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

Linda L. Hestenes*, Victoria Kintner-Duffy, Yudan Chen Wang, Karen La Paro, Sharon U. Mims, Danielle Crosby, Catherine Scott-Little, Deborah J. Cassidy

Department of Human Development & Family Studies, University of North Carolina at Greensboro, United States

ARTICLE INFO

ABSTRACT

Quality Rating and Improvement Systems (QRIS) include the assessment of classroom quality as one component of how early childhood programs are monitored and licensed in many states across the United States. However, varying measures and foci of quality exist and have led to challenges in accurately depicting program quality across programs and improvement efforts. The current validation study explores several measures of classroom quality and their associations with components and overall star ratings of the North Carolina QRIS and preschool children's social-emotional outcomes within center-based child care programs. Data for this study were collected in 2009, 10 years after the start of North Carolina's QRIS. Results indicate that individual levels of star ratings did not generally represent distinctive levels of classroom quality, but did differentiate classrooms at the lower and higher levels of quality. Structural features of the environments such as teacher education and teacher–child ratio were associated with classroom quality across these measures in the expected directions; however, teacher experience was not. Further, children's social-emotional outcomes were predicted to a varying degree by star levels and different aspects of classroom quality as represented by these various measures. Results are discussed in terms of the differing levels of quality and teaching processes in classrooms. Future directions for research are presented to contribute to an increased understanding of QRIS and children' experiences in early care and education programs.

© 2014 Elsevier Inc. All rights reserved.

Introduction

For the past decade, Quality Rating and Improvement Systems (QRIS) have dramatically changed how early childhood programs are monitored and licensed in most states across the United States. Although there are often multiple dimensions and components to a state's QRIS, one key aspect of these systems is the assessment of quality within classrooms. The underlying belief that the quality of child care environments impacts children's developmental outcomes (Copple & Bredekamp, 2009; NICHD Early Child Care Research Network, 2002; NICHD Early Child Care Research Network, 2010) highlights the importance of studying classroom assessment measures.

Numerous states have used the Environment Rating Scales (the ITERS-R, Harms, Cryer, & Clifford, 2006; ECERS-R, Harms, Clifford, & Cryer, 2005; or SACERS, Harms, Jacobs, & White, 1995) as the primary measures of classroom quality within their accountability and licensing systems (Tout et al., 2010). However, sole reliance on this group of scales may fail to capture important aspects of quality that contribute to children's experiences and outcomes (La Paro, Thomason, Lower, Kintner-Duffy, & Cassidy, 2012; Layzer & Goodson, 2006). As multiple dimensions of classroom environments are likely to matter for children's outcomes, it is necessary to investigate a broader range of instruments and consider how these tools may contribute to overall quality and ultimately to a QRIS. These investigations of multiple measures must include information on how the tools relate to the components and goals of a particular QRIS, as well as assess their association with child

* Corresponding author at: P.O. Box 26170, Department of Human Development and Family Studies, University of North Carolina at Greensboro, Greensboro, NC 27402-6170, United States. Tel.: +1 336 256 0093.
  E-mail address: llhesten@uncg.edu (L.L. Hestenes).

outcomes. The purpose of this study was to contribute to on-going efforts to validate North Carolina's QRIS by examining the associations among a variety of quality assessment tools, and the linkages between these tools, components of the QRIS and overall star ratings, and preschool children's social-emotional outcomes. We first present a framework for validating a QRIS, followed by a review of the literature related to the constructs examined within this validation study and the rationale for the emphasis on social-emotional outcomes, and then a description of the study.

Quality Rating and Improvement Systems

Quality Rating and Improvement Systems (also referred to as Tiered Quality Rating and Improvement Systems) were developed as a comprehensive approach to understanding, assessing, and improving early care and education programs at the state level. The impetus for QRIS arose from a need to improve quality in a systemic manner that recognized programs' progress toward an achievable top tier. A majority of states have established financial incentives along with support systems to foster quality improvement from one tier to the next higher tier (Tout et al., 2010). Furthermore, QRIS enabled families to have more information about the level of quality in which they were enrolling their child and choose higher quality if available and affordable. Indeed, most states provide specific information to parents about how programs earn their rating (Tout et al., 2010).

