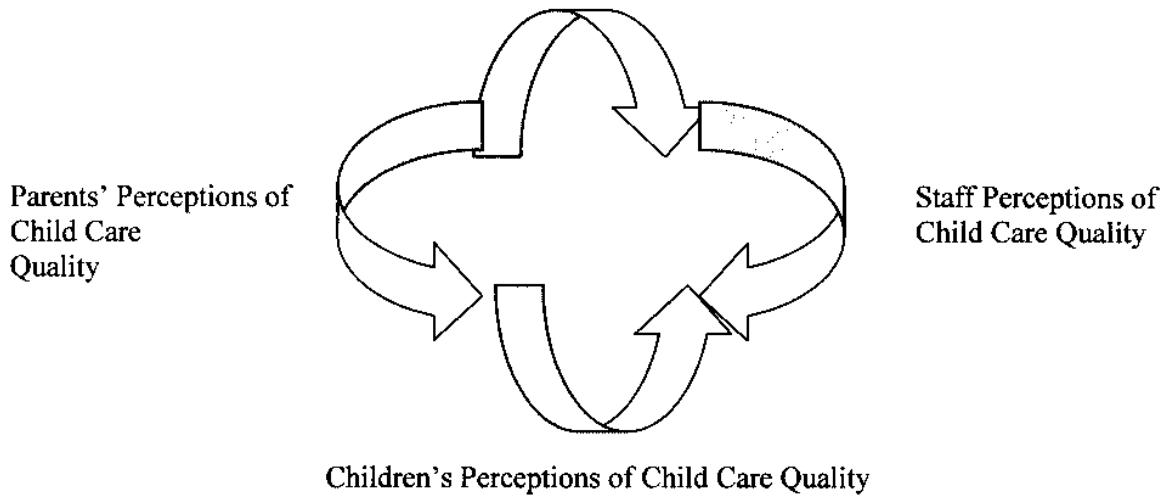
7.4. Is there anything you do to extend children's dramatic play?

| | Inadequate 1 | 2 | Minimal 3 | 4 | Good 5 | 6 | Excellent 7 |
|---|-----------------|---|--------------|---|-----------|---|----------------|
| 28. Promoting acceptance of diversity* | | | | | | | |
| 1.1 No racial or cultural diversity visible in materials (Ex. all toys and pictures are of one race, all print materials are about one culture, all print and audio materials are in one language where bilingualism is prevalent). | | | | | | | |
| 1.2 Materials present only stereotypes of races, cultures, ages, abilities, and gender. | | | | | | | |
| 1.3 Staff demonstrate prejudice against others (Ex. against child or other adult from difference race or cultural group, against person with disability). * | | | | | | | |
| 3.1 Some racial and cultural diversity visible in materials (Ex. multi-racial or multi-cultural dolls, books, or bulletin board pictures, music tapes from many cultures; in bilingual areas some materials accessible in children's primary language). * | | | | | | | |
| 3.2 Materials show diversity (Ex. different races, cultures, ages, abilities, or gender) in a positive way. * | | | | | | | |
| 3.3 Staff intervene appropriately to counteract prejudice shown by children or other adults (Ex. discuss similarities and differences; establish rules for fair treatment of others), <i>or</i> no prejudice is shown. | | | | | | | |
| 5.1 Many books, pictures, and materials accessible showing people of different races, cultures, ages, abilities, and gender in non-stereotyping roles (Ex. both historical and current images; males and females shown doing many different types of work including traditional and non-traditional roles). * | | | | | | | |
| 5.2 Some props representing various cultures included for use in dramatic play (Ex. dolls of different races, ethnic clothing, cooking and eating utensils from various cultural groups). * | | | | | | | |
| 7.1 Inclusion of diversity is part of daily routines and play activities (Ex. ethnic foods are a regular part of meals/snacks; music tapes and songs from different cultures included at music time). | | | | | | | |
| 7.2 Activities included to promote understanding and acceptance of diversity (Ex. parents encouraged to share family customs with children; many cultures represented in holiday celebration). * | | | | | | | |

Appendix C

Alternate Models of Quality

Researcher/ Professional Perceptions of Child Care Quality



Katz's Model on Quality
(as cited in Ceglowski, 2004)

