THE CHILD CARE QUALITY RATING SYSTEM (QRS) ASSESSMENT

Measuring Quality Across Three Child Care Quality Rating and Improvement Systems: Findings from Secondary Analyses

OPRE Report 2011-30

August 2011

THE CHILD CARE QUALITY RATING SYSTEM (QRS) ASSESSMENT

Measuring Quality Across Three Child Care Quality Rating and Improvement Systems: Findings from Secondary Analyses

OPRE Report 2011-30
August 2011

Mathematica Policy Research
Lizabeth Malone
Gretchen Kirby
Pia Caronongan
Kimberly Boller

Child Trends
Kathryn Tout

Prepared for:

Office of Planning, Research and Evaluation
Administration for Children and Families
Department of Health and Human Services
370 L'Enfant Plaza Promenade, SW
7th Floor West, Room 7A011
Washington, D.C. 20447

Project Officers:
Ivelisse Martinez-Beck
Kathleen Dwyer

Prepared by:

Mathematica Policy Research
600 Maryland Ave., S.W., Suite 550
Washington, DC 20024-2512

Project Director:
Gretchen Kirby, Mathematica

Co-Principal Investigators:
Kimberly Boller, Mathematica
Kathryn Tout, Child Trends

This document was prepared under Contract #HHSP233200800394G with the Administration for Children and Families, U.S. Department of Health and Human Services. The views expressed in this publication are those of the authors and do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services. This report and other reports sponsored by the Office of Planning, Research and Evaluation are available at http://www.acf.hhs.gov/programs/opre/index.html.

Suggested citation: Malone, L., Kirby, G., Caronongan, P., Tout, K., Boller, K. (2011). *Measuring Quality Across Three Child Care Quality Rating and Improvement Systems: Findings from Secondary Analyses*. OPRE Report # 2011-30. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

OPRE
Office of Planning, Research and Evaluation
Administration for Children and Families

MATHEMATICA Policy Research

Child TRENDS

ACKNOWLEDGMENTS

This report was produced as part of the QRS Assessment Project, funded by the Office of Planning, Research and Evaluation in the Administration for Children and Families, U.S. Department of Health and Human Services. We are grateful to our federal project officers, Ivelisse Martinez-Beck and Kathleen Dwyer, who provided valuable guidance and feedback pertaining to all aspects of this work. We also benefited from the knowledge shared by members of the QRS Assessment expert panel, which helped illuminate the gaps in knowledge and launched our efforts to identify QRIS data that could address key questions. This expert panel included Sheri Azer Fischer, J. Lee Kreader, Kelly Maxwell, Deb Swenson-Klatt, Kathy Thornburg, and Gail Zellman as well as members of the project team, Shannon Christian and Karen Tvedt from Christian and Tvedt Consulting.

We would like to thank other members of the QRS Assessment team, Heather Zaveri at Mathematica, for her thoughtful contributions in the design and selection stage, along with Mark Strayer, Jacqueline Agufa, Francesca Palik, and Aaron Horvath at Mathematica, for their assistance in the analysis programming. Other colleagues at Mathematica assisted in the preparation and production of this report. Amanda Bernhardt and Laura Bernstein provided editorial support. Donna Dorsey provided administrative and secretarial support.

Finally, this report would not have been possible without the generosity of QRIS administrators, staff, and evaluators who provided access to QRIS data sources and documentation, and patiently fielded multiple follow-up questions from our research team. We offer our gratitude to the individuals in each of the three states who were gracious with their time and knowledge to support this analysis, specifically: in Miami-Dade, Lisa Pittman of the Children's Trust, Bill Kirchhoff of the Children's Forum, and Nicole Bardin and Carlos Valdes of the Early Learning Coalition of Miami-Dade/Monroe; in Illinois, Megan Fitzgerald of the Illinois Department of Human Services, Joellyn Whitehead of the Illinois Network of Child Care Resource and Referral Agencies, and Linda Butkovich of the National-Louis University; and, in Tennessee, David Shirk, Barbara Wall, and Gary Smith of the Tennessee Department of Human Services; and Bingham Graves, Karen Homer, and Maryanne Cunningham with the University of Tennessee-Knoxville.

CONTENTS

EXECUTIVE SUMMARY xiii

I INTRODUCTION 1

A. Research Strategy and Topics 2

B. Study Methods 4

C. Analytic Approach, Study Scope, and Limitations 6

D. Roadmap to the Report 11

II QUALITY COMPONENTS: DESCRIPTIONS ACROSS THREE QRIS 13

A. Analytic Goals 13

B. Analytic Strategy and Considerations 13

C. Key Findings 14

D. Detailed Findings 15

1. What percentage of center providers obtain the highest quality rating level and what are average ERS scores across providers? 15
2. What quality rating components are typically included (and, therefore, are measured or evaluated) in the rating process across QRIS? 15
3. What is the prevalence of quality rating components as measured on a common metric across QRIS? For example, what percentage of center providers meet each quality rating component (e.g., curriculum use)? 17
4. How does the prevalence of quality rating components differ between the base and highest quality rating levels across QRIS? 18

III QUALITY COMPONENTS: DIFFERENCES BETWEEN RATING LEVELS AND TYPES OF CENTERS 29

A. Analytic Goals 29

B. Analytic Strategy and Considerations 29

C. Key Findings 30

C. Detailed Findings 31

III *(conitnued)*

1. How does the prevalence of quality rating components differ between quality rating levels within Florida, Miami-Dade Quality Counts? 31
2. How does the prevalence of quality rating components differ between quality rating levels within the Tennessee Star-Quality child care program? 33
3. What similarities exist by quality rating level in the prevalence of quality rating components across QRIS? 34
4. What is the prevalence of quality rating components in child care centers by Head Start status in Florida, Miami-Dade? 36
5. What is the prevalence of quality rating components in child care centers by accreditation status in Florida, Miami-Dade? 37

IV UNIQUE CONTRIBUTIONS OF QUALITY COMPONENTS IN PREDICTING OBSERVED QUALITY 49

A. Analytic Goals 49

B. Analytic Strategy and Considerations 49

1. Specifying a proxy for quality (the outcome variable) 49
2. Specifying the sample for analysis 50
3. Specifying the quality components (the explanatory variables) 50
4. Specifying the models 52

C. Key Findings 52

D. Detailed Findings 54

1. What is the unique contribution of each quality rating component in predicting the combined ERS score? 54
2. What patterns in the associations between quality rating components and observed quality are present for ITERS-R and ECERS-R scores separately? 55
3. What patterns in the associations between quality rating components and observed quality are present for ECERS-R factor scores: Teaching and Interactions and Provisions for Learning? 55
4. What patterns in the associations between quality rating components and observed quality are present across the different QRIS? 56

