

Validation of Quality Rating and Improvement Systems (QRIS): Examples from Four States Executive Summary

Introduction

“Validation of a QRIS is 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” (Zellman & Fiene, 2012, p. 1). In a recent Brief produced through the Quality Initiatives Research and Evaluation Consortium – INQUIRE – Zellman and Fiene (2012) provide a framework for QRIS validation and examples of the activities that could be conducted as part of validation efforts. The current Brief serves as a companion to the 2012 INQUIRE Brief by providing detailed examples and findings from the validation activities in four states: Indiana, Maine, Minnesota and Virginia. The purpose of this Brief is to demonstrate how different states have approached QRIS validation, to compare findings, and to highlight challenges in designing and conducting QRIS validation studies.

Approach 1. Examine the Validity of Key Underlying Concepts: Indiana and Maine

This approach involves examining the elements or concepts that are to be included in program ratings, using empirical and/or expert support.

- **Indiana:** Evaluators referred to the literature and other QRIS reports to review the validity of the QRIS quality standards. The review concluded that 75% of the quality indicators examined had substantial evidence for their validity.
- **Maine:** Evaluators conducted “concept mapping” to determine how parents, providers, and professionals viewed the QRIS standards. They also looked to the literature to support the standards. The concept mapping revealed similarly high ratings of several standards by all raters.

Summary of validation approach 1: Both Indiana and Maine conducted reviews of the concepts to be included in the QRIS ratings. Evaluators in both states found evidence supporting the inclusion of certain concepts in the state QRIS.

Approach 2. Examining the Measurement Strategies Used to Assess Quality: Minnesota and Virginia

The second approach to validation discussed by Zellman and Fiene (2012) is to examine the properties of the measures used to assess quality. This type of validation involves looking at the attributes of individual measures and how they are combined to produce a summary rating of quality. This might include examining the distribution and variance of scores for indicators or looking at the inter-correlations among indicators.

Virginia and Minnesota examined the distribution of indicator scores as a means to determine whether a measure provided enough variability to distinguish levels of quality.

- **Minnesota:** The distributions of scores on QRIS categories and total QRIS rating were varied and generally not normally distributed.
- **Virginia:** Total points and half of the categories (staff qualifications and interactions) were relatively normally distributed. The other two categories (ratio and environment) were less so.
- The skewness found in some indicator scores in both states suggests that those measures may be weak in distinguishing levels of quality in the QRIS.

Virginia and Minnesota also examined the correlations among QRIS indicators.

- **Minnesota and Virginia:** Correlations among indicators ranged from non-significant to moderately high correlations.
- The review of correlations in Minnesota and Virginia revealed that components are related to each other but are not providing duplicative information. The small to moderate correlations provide some confidence that the overall categories of quality measures are contributing unique information to the rating.

Summary of validation approach 2: Both Minnesota and Virginia examined the distributions of scores and the correlations among indicators. They found variation in distributions of scores and small to moderately correlated indicators.

Approach 3. Assess the Outputs of the Rating Process: Indiana, Maine, Minnesota, Virginia

A third approach to QRIS validation involves assessing the actual ratings (the “outputs” of the rating process) to understand the degree to which they are producing levels of quality that are distinct in meaningful ways. The degree to which ratings produced distinct levels of quality were examined across the four states by analyzing how ratings were linked to scores on the Early Childhood Environment Rating Scale – Revised (ECERS-R; Harms, Clifford, & Cryer, 1998), a global measure of preschool classroom quality.

- **Indiana:** Differences in mean ECERS-R scores across QRIS levels were noted. Level 4 programs had higher scores than Levels 1 and 2.
- **Maine:** No significant differences between mean ECERS-R scores across QRIS levels were noted.
- **Minnesota:** Differences in mean ECERS-R scores across QRIS levels were found. Level 2 programs had lower scores than Levels 3 and 4.
- **Virginia:** Differences in mean ECERS-R scores were found across Levels 2, 3, and 4.
- Mean ECERS-R scores fell below the “good” level of quality in all four states.

Summary of validation approach 3: Overall, the cross –state findings indicate that ECERS-R scores can discriminate quality levels when the tool or similar indicators are used in the rating process (as seen in three of the four states). The findings in Maine, showing no correlation between quality levels and ECERS-R scores, provide a cautionary note that quality levels may not be distinct on the ECERS-R if indicators used in the QRIS do not align closely with the measure of environmental quality that is used in the validation process.

Approach 4: Examine How Ratings are Associated with Children’s Development: Indiana, Minnesota, Virginia

The fourth approach to validation examines the association between QRIS ratings and children’s developmental gains. Validation studies using this approach examine whether the QRIS ratings and quality components that comprise the ratings are related in expected ways to measures of children’s development and differences in their patterns of growth. These studies are challenging because they must be conducted with a clear understanding of how the particular QRIS operates and characteristics of the quality components that make up the ratings, methods to account for selection biases (in programs, parents and children), recognition that effect sizes are modest in research examining dimensions of quality and children’s development, and data collection that allows for analysis of children’s gains rather than point-in-time measurements. It is important to review the results with these challenges in mind.

- **Indiana:** No consistent, strong associations between QRIS quality level and young children's development and learning were found. There were some relations between measures of observed quality and child development.
- **Minnesota:** No systematic evidence of strong relations between quality ratings, measures of program quality and children's developmental progress was found.
- **Virginia:** Some evidence was found for an association between QRIS rating and growth in pre-literacy skills in prekindergarten.
- Limitations and implications of these types of analyses are discussed.

Summary of validation approach 4: States are just beginning to examine the relation between QRIS ratings and child development. Examples from Indiana, Minnesota, and Virginia show the difficulty and limitations of these analyses and, as a result, reveal limited evidence for associations between QRIS ratings and child development.

Conclusion

The picture that emerges from the synthesis of findings across the four states and across the validation approaches is mixed. For instance, the results of efforts to validate the quality standards and indicators in QRIS generally have been successful. Efforts to review how well measures are functioning, however, reveal concerns about limited variation on some measures and QRIS structures that are producing skewed distribution of programs. There are some indications that QRIS levels are distinct with respect to measures of observed quality, but only in the QRIS that used the measures as part of the system. Finally, validation studies that included measures of children's developmental progress indicate only limited support for linkages between these measures of children's growth, QRIS ratings and program quality elements. The findings suggest that further work is needed to strengthen the ability of QRIS ratings to serve as meaningful markers of program quality.

A key theme discussed in the brief is that the information gained from validation efforts can serve as a critical tool for guiding initial design of QRIS, redesign efforts and continuous quality improvement. Zellman and Fiene (2012) emphasize that validation studies do not produce "yes" or "no" answers about QRIS but provide data that can support QRIS in a process of refining and improving. As such, validation efforts must be timed appropriately and aligned with a clear decision-making framework for how the findings will be used. In the four states highlighted in this Brief, researchers partnered with state agency leaders and other QRIS stakeholders to assist in developing a validation plan that could support QRIS development as well as a process for reviewing and interpreting findings so that the results could be applied appropriately. As states continue implementation of QRIS, administrators and stakeholders are encouraged to engage in validation efforts that can inform their systems and move progressively toward the provision of effective services.