References

- About, F.E., & Levy, S.R. (2000). Interventions to reduce prejudice and discrimination in children and adolescents. In S. Oskamp (Ed.) *Reducing Prejudice and Discrimination* (pp. 269-293). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Blau, D. M. (2003). Do child care regulations affect the child care and labor markets? *Journal of Policy Analysis and Management*, 22, 443–465.
- Blau, D. M. (2007). Unintended consequences of child care regulations. *Labor Economics*, 14, 513–538.
- Blair, J. (2013, June 12). States fold teaching into preschool rating factors. *Education Week*, 31-34.
- Bryant, D., Maxwell, K., Taylor, K., Poe, M., Peisner-Feinberg, E., & Bernier, K. (2003). Smart Start and preschool child care quality in NC: Change over Time and relation to children's readiness. Chapel Hill, NC: FPG Child Development Institute.
- Burchinal, M., Kainz, K., Cai, Y. (2011) How well do our measures of quality predict child outcomes?: A meta-analysis and coordinated analysis of data from large-scale studies of early childhood settings. In M. Zaslow, I. Martinez-Beck, K. Tout, T. Halle (Eds.) *Quality Measurement in Early Childhood Settings*. Baltimore, MD: Brookes Publishing.
- Cassidy, D. J., Hestenes, L. L., Hansen, J. K., Hegde, A., Shim, J., & Hestenes, S. (2005). Revisiting the two faces of child care quality: Structure and process. *Early Education & Development*, 16(4), 505-520.
- Ceglowski, D. (2004). How stakeholder groups define quality in child care. *Early Childhood Education Journal*, 32(2), 101-111.
- Ceglowski, D., & Bacigalupa, C. (2002). Four perspectives on child care quality. *Early Childhood Education Journal*, 30(2), 87-92.
- Clifford, R., Reszka, S. and Rossbach, H. (2010). Reliability and validity of the Early Childhood Environment Rating Scale. Chapel Hill: FPG Child Development Institute, University of North Carolina
- Child Trends. (2010). Quality Rating and Improvement Systems for early care and education. *Early Childhood Highlights*, 1(1) 1–4.
- Cost, Quality, & Child Outcomes Study Team (1995). Cost, quality and child outcomes in child care centers: Public report. Economics Department, University of Colorado at Denver.
- Connors, M. A. & Morris, P. A. (2014). 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*. Advance online publication: <http://dx.doi.org/10.1016/j.ecresq.2014.05.006>

- Dahlberg, G., Moss, P., & Pence, A. R. (1999). *Beyond quality in early childhood education and care: Postmodern perspectives*. London: Falmer Press.
- Derman-Sparks, L., & Ramsey, P. G. (2011). *What if all the kids are white?: Anti-bias multicultural education with young children and families*. New York: Teachers College Press.
- Duncan G. J., Jenkins, J. M., Auger, A., Burchinal, M., Domina, T., & Bitler, M. (2014). Boosting school readiness with preschool curricula. [Manuscript in preparation]
- Elicker, J., Ruprecht, K.M., Langill, C., Lewsader, J., Anderson, T. & Brizzi, M. (2013). Paths to QUALITY™: Collaborative evaluation of a new child care quality rating and improvement system. *Early Education and Development*, 24, 42-62.
- Emlen, A. (1999). *From a parent's point of view: Measuring the quality of child care*. Portland, OR: Portland State University, Regional Research Institute for Human Services.
- Emlen, A. (2010). *Solving the childcare and flexibility puzzle: How working parents make the best feasible choices and what that means for public policy*. Boca Raton, Fla.: Universal.
- Frank Porter Graham Child Development Institute. (2003). A whole new yardstick. *Early Developments, Vol. 7 [Rating early childhood environments]*, 8–11.
- Fenech, M. (2006). The impact of regulatory environments on early childhood professional practice and job satisfaction: A review of conflicting discourses. *Australian Journal of Early Childhood*, 31(2), 49-57.
- Fenech, M. (2011). An analysis of the conceptualization of 'Quality' in early childhood education and care empirical research: Promoting 'blind spots' as foci for future research. *Contemporary Issues in Early Childhood*, 12(2), 102.
- Gordon, R. A., Fujimoto, K., Kaestner, R., Korenman, S., & Abner, K. (2013). An assessment of the validity of the ECERS-R with implications for measures of child care quality and relations to child development. *Developmental Psychology*, 49(1), 146-160.
- Hamre, B. (2014). Teachers' daily interactions with children: An essential ingredient in effective early childhood programs. *Child Development Perspectives*, 8(4), 223-230.
- Harms, T., & Clifford, R. (1983). Assessing preschool environments with the Early Childhood Environment Rating Scale. *Studies in Educational Evaluation*, 8, 261-269
- Harms, T., Clifford, R., & Cryer, D. (2005). *Early Childhood Environment Rating Scale* (Rev. ed.). New York: Teachers College Press.
- Hestenes, L. L., Kintner-Duffy, V., Wang, Y. C., Paro, K. L., Mims, S. U., Crosby, D., Scott-Little, C., & Cassidy, D. J. (2014). Comparisons among quality measures in child care settings: Understanding the use of multiple measures in North Carolina's QRIS and their links to social-emotional development in preschool children. *Early Childhood Research Quarterly*. Advance online publication. <http://dx.doi.org/10.1016/j.ecresq.2014.06.003>