V QUALITY PROFILES: PATTERNS IN QUALITY COMPONENTS 65

A. Analytic Goals 65

B. Analytic Strategy and Considerations 65

1. Selecting cross-QRIS quality profile components 65

V *(conitnued)*

C. Key Findings....66

D. Detailed Findings 67

1. What is the prevalence of centers demonstrating select components, individually or in combination, on the high end of quality?....67
2. How might centers be categorized into profiles of quality based on the number and combination of quality components that demonstrate significant associations with observed quality?....68
3. Do the profiles map to existing rating levels in each of the two QRIS in ways that create distinctions in quality? 69

VI SUMMARY AND FUTURE DIRECTIONS FOR RESEARCH....73

A. Programs in the QRIS....73

B. Patterns in Quality Improvement....74

C. Rating Levels 75

D. Quality Components....75

E. Quality Outcome 77

REFERENCES 79

APPENDIX A: SITE SAMPLE AND DATA FILE CONSTRUCTION FOR THE THREE QRISA-**ERROR! BOOKM**

APPENDIX B: ANALYTIC VARIABLE CONSTRUCTIONB-**ERROR! BOOKMARK NOT DEFINED.ERROR! B**

APPENDIX C: QRIS STANDARDS FOR CHILD CARE CENTERS IN ILLINOIS, MIAMI-DADE, FLORIDA, AND TENNESSEEC-**ERROR! BOOKMARK NOT DEFINED.ERROI**

APPENDIX D: METHODOLOGICAL NOTES BY CHAPTERD-**ERROR! BOOKMARK NOT DEFINED.ERROR**

APPENDIX E: DETAILED TABLES FOR CHAPTER II E-1

APPENDIX F: DETAILED TABLES FOR CHAPTER IV F-1

APPENDIX G: DETAILED TABLES FOR CHAPTER V H-1

APPENDIX H: DETAILED TABLES FOR CHAPTER V H-1

TABLES

I.1 Summary Characteristics of QRIS Included in Secondary Data Analysis.... 6

I.2 QRS Assessment Project Secondary Data Analysis Variables, by Quality Category 7

I.3 Quality Rating Level Categories Used in the Secondary Data Analysis..... 9

II.1 Quality Rating Levels, Observed Quality, and Quality Components Across Three QRIS 24

II.2 Percentage of Centers That Are Not Evaluated on Quality Rating Components At Base and High QRIS Rating Levels Across Three QRIS[a] 26

II.3 Percentage and Means (SDs) of Centers That Meet Quality Rating Component Criteria, At Base and High QRIS Rating Levels Across Three QRIS[a] 27

III.1 Quality Rating Components Among Florida, Miami-Dade Child Care Centers, by Rating Level[a] 39

III.2 Quality Rating Components Among Tennessee Child Care Centers, by Rating Level 41

III.3 Quality Rating Components Among Base QRIS Rating Level-Florida, Miami-Dade And Tennessee 43

III.4 Quality Rating Components Among Middle QRIS Rating Level(s)-Florida, Miami-Dade And Tennessee 44

III.5 Quality Rating Components Among Highest QRIS Rating Level(s)-Florida, Miami-Dade And Tennessee 45

III.6 Quality Rating Levels and Components in Florida, Miami-Dade, by Head Start Status 46

III.7 Quality Rating Components Among Florida, Miami Dade Child Care Centers, By Accreditation Status 48

IV.1 Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ECERS-R+ITERS-R Total Scores, by Cross-Site and Site-Specific Models..... 58

IV.2 Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ITERS-R Total Scores, by Cross-Site and Site-Specific Models..... 59

IV.3 Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ECERS-R Total Scores, by Cross-Site and Site-Specific Models..... 60

IV.4 Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ECERS-R Teaching and Interactions Scores, by Cross-Site and Site-Specific Models 61

IV.5 Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ECERS-R Provisions for Learning Scores, by Cross-Site and Site-Specific Models 62

IV.6 Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ECERS-R+ITERS-R Total Scores, by QRIS 63

V.1 Percentage of Centers Demonstrating Quality Profile Components Across Two QRIS 70

V.2 Percent of Centers Meeting a Particular Count of Quality Profile Components Across Two QRIS 70

V.3 Correlations Among Quality Profile Components Across Two QRIS..... 71

V.4 Quality Profiles Across Two QRIS 71

V.5 Percentage of Centers at Each Quality Profile by Rating Level, Miami-Dade 72

V.6 Percentage of Centers at Each Quality Profile by Rating Level, Tennessee 72

FIGURES

II.1 Percentage of Center-based Providers by Rating Level Across Three QRIS 20

II.2 ERS Scores for Center-based Providers Across Three QRIS 20

II.3 Family Partnerships, Staff Benefits, and Program Management Quality Components Across Three QRIS 21

II.4 Staff Management and Salary Scale Quality Components Across Three QRIS 21

II.5 Individualization Quality Components Across Three QRIS 22

II.6 Child to Staff Ratios for Center-based Providers Across QRIS 23

II.7 Director and Staff Education in Center-based Providers Across Three QRIS 23

EXECUTIVE SUMMARY

States and communities have adopted Quality Rating and Improvements Systems (QRIS) as a tool to promote, measure, and monitor the quality of early child care programs. The primary goal of a QRIS is to improve quality across a range of child care programs, with the intent of providing positive experiences for all children. The fundamental design decision then is to define and measure quality.

States and localities have used the information and knowledge that is available from research, program administrator's experiences, and key stakeholders to design systems they hope are valid and meaningful in defining quality for providers, parents, and children alike. At this time, information about the components to include in a QRIS, in what combination, and at what cut-points per level is lacking. In an environment in which adoption, implementation, and refinement of QRIS are moving quickly but the research base to inform decision-making is slim, the Office of Planning, Research, and Evaluation (OPRE), within the Administration for Children and Families, initiated the Child Care Quality Rating Systems (QRS) Assessment project.

This report, produced as part of the QRS Assessment, presents findings from an exploratory analysis of administrative data from three QRIS, which examines the prevalence of quality components across providers and how they function in relation to observed quality. As QRIS enters its second decade, it is clear there has been a growing sophistication in data collection, providing opportunities for in-depth analysis of distinctive QRIS as well as cross-QRIS analysis. This analysis is developmental in nature; the findings are tenable within the limitations and scope, but should be interpreted with caution and are not confirmatory. The larger contribution of this work ties back to the intent of the Assessment project as a whole—to contribute to and build avenues for future analysis that can support a growing body of research that will inform decision making.

A. Study Methodology

The objective of this work focuses on exploratory analysis to inform development and practice related to QRIS standards and ratings. The analysis builds on the categorization of quality categories that was developed in the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010), which identified 13 categories that capture the range of components used by QRIS to define quality within the rating structure (Exhibit ES.1).