Virtually every state is currently engaged in some type of QRIS activity, with 37 states implementing a statewide QRIS, 2 states with regional systems, 7 states planning a QRIS, 3 states piloting a system, and 1 state that would require legislation to develop a QRIS (QRIS National Learning Network, 2014). Although each state's system functions in different ways, five components are essential to QRIS: (1) standards of quality such as requirements for teacher qualifications and program characteristics that are used to determine program ratings; (2) accountability measures used to assess quality standards and ensure compliance with state regulations; (3) support for programs to improve quality (e.g., technical assistance or professional development for teachers); (4) subsidy funds and other financial incentives for programs that participate in ratings and/or quality improvement support; and (5) education efforts for families participating in child care programs (Mitchell, 2009; Tout et al., 2010). Across the states, each of these QRIS components is given a different emphasis, attention and allocation of resources, creating significant variation in systems.

It is important to recognize that in this first wave of QRIS, most early adopting states (like North Carolina) did not consider the highest tier as the ultimate in quality care and education in terms of achieving positive outcomes for all children in all settings, but rather the first set of steps to move a system of care and education forward. Ultimately, the hope for these QRIS efforts is that program quality improvements will lead to improved outcomes for children. Yet, the processes by which programs achieve higher quality and quality, in turn, contributes to positive child outcomes are likely dependent on multiple aspects of QRIS, such as the programs participating, and the children's experiences within and outside of the early childhood programs in which they are enrolled. Given that QRIS are complex systems with multiple goals and represent a relatively new policy development, it is important to investigate systematically whether they provide useful and credible information to programs, policy makers, and parents. The goal of this study was both to contribute to North Carolina's effort to validate and strengthen the state's QRIS and to provide information to inform the national discourse on QRIS.

Validation of QRIS

Zellman and Fiene (2012) define validation of a QRIS as "a multi-step process that assesses the degree to which design decisions about *program quality standards and measurement strategies* are resulting in accurate and meaningful ratings" (p. 4). Although a QRIS validation study can address many aspects of the QRIS, the over-arching purpose is to provide data regarding how well the system is functioning and whether the system is meeting its established goals. QRIS are founded on the assumption that the quality of early childhood programs can be measured, that the results from the measurements are credible, and that differences between programs that are measured through the QRIS are accurate and associated with meaningful outcomes. Validation studies provide data to indicate the extent to which these assumptions are supported within the system.

Zellman and Fiene (2012) describe four approaches or processes to validating a QRIS:

(1) Examine the validation of key underlying concepts within the QRIS to evaluate whether the program features addressed by the system are meaningful and valid.
(2) Examine the data collection process and the measures used within the QRIS to test whether they are accurate and psychometrically sound.
(3) Examine the ratings produced by the QRIS to determine if they are assessing quality effectively and are functioning to distinguishing between the quality level of programs in the expected manner.
(4) Examine how the program ratings are associated with child outcomes.

Thus, QRIS validation studies are conducted for different purposes. States that are early in the design phase of their QRIS may conduct validation studies to establish and validate their quality standards. For example, numerous states have conducted expert reviews, focus groups, and literature reviews to validate the constructs addressed within their QRIS early in their planning and implementation process (Zellman & Fiene, 2012). States in the process of implementing a new QRIS can use validation studies to pilot their processes and the measures used to collect data. Three states recently engaged in these types of activities are Kentucky (Isner, Soli, Rothenberg, Moodie, & Tout, 2012), Colorado (Zellman, Perlman, Le, & Setodji, 2008), and Tennessee (Denny, Hallam, & Homer, 2012). Finally, states with fully implemented systems can conduct validation studies to document outcomes of the system and guide revisions or a redesign of the system (Tout & Starr, 2013). For example, in California, validation data were used to model how programs might be distributed across different levels within the QRIS and to inform decisions about cut scores or thresholds (Karoly & Zellman, 2012). Other states have evaluated the link between QRIS quality levels and child outcomes. Among these, Minnesota (Tout et al., 2011) and Indiana (Elicker, Langill, Ruprecht, Lewsader, & Anderson, 2011) have found limited support for such associations, while Missouri's validation study indicated larger gains on measures of social-emotional development for children participating in programs with higher versus lower quality ratings (Thornburg, Mayfield, Hawks, & Fuger, 2009).

Although the types of data collected and the methodology used for each of the four QRIS validation approaches vary from state to state, the underlying purpose is common: To study empirically the components, processes, and outcomes of the QRIS to determine if it is working as intended and has documented positive outcomes for programs, families, and/or children (Tout & Starr, 2013; Zellman & Fiene, 2012). The current study focuses on the third and fourth validation approaches presented by Zellman and Fiene by exploring