- Heckman, J., Pinto, R., & Savelyev, P. (2012). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 2052-2086.
- Hofer, K. G. (2010). How measurement characteristics can affect ECERS-R scores and program funding. *Contemporary Issues in Early Childhood*, 11(2), 175.
- Hooks, L. M., Scott-Little, C., Marshall, B. J., & Brown, G. (2006). Accountability for quality: One state's experience in improving practice. *Early Childhood Education Journal*, 33(6), 399-403.
- Hotz, V. J., & Xiao, M. (2011). The impact of regulations on the supply and quality of care in child care markets. *American Economic Review*, 101(5), 1775–1805.
- Jeon, L., Buettner, C. K., & Hur, E. (2014). Examining pre-school classroom quality in a statewide Quality Rating and Improvement System. *Child & Youth Care Forum*, 43(4), 469-487.
- Katz, L. (1993). *Multiple perspectives on the quality of early childhood programs*. ERIC Document Reproduction Service No: ED355 041.
- Kim, J., & Fram, M. S. (2009). Profiles of choice: Parents' patterns of priority in child care decision-making. *Early Childhood Research Quarterly*, 24(1), 77-91.
- Kauerz, K., & Thorman, A. (2011). *QRIS and P-3: Creating synergy across systems to close achievement gaps and improve opportunities for young children*. Available at http://www.buildinitiative.org/files/QRIS_P-3brief.pdf
- Layzer, J. I., & Goodson, B. (2006). The "quality" of early care and education settings: Definitional and measurement issues. *Evaluation Review*, 30(5), 556-576.
- Le, V., Schaack, D. D., & Setodji, C. M. (2014). Identifying baseline and ceiling thresholds within the Qualistar Early Learning Quality Rating and Improvement System. *Early Childhood Research Quarterly*. Advance online publication. <http://dx.doi.org/10.1016/j.ecresq.2014.03.003>
- Leslie, L.A., Ettenson, R., & Cumsille, P. (2000). Selecting a child care center: What really matters to parents? *Child & Youth Care Forum*, 29(5), 299-322.
- Love, J. M., Schochet, P. Z., Meckstroth, A. L. (1996). *Are they in any real danger? What research does — and doesn't—tell us about child care quality and children's well-being*. Princeton, NJ: Mathematica Policy Research
- Lower, J. K. (2010). *Socioeconomic disparities in North Carolina communities: Issues of access and quality of licensed child care* (Unpublished doctoral dissertation). University of North Carolina-Greensboro, Greensboro, NC.
- Lubeck, S. (1998). Is Developmentally Appropriate Practice for Everyone? *Childhood Education*, 74(5), 283-292.
- Mathers, S., Linskey, F., Seddon, J., & Sylva, K. (2007). Using quality rating scales for professional development: experiences from the UK. *International Journal of Early Years Education*, 15(3), 261-274.
- Mitchell, A. (2012). *Financial Incentives in Quality Rating and Improvement Systems: Approaches and Effects*. Retrieved from <http://www.qrisnetwork.org/sites/all/files/resources/gscobb/2012-05-24%2015:13/Approaches%20to%20Financial%20Incentives%20in%20QRIS.pdf>