Exhibit ES.1. Thirteen Categories of Quality Components Used by 26 QRIS

Licensing compliance	Family partnerships
Ratio and group size	Administration and management
Health and safety	Cultural and linguistic diversity
Curriculum	Accreditation
Environment	Provisions for special needs
Child assessment	Community involvement
Staff qualifications	

Source: Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010), produced as part of the QRS Assessment project.

Study Methods. We selected three QRIS for inclusion in the secondary data analysis based on criteria defining data coverage, access, and documentation. These are: Florida, Miami-Dade; Illinois; and Tennessee. A summary of characteristics of these QRIS and the sample sizes of child care centers used in the analysis is presented in Table ES.1. Within a site, we focused on centers that

served children birth through five years. We did not include centers serving only school-age children. Data came from three sources: (1) QRIS rating administrative databases, (2) professional development registries, and (3) observation databases.

Table ES.1. Summary Characteristics of QRIS Included in Secondary Data Analysis

	Miami-Dade, Florida	Illinois	Tennessee
QRIS name	Quality Counts	Quality Counts	Star-Quality Child Care Program
Oversight agency	Early Learning Coalition of Miami-Dade/Monroe	Illinois Department of Human Services	Tennessee Department of Human Services
Year of full implementation	2008	2007	2001
Number of rating levels	5	4	3
Structure of rating levels	Combination	Building Block	Combination
Validity period of rating level	Initial: 2 years Subsequent: 13 months	1 year	1 year
Voluntary	Yes	Yes	Yes; all licensed providers must participate in Report Card program but progression beyond the Report Card is voluntary
Providers included in the analysis	Center-based including participating Head Start programs	Center-based including participating Head Start programs	Center-based, including Head Start programs, that voluntarily participate beyond the Report Card program
Analysis time period	October 2008–September 2010	April 2008–August 2010	October 2009–September 2010
Sample size	253[a]	350	1,369

Source: Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); secondary data analysis conducted as part of the QRS Assessment project.

[a] As the newest QRIS included in the analysis, 137 of the 253 centers had initial baseline ratings that are re-evaluated after 2 years..

Research Topics and Analytic Methods. Using the defined Compendium quality categories as an analysis framework, the secondary data analysis examined the following four research topics:

1. ***Examining quality rating components across QRIS*** (through univariate analyses--means, ranges, percentages--of the individual components)
2. ***Examining differences in quality rating components between quality rating levels and for certain types of centers*** (through bivariate analyses such as tests of significance)

3. ***Exploring the unique contribution of each quality rating component on observed quality*** (through multivariate analyses employing regression models with observed quality as the outcome and the quality rating components as explanatory variables)
4. ***Describing patterns of quality*** (through univariate descriptive analyses to develop quality profiles based on select quality components meeting particular thresholds)

The research questions themselves evolved and were refined over time, adapting to the coverage and detail of available data, and challenges to defining variables across systems. Nonetheless, two fundamental concepts remained in place throughout the process and distinguish this analysis from other work that has been conducted to date: (1) development and use of a common metric across QRIS for defining variables within the quality categories, and (2) use of a "not evaluated" category within a component to capture whether the evaluation of a quality component has bearing on the measurement of quality.

Analytic Approach. Using the framework of the quality components as defined in the Compendium and the parameters presented by the administrative data available, we specified the quality component variables, quality outcome (observed quality), and a categorization for the rating levels to form the basis for the analysis.

- **Quality rating components.** We created common measures across the QRIS, generally one variable for 10 of the 13 quality component categories identified in the Compendium. Where possible, we retained the continuous numeric values for maximum variability in conducting univariate and bivariate analyses and then created binary variables for use in the multivariate analysis.
- **Observed quality.** Observed quality (as measured by the Early Childhood Environment Rating Scales [ERS]) can be found as a quality component in many QRIS. The analysis was limited to existing data collected by QRIS. Observed quality as represented by the ERS is used in this analysis to examine the correlation and association of all other quality rating components to observed environment quality scores as an objective measure of quality.
- **Quality rating levels.** The three sites vary in the number of quality level ratings they assign. A grouping of levels across the sites was of interest in order to examine patterns in the prevalence of quality components within and between levels across the QRIS. We collapsed the rating levels from each of the three QRIS into three groups—base, middle, and high.

Limitations. The use of administrative data and the selection of a small number of select QRIS have some inherent limitations: (1) the data were not collected as part of a study designed to answer specific research questions relevant to QRIS quality measurement design and refinement, and (2) the findings are not representative of QRIS beyond those included in the analysis. Other limitations stem from analytic decisions that were made to accomplish a cross-QRIS analysis: (1) these analyses pertain to center-based care only, and (2) in defining and modeling quality components across QRIS, site-specific components may be muted. That is, the cross-QRIS quality components are often summative in nature, capturing multiple site-specific QRIS indicators within a broader measure.

We address the research topics from a descriptive perspective with exploratory findings that are correlational in nature. The findings from this analysis are not confirmatory in defining the components to measure to maximize state or county resources and still achieve reliable and valid overall quality ratings. In time, these findings in combination with other current and future work will

contribute to a body of research that can be used to make specific programmatic and policy-relevant decisions.

B. Quality Components: Descriptions Across Three QRIS

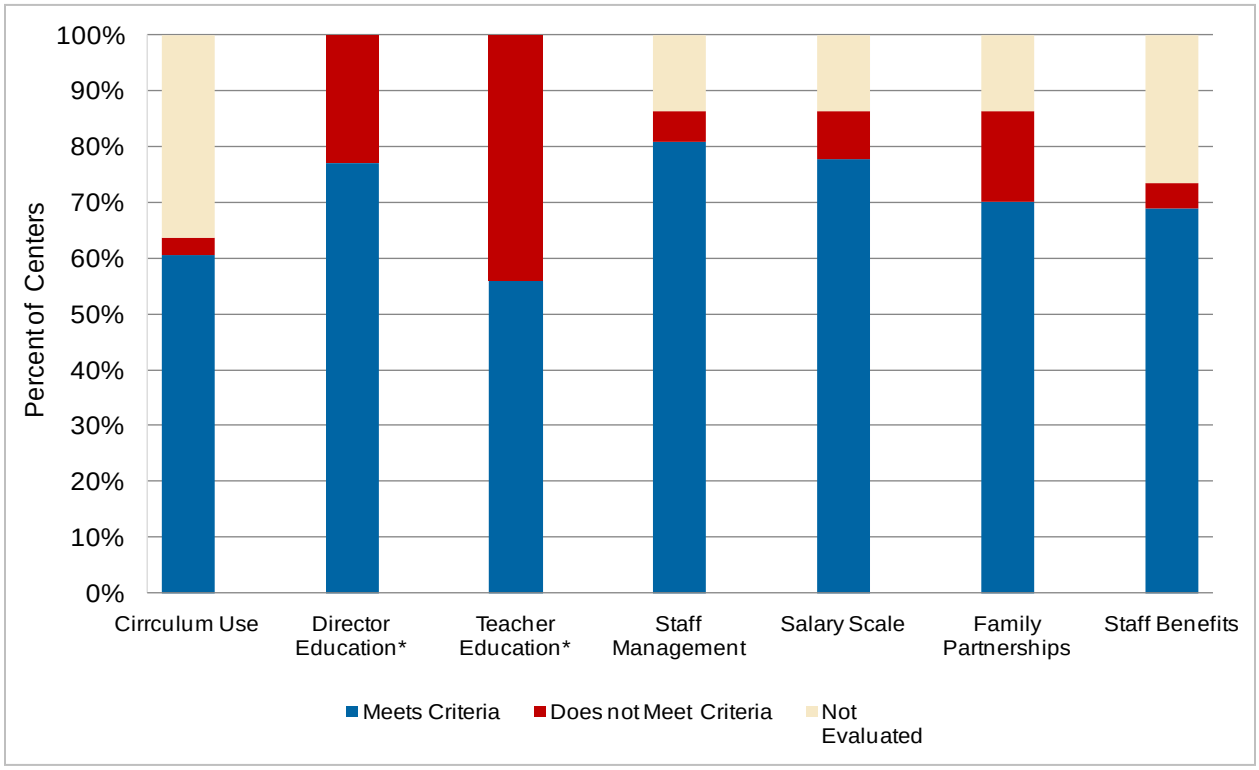
This descriptive analysis examines the emphasis states place on each quality component in the QRIS rating process and their prevalence across the three QRIS. The analysis addresses the following questions:

- What quality rating components are typically included (and, therefore, are measured or evaluated) in the rating process across QRIS?
- What is the prevalence of quality rating components as measured on a common metric across QRIS? For example, what percentage of center providers meet each quality rating component (e.g., curriculum use)?
- How does the prevalence of quality rating components differ between the base and highest quality rating levels across QRIS?

Key Findings:

- Seven quality components are demonstrated by more than half of all center-based providers across the three QRIS (Figure ES.1). They are:
  - Use of a curriculum based on early learning guidelines or standards
  - Director with some college or higher level of education
  - Majority (50 percent or more) of teaching staff that have some college or higher level of education
  - Implementation of staff management practices (such as staff meetings, annual staff performance evaluations, orientation, or staff development plans)
  - Use of a salary scale that differentiates pay by an individual's education, training, or experience
  - Implementation of the full range of family partnership activities specified by the QRIS
  - Offering of the full range of staff benefits specified by the QRIS
- Four components reflecting the degree to which providers individualize services to meet the needs of children and families are not often evaluated—more than 50 percent of centers across the three QRIS are not evaluated on these components. They are (1) conducting child assessment and screening, (2) use of child assessment results to guide planning, (3) communicating with families in their home language, and (4) planning for children with special needs (Figure ES.2).
- Center-based providers at the high rating level have observed quality and the presence of quality components that are higher, on average, than those at the base level (Figure ES.3). For example, the average combined ERS score (ECERS-R+ITERS-R) for providers at the high level is 5.1 compared to 3.3 for providers at the base level.

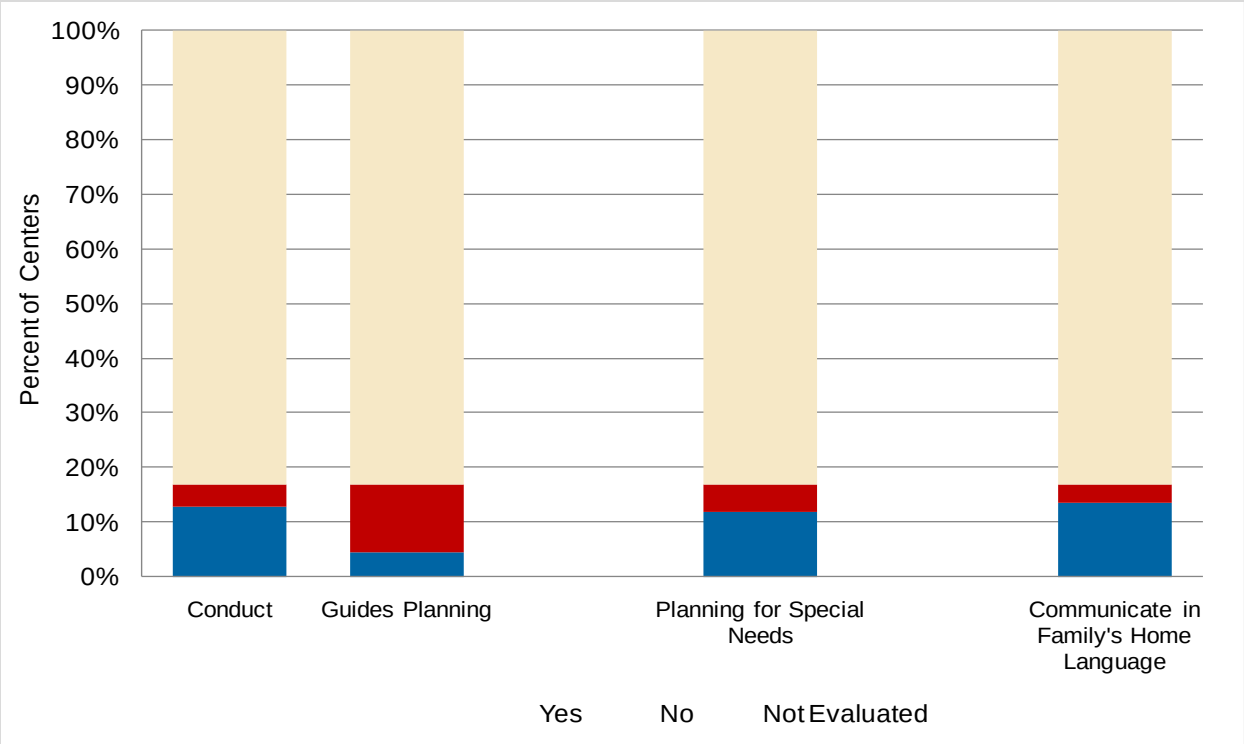
Figure ES.1. Quality Components Demonstrated by Center- Based Providers



Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

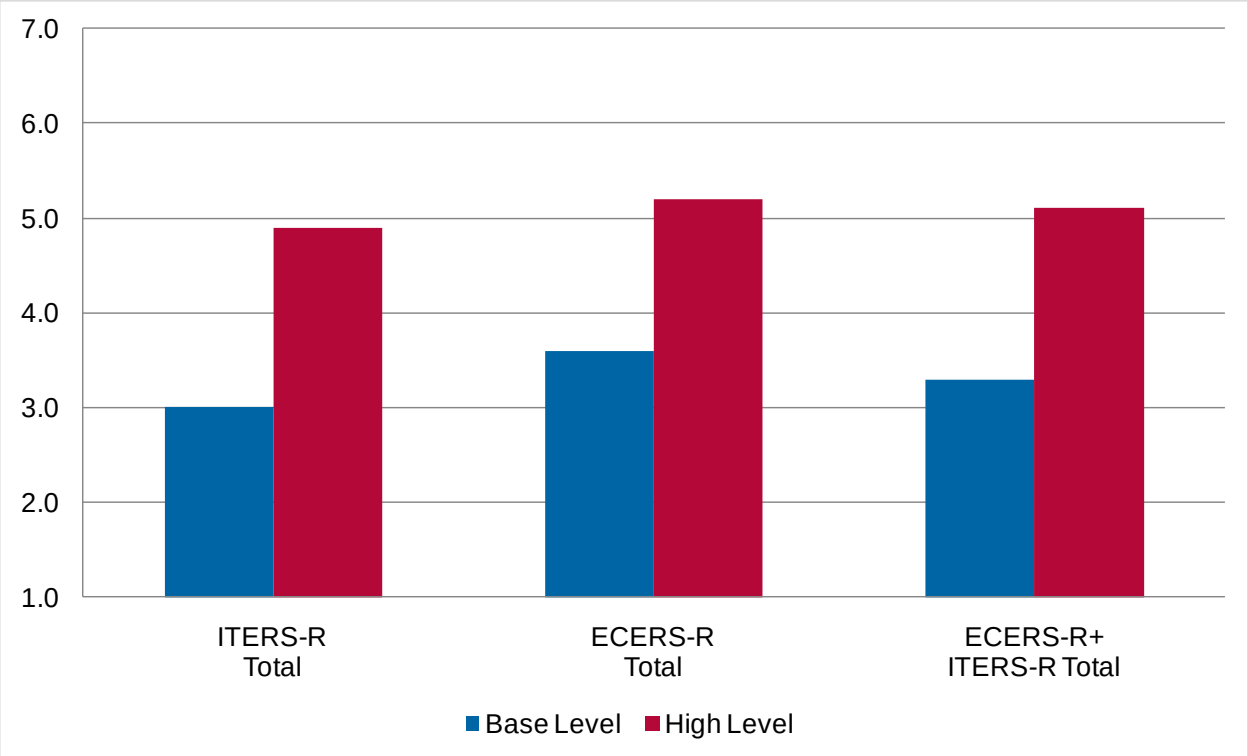
*Percentage with education of some college or higher.

Figure ES.2. Individualization Quality Components Across Three QRIS



Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

Figure ES.3. Individualization Quality Components Across Three QRIS



Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

C. Quality Components: Differences Between Rating Levels and Types of Centers

We use bivariate analyses to explore the following:

- How does the prevalence of quality rating components differ between quality rating levels within each QRIS?
- What similarities exist by quality rating level in the prevalence of quality rating components across QRIS?
- What is the prevalence of quality rating components by different types of centers (such as Head Start status or accreditation status)?

Key Findings:

- Centers within each rating level are generally distinct in the prevalence of quality rating components. On average, centers in the middle level demonstrate quality components at higher rates than those at the base, and centers at the highest level demonstrate quality components at higher rates than those in the middle.
- There are some differences between the two states in these patterns, however. With few exceptions, centers consistently demonstrate increasing rates in each of the quality components from one of the three levels (base, middle, high) to the next in Tennessee. In Miami-Dade, a number of quality components demonstrate a significant difference from only one level to the next, rather than between each of the three levels.

- The prevalence of quality components among centers at the same rating levels but in different QRIS is not entirely the same, but the differences that do exist tend to be smaller at the highest level (with the notable exception of teacher education). Licensing standards, subsidy reimbursement policies, professional development systems and supports, and simply the longevity of the QRIS, influence the prevalence of quality components at each rating level such that significant differences exist between Miami-Dade and Tennessee.
- Head Start programs generally do not differ from other centers participating in the QRIS in Miami-Dade in ERS scores. They are, however, more likely to have in place quality components that include individualization practices (such as those around curriculum, child assessment, planning for children with special needs, and communicating with families in their home language). Head Start programs are also more likely than other centers to have directors and teaching staff with some college or more, to demonstrate all family partnership strategies specified by the QRIS, and to use a differentiated salary scale.
- Accredited centers are more likely to demonstrate many of the quality components over non-accredited centers. However, accredited centers do not generally differ from other centers in the education levels of directors or teaching staff. Accredited centers have higher ERS scores, on average, than other centers and are more likely to be at the highest rating level.

D. Unique Contributions of Quality Components in Predicting Observed Quality

Through multivariate analyses, we examine the unique contribution of each quality rating component in predicting observed environment quality, as measured by the ERS, to answer the following questions:

- What is the association between each quality rating component and observed quality when all other components are held constant?
- What patterns in these associations are present across different ERS scales and different QRIS?

Key Findings:

- Across sites and different specifications of the ERS outcome, the use of a differentiated salary scale and accreditation status were consistently associated with higher scores (Table ES.2). Centers with differentiated salary scales were found to have higher ERS scores compared to centers that do not offer a similar benefit. Accredited centers were consistently found to have higher ERS scores than those not accredited.
    - These associations do not indicate that certain quality components cause higher ERS scores or vice versa.
    - The unique contribution of these two components may reflect characteristics not observed or not directly measured such as motivation, or time and resources that can also influence the presence of other quality initiatives overall.