- Mocan, N. (2007). Can consumers detect lemons? An empirical analysis of information asymmetry in the market for child care. *Journal of Population Economics*, 20(4), 743-780.
- Moss, P. (1994). Defining quality: values, stakeholders and processes. In P. Moss & A. Pence (Eds). *Valuing Quality in Early Childhood Services* (pp 1-9). London: Paul Chapman Publishing Ltd.
- Nagasawa, M., Peters, L., & Swadener, B. B. (2014) The costs of putting quality first: Neoliberalism, (ine)quality, (un)affordability, and (in)accessibility? In M. N. Bloch, B. B. Swadener, & G. S. Canella (Eds.), *Reconceptualizing early childhood care & education: a reader : critical questions, new imaginaries and social activism* (pp. 279-290). Bern, Switzerland: Peter Lang International Academic Publishers.
- National Center on Child Care Quality Improvement. (n.d.). *Quality Rating & Improvement System Resource Guide*. Retrieved October 1, 2014, from <https://ocqrsguide.icfwebservices.com/index.cfm>
- Oliveira, P. (2006). The Child Care Subsidy Program Policy and Practice: Connecticut Child Care Providers Identify the Problems. New Haven, CT: Connecticut Voices for Children.
- Perlman, M., Zellman, G. L., & Le, V. (2004). Examining the psychometric properties of the Early Childhood Environment Rating Scale-Revised (ECERS-R). *Early Childhood Research Quarterly*, 19(3), 398-412.
- Peyton, V., Jacobs, A., O'Brien, M., & Roy, C. (2001). Reasons for choosing child care: associations with family factors, quality, and satisfaction. *Early Childhood Research Quarterly*, 16(2), 191-208.
- Phillips, D., Lande, J., & Goldberg, M. (1990). The state of child care regulation: A comparative analysis. *Early Childhood Research Quarterly*, 5(2), 151-179.
- North Carolina Quality Rating and Improvement System Advisory Committee. (2012). The North Carolina Quality Rating and Improvement System (QRIS) Advisory Committee executive summary. Raleigh, NC: North Carolina Division of Child Development and Early Education.
- QRIS National Learning Network. (2009). QRIS Goals and/or Objectives. Retrieved November 29, 2014 from <http://qrisnetwork.org/category/feed/qris-goals>
- Rigby, E., Ryan, R. M., & Brooks-Gunn, J. (2007). Child care quality in different state policy contexts. *Journal of Policy Analysis and Management*, 26(4), 887-908.
- Sabol, T. J., & Pianta, R. C. (2014). Validating Virginia's quality rating and improvement system among state-funded pre-kindergarten programs. *Early Childhood Research Quarterly*. Advance online publication. <http://dx.doi.org/10.1016/j.ecresq.2014.03.004>
- Sakai, L. M., Whitebook, M., Wishard, A., & Howes, C. (2003). Evaluating the Early Childhood Environment Rating Scale (ECERS): Assessing differences between the first and revised editions. *Early Childhood Research Quality*, 18(4), 427-445.

- Scarr, S., Eisenberg, M., & Deater-Deckard, K. (1994). Measurement of quality in child care centers. *Early Childhood Research Quarterly*, 9(2), 131–151.
- Shlay, A. B., Tran, H., Weinraub, M., & Harmon, M. (2005). Teasing apart the child care conundrum: A factorial survey analysis of perceptions of child care quality, fair market price and willingness to pay by low-income, African American parents. *Early Childhood Research Quarterly*, 20(4), 393-416.
- Shonkoff, J. P. & Phillips, D. A. (Eds.). (2000). From neurons to neighborhoods: The science of early childhood development. Washington, DC: National Academy of Sciences.
- Stoney, L. (2012). Unlocking the potential of QRIS: trends and opportunities in the Race to the Top – Early Learning Challenge applications. Retrieved from <http://www.qrisnetwork.org/sites/all/files/resources/gscobb/2012-03-07%2008:29/LouiseStoneyMemo.pdf>
- Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K., & Totsika, V. (2006). Capturing quality in early childhood through Environmental Rating Scales. *Early Childhood Research Quarterly*, 21(1), 76-92.
- Tarrant, K. & Huerta, L. A. (2014). Substantive or symbolic stars: Quality rating and improvement systems through a new institutional lens. *Early Childhood Research Quarterly*. Advance online publication. <http://dx.doi.org/10.1016/j.ecresq.2014.04.00>
- Thornburg, K., Mayfield, W. A., Hawks, J. S., & Fuger, K. L. (2009). The Missouri quality rating system school readiness study. Columbia, MO: Center for Family Policy and Research, University of Missouri Columbia.
- Tobin, J. (2005). Quality in early childhood education: An anthropologist's perspective. *Early Education & Development*, 16(4), 421-434.
- Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G., & Boller, K. (2010). Compendium of quality rating systems and evaluations. Washington, DC: Child Trends.
- U.S. Department of Education, Office of Postsecondary Education. (2014). Teacher Shortage Areas Nationwide Listing 1990-1991 through 2014-2015. Retrieved from <http://www2.ed.gov/about/offices/list/ope/pol/tsa.doc>
- Warash, B. G., Markstrom, C. A., & Lucci, B. (2005). The Early Childhood Environment Rating Scale-Revised as a tool to improve child care centers. *Education*, 126(2), 240-250.
- Wiltz, N. W., & Klein, E. L. (2001). “What do you do in child care?” children’s perceptions of high and low quality classrooms. *Early Childhood Research Quarterly*, 16(2), 209-236.
- Zellman, G. L., Perlman, M., Le, V., 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. Santa Monica, CA: RAND Corporation.