Table ES.2. Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ECERS-R+ITERS-R Total Scores, by Cross-Site and Site-Specific Models

	Cross-site model		Florida, Miami-Dade				Tennessee	
			Cross-site common components		Individualization components			
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	2.63	0.11***	2.54	0.20***	2.52	0.20***	4.04	0.15***
Curriculum-staff training (vs. not all staff)								
All teaching staff	0.05	0.08	-0.05	0.11	-0.08	0.11	0.17	0.12
Some teaching staff	-0.14	0.11	0.01	0.19	0.02	0.19	-0.10	0.15
Not evaluated	0.13	0.08					0.23	0.12+
Child-to-staff ratio (3-year-old - 8:1 or lower)	0.12	0.04**	0.08	0.10	0.08	0.10	0.12	0.04**
Teacher education (vs. less than 25 percent)								
50 percent or greater with some college or higher	0.25	0.08**	0.21	0.12+	0.20	0.12+	0.35	0.12**
25 to 50 percent with some college or higher	0.12	0.08	0.09	0.10	0.10	0.10	0.23	0.13+
Director education (some college or higher vs. less)	0.03	0.06	0.05	0.10	0.03	0.10	0.04	0.08
Family partnerships (all vs. less than all)	0.12	0.06+	0.19	0.12	-0.01	0.16	0.12	0.07+
Staff management	0.06	0.10	0.25	0.22	0.25	0.22	-0.10	0.14
Salary scale	0.41	0.09***	0.44	0.12***	0.39	0.13**	0.43	0.12***
Accredited center	0.34	0.06***	0.33	0.09***	0.32	0.09***	0.36	0.07***
Individualization; assessment, planning for special needs, family communication (vs. two or less)								
More than half	0.18	0.11+			0.31	0.15*		
Not evaluated	1.49	0.12***						
R-square	0.51		0.30		0.31		0.12	
Adjusted R-square	0.50		0.27		0.28		0.11	
Number of centers	1327		227		227		1100	

Source: Analysis of data from Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- Three other components—teacher education, child-to-staff ratios, and family partnerships—demonstrated unique contributions across models, but not consistently and not at the level of significance seen for the salary scale and accreditation components.

The differences in ERS scores associated with individual quality components were all less than half a point in magnitude. Half a point is equivalent to the difference in ERS scores between a level one and level two provider in Florida, Miami-Dade and Tennessee. Within the context of QRIS as currently designed, the largest estimated difference in ERS scores associated with any of the components is smaller than what the systems would require to move up one quality rating level.

E. Quality Profiles: Patterns in Quality Components

The influence on quality and, in turn, child outcomes may best be accounted for by considering more than one component at a time. We begin to explore a composite view, describing the patterns for a subset of cross-QRIS quality components across two QRIS (Florida, Miami-Dade and Tennessee), by addressing the following questions:

- What is the prevalence of centers demonstrating select components on the high end of quality?
- How might centers be categorized into profiles of quality based on the number and combination of quality components that demonstrate significant associations with observed quality?
- Do the profiles map to existing rating levels in each of the two QRIS in ways that create distinctions in quality?

We selected five quality components from which to build profiles of providers in the two QRIS. Four components—use of a differentiated salary scale, family partnership strategies, teacher education levels, and child-to-staff ratios—consistently or frequently demonstrated unique contributions in predicting observed quality in the multivariate analysis. We add observed quality (the ERS score) back as a component, rather than an outcome, to contribute to the profiles since the learning environment is often included as a rating component in QRIS. In general, these five components are among those that are commonly included in rating systems and therefore are relevant to a broader set of QRIS.

Key Findings:

Based on five selected quality components and associated thresholds, centers in two QRIS fall into seven distinct quality profiles. The profiles are differentiated by (1) the number of quality components that meet the thresholds (that is, one to five), (2) ability of the center to meet the learning environment threshold of a combined ERS score of 5.0 or higher, and (3) ability to meet the requirements of both the family partnership component and use of a differentiated salary scale.

- The most prevalent profile, in which 27 percent of centers fall, is the one in which all five quality component thresholds are met (Table ES.3). This suggests that for nearly three-quarters of the centers, these five quality components and associated thresholds are sufficiently demanding, demonstrated by the variation in centers meeting different numbers and combinations of the thresholds.

- The next most common profile demonstrates mixed components. The 24 percent of centers in this profile did not meet the learning environment threshold, but have the pairing of the family partnership and salary scale components (Table ES.3). They may also meet the threshold for one other component (either teacher education or child-to-staff ratio, but not both).
- Very few centers—only three percent—do not meet any of the five quality component thresholds (Table ES.3).
- The overall quality rating levels assigned to centers in each of the two QRIS map to the profiles such that providers assigned to the higher quality profiles have higher rating levels, and vice versa (Tables ES.4 and ES.5). This suggests that the five components and the associated thresholds may indeed be similar drivers in creating distinctions in quality across different QRIS.

Table ES.3. Quality Profiles Across Two QRIS

Profile		Description	Number of Threshold Components	ERS Meet Threshold	Percentage of Centers
A	Maximizing quality	All components meet threshold	5	Yes	27.1
B	High operational quality	ERS not meet threshold, all others do	4	No	17.0
C	High center environment	ERS meet threshold as well as at least family partnerships and salary scale	3 - 4	Yes	13.2
D	High observed quality plus	ERS meet threshold and have at least one other component, but lack the family partnership-salary scale combination	2 - 4	Yes	3.7
E	Mixed center environment	ERS not meet threshold but have at least family partnerships and salary scale	2 - 3	No	24.4
F	Lower observed quality plus	ERS not meet threshold but have at least one component, but lack the family partnership-salary scale combination	1 - 3	No	11.4
G	Base quality	No components meet thresholds	0	No	3.2
Number of centers					1,334

Source: Analysis of data from Florida, Miami-Dade and Tennessee conducted as part of the QRS Assessment project.

Table ES.4. Percentage of Centers at Each Quality Profile by Rating Level, Miami- Dade

		Level 1	Level 2	Level 3	Level 4	Level 5	Number of Centers
A	Maximizing quality	0.0	0.0	0.0	0.0	0.0	0
B	High operational quality	0.0	0.0	0.9	1.3	0.4	6
C	High center environment	0.0	0.0	0.0	0.9	1.3	5
D	High observed quality plus	0.0	0.0	0.0	0.4	0.0	1
E	Mixed center environment	0.0	8.1	26.9	17.5	3.4	131
F	Lower observed quality plus	2.1	11.5	9.8	1.3	0.4	59
G	Base quality	5.1	6.4	2.1	0.0	0.0	32
Number of centers		17	61	93	50	13	234

Source: Analysis of data from Miami-Dade, Florida as part of the QRS Assessment project.

Note: Shaded cells indicate the highest concentration of providers for each quality profile.

Table ES.5. Percentage of Centers at Each Quality Profile by Rating Level, Tennessee

		Level 1	Level 2	Level 3	Number of Centers
A	Maximizing quality	0.0	0.1	32.8	362
B	High operational quality	0.0	0.8	19.3	221
C	High center environment	0.0	1.4	14.2	171
D	High observed quality plus	0.0	2.5	1.9	48
E	Mixed center environment	0.1	7.4	10.2	194
F	Lower observed quality plus	0.8	5.5	2.2	93
G	Base quality	0.8	0.2	0.0	11
Number of centers		19	195	886	1100

Source: Analysis of data from Tennessee as part of the QRS Assessment project.

Note: Shaded cells indicate the highest concentration of providers for each quality profile.

F. Future Directions for Research

This analysis covered a number of topics related to quality measurement through use of existing administrative data from a small and select number of QRIS. As a descriptive, exploratory analysis it provides useful and interesting information, but it also identifies areas in need of further research to add both context and evidence.

Programs in the QRIS. An important issue in the evaluation of QRIS as well as the interpretation of any research and evaluation is which centers select into voluntary QRIS. Specifically, do centers that are high quality based on the definition of the QRIS choose to participate at a higher rate than those that may be rated as low quality? QRIS stakeholders and evaluators need to further explore and understand the characteristics of programs that participate in QRIS and those that do not in order to fully examine changes in quality and, ultimately, changes in child outcomes that may be associated with QRIS.

Patterns in quality improvement. Based on data collected from a cross-section of child care centers in a few, select QRIS it appears that there are distinctions between the base, middle, and high rating levels in that providers at the higher rating levels are more likely to demonstrate the quality components examined than those at the lower levels. What we cannot determine from this analysis, however, is whether and how these components have changed for providers over time. That is, we cannot assess the degree to which this snapshot of providers by level may be an indication of patterns that have persisted or evolved over time.

Rating levels. We find more similar proportions of centers meeting quality components when we compare centers at the *highest* rating level in the two sites (except for teacher education), building on the findings in a companion qualitative report that QRIS requirements are more similar at the highest level (Caronongan et al. 2011). This may provide some confidence in the face validity of QRIS indicators and suggest that they are implemented with some consistency across QRIS. (Though we stress that these findings are based on just two QRIS.) This also supports the importance of continuing to conduct validation studies of QRIS with a focus on the highest level, at least to start, and an examination of child outcomes as well as other indicators such as parent satisfaction and service to low-income children.

Quality components. Critical to decisions of design and refinement of QRIS are issues related to which quality components to measure and in what combination. This analysis provides some suggestive findings from two QRIS. Overall, the patterns of findings in this analysis confirm the importance of domains long considered predictors of quality (such as teacher education, child-to-staff ratios, and salary). The findings also support the inclusion of other domains with a smaller evidence base in early childhood, such as the use of family partnership strategies and of practices that promote the individualization of services for children and families. These findings can help structure an agenda for next steps in both research and, potentially practice, through further examination of the resources necessary for quality improvement, measurement refinement of particular components, and definition to emerging components. The larger contribution of this work may be the introduction of a framework for using a common metric across QRIS that can be used in future evaluations to continue building knowledge in this area.

Quality outcomes. Lastly, but most importantly, the field needs to continue building knowledge to better understand how the ERS are related to the other components of quality that are measured in QRIS, as well as to continue to seek other outcome measures that capture the dimensions of quality most closely linked with child outcomes.

I. INTRODUCTION

States and communities have adopted Quality Rating and Improvements Systems (QRIS) as a tool to promote, measure, and monitor the quality of early child care programs. State motivations for developing and implementing QRIS are rooted in increasing pressures policymakers face as public accountability efforts in early childhood and elementary education focus on reducing the large disparities in how children are faring based on a family's economic circumstances (ACF 2002; Halle et al. 2009; Love et al. 2005). The primary goal of a QRIS is to improve quality across a range of child care programs, with the intent of providing positive experiences for all children. The fundamental design decision then is to define and measure quality.

QRIS are dynamic—with the number of new systems increasing and refinements to existing systems continually evolving—and the pressure for increased data collection and research is mounting. In an environment in which adoption, implementation, and refinement of QRIS are moving quickly but the research base to inform decision-making is slim, the Office of Planning, Research, and Evaluation (OPRE), within the Administration for Children and Families, initiated the Child Care Quality Rating Systems (QRS) Assessment project. This work, contracted to Mathematica Policy Research, Child Trends, and Christian and Tvedt Consulting, is intended to (1) gather and analyze existing and new information on QRIS implementation and research to inform decision making on QRIS development and refinement and (2) build the capacity for ongoing monitoring and evaluation within and across systems.

The Compendium of QRS and Evaluations—the first product of the Assessment project—documented that broad quality categories are similar across 26 QRIS (Tout et al. 2010). However, the exact definitions of the components within the categories and how they are measured varies widely. A qualitative report on quality measurement practices, also produced under the QRS Assessment project, describes this variation, which is due to different state contexts with regard to licensing and existing infrastructure (the foundation of QRIS), the emphasis a QRIS places on accreditation, and the perceived status of the level of current quality among child care providers and the gaps that need to be addressed (Caronongan et al., 2011). This report presents results of a cross-QRIS secondary data analysis that develops a common framework for examining the prevalence of quality components and their association with observed quality within the variation that exists across systems.

States and localities have used the information and knowledge that is available from research, program administrator's experiences, and key stakeholders to design systems they hope are valid and meaningful to providers, parents, and children alike. At this time, information about the components to include in a QRIS, in what combination, and at what cut-points per level is lacking. Little empirical work has brought together and examined data on the use of quality components for the purpose of producing a systematic rating as a global quality measure (which the QRIS rating level aims to provide) or determining their linkage to an objective measure of observed quality. The existing validation studies present vital information for each individual system examined (Smart Start Evaluation Team 2001; Elicker et al. 2007; Norris et al. 2003; Norris and Dunn 2004; Zellman et al. 2008). However, empirical gaps in quality measurement across QRIS emerge—states continue to

seek guidance in how best to measure quality and make meaningful distinctions between quality levels.[1]

As QRIS enters its second decade, it is clear there has been a growing sophistication in data collection, providing opportunities for in-depth analysis of distinctive QRIS as well as cross-QRIS analysis. This report presents findings from an exploratory analysis of administrative data from three QRIS, which examines the prevalence of quality components across providers and how they function in relation to observed quality. This analysis is developmental in nature; the findings are tenable within the limitations and scope, but should be interpreted with caution and are not confirmatory. The larger contribution of this work ties back to the intent of the Assessment project as a whole—to contribute to and build avenues for future analysis that can support a growing body of research that will inform decision making.

A. Research Strategy and Topics

The objective of this work focuses on exploratory analysis to inform development and practice related to QRIS standards and ratings. The analysis builds on the categorization of quality categories that was developed in the Compendium, which identified 13 categories that capture the range of components used by QRIS to define quality within the rating structure (Exhibit I.1). Of these categories, nearly half are included in 20 or more of the QRIS included in the Compendium. The most commonly included categories are licensing compliance (26), environment (23), staff qualifications (25), family partnership (23), administration and management (23), and accreditation (20). In comparison, four categories are included in fewer than 10 of the QRIS examined: health and safety (3), cultural and linguistic diversity (9), provisions for children with special needs (9), and community involvement (6). We identified select QRIS for which existing administrative data could support in-depth analysis of the quality categories and their components across systems.

Exhibit I.1 Thirteen Categories of Quality Components Used by 26 QRIS

Licensing compliance	Family partnerships
Ratio and group size	Administration and management
Health and safety	Cultural and linguistic diversity
Curriculum	Accreditation
Environment	Provisions for special needs
Child assessment	Community involvement
Staff qualifications	

Source: Compendium of QRS and Evaluations (Tout et al. 2010), produced as part of the QRS Assessment project.

Research topics. Using the defined Compendium quality categories as an analysis framework, the secondary data analysis examined the following four research topics:

1. Examining quality rating components across QRIS:

[1] The Quality Initiatives Research and Evaluation Consortium (INQUIRE) funded by OPRE is producing two briefs on the topic of QRIS validation. One brief outlines different types of validation strategies. The other brief examines and compares the validation strategies used in four QRIS: Indiana, Maine, Minnesota and Virginia. The Briefs will be available late in 2011.

- What percentage of center providers obtain the highest quality rating level and what are the average ERS scores across providers?
- What quality rating components are typically included (and, therefore, are measured or evaluated) in the rating process across QRIS?
- What is the prevalence of quality rating components as measured on a common metric across QRIS? For example, what percentage of center providers meet each quality rating component (e.g., curriculum use)?
- How does the prevalence of quality rating components differ between the base and highest quality rating levels across QRIS?

2. Examining differences in quality rating components between quality rating levels and for certain types of centers:
   - How does the prevalence of quality rating components differ between quality rating levels within a site?
   - What similarities exist by quality rating level in the prevalence of quality rating components across QRIS?
   - What is the prevalence of quality rating components by different types of centers (such as Head Start status or accreditation status)?
3. Exploring the unique contribution of each quality rating component on observed quality:
   - What is the association between each quality rating component-with observed quality when all other components are held constant?
   - What patterns in these associations are present across different scales and different QRIS?
4. Describing patterns of quality:
   - What is the prevalence of centers demonstrating select components on the high end of quality?
   - How might centers be categorized into profiles of quality based on the number and combination of quality components that demonstrate unique contributions in predicting observed quality?
   - Do the profiles map to existing rating levels in the QRIS in ways that may create distinctions in quality?

The research questions themselves evolved and were refined over time, adapting to the coverage and detail of available data, and the nuances to variable definition (and its limitations) across systems. Nonetheless, two fundamental concepts remained in place throughout the process and distinguish this analysis from other work that has been conducted to date: (1) development and use of a common metric across QRIS for defining variables within the quality categories, and (2) use of a "not evaluated" category within a component to capture whether the evaluation of a quality component has bearing on the measurement of quality.

Use of a common metric for defining quality components. In developing a common metric for defining each of the quality components, we can better understand the role that each may play in contributing to quality across QRIS and varying contexts. While site-specific quality components have been designated to achieve particular levels, this approach places the components on a different metric to examine the full variation of components in predicting observed quality (for example, using the actual percentage of teaching staff that have some college education or more rather than a yes/no indicator that a specific rating level requirement was met for education among staff). The common metric captures the intent of each component despite the site-specific nuances. In so doing, the analysis does not analyze or replicate any one particular approach. Rather, it takes a broader perspective—albeit one that builds on the concepts of select QRIS—to examine the prevalence of quality components and their association with observed quality levels.

Use of a "not evaluated" category within quality components. Within and across the existing 26 QRIS profiled in the Compendium, providers may differ in the presence or level of a host of quality components. For example, providers may or may not have a curriculum in place. The 26 QRIS also vary in the requirements in place for quality components; that is, not all QRIS specify each category of components as a part of its rating, and some require particular components only for providers applying for particular rating levels. To support a rich cross-QRIS analysis, it is important to distinguish between providers who may not possess a component (for example, did not have a curriculum) and those who were not evaluated on that component (for example, the QRIS does not assess whether a curriculum is present), especially as the provider may in fact possess that component but it is not evident in the QRIS database (by design). A "not evaluated" category, therefore, was developed in order to capture this distinction. A provider is coded as "not evaluated" for a quality component when a QRIS does not prescribe the measurement of that component but another QRIS does, or when a given provider within a QRIS is not evaluated on that component (based on rating level). This approach enables an analysis of whether *not* including a component explains variation in quality.

B. Study Methods

The project team followed guidance from an expert panel and federal project officers to specify the scope and goals of the analysis so it would be in line with the intent of the QRS Assessment project. Given the parameters on resources and scope, the team worked to identify three QRIS through a two-stage selection process that had the data availability, coverage, and quality that would support the analysis.

In the first stage, we developed a set of four criteria based on QRIS characteristics from the Compendium that were indicative of program stability and structure that could support the analysis: (1) QRIS that were currently active, not a pilot, and had been implemented for a least one year; (2) QRIS that used a multi-level rating score (this eliminated one QRIS that did not assign levels and another that used only one level); (3) QRIS that followed either a building block approach where items on one level must be met before being eligible for the next level or a combination approach that uses building blocks for individual components to support developing a common metric and comparison by rating levels; and (4) QRIS that include an observational measure of quality that is independent from the sources of evidence supporting determination of the presence of other quality components. A fifth criterion—the presence of a QRIS database as the main data source—was gathered from the QRS assessment screener on data and monitoring information (internal OPRE document). Nine QRIS met the five criteria for selection in the first stage: Miami-Dade, Florida, Palm Beach, Florida, Illinois, Indiana, Kentucky, Maryland, Oklahoma, Pennsylvania, and Tennessee.

In the second stage of selection, two researchers conducted 60-minute conference calls with QRIS representatives to review construction, coverage, access, and documentation of data collected as part of the quality rating process. Between April and June 2010, the researchers conducted screener calls with eight of the selected QRIS (all but Kentucky),[2] focusing on confirming information sources and databases (data access), the presence of data necessary to the analysis (data coverage), and supporting documentation (data documentation). Specifically, the team reviewed the coverage of existing databases for the final rating, Environment Rating Scales (ERS) scores (including overall scores, subscales, and item-level data), and other quality components, prioritizing those QRIS with the most complete and detailed inventory of each of the quality components. We considered access in terms of the reports or format of files that could be exported for analysis, linkages across data sources (such as between the QRIS database and the professional development registry), and the ease in obtaining data from the sites and their data collection or management partners. Lastly, we prioritized sites with documentation such as manuals or codebooks on how data are collected, entered, or stored. This information helped in assessing the degree of data manipulation and cleaning necessary to define analytic variables on a common metric.

Selected QRIS and sample for analysis.[3] Based on these criteria of data coverage, access, and documentation, we selected three QRIS for inclusion in the secondary data analysis. These are: Florida, Miami-Dade, Illinois, and Tennessee. A summary of characteristics of these QRIS and the sample sizes of child care centers used in the analysis is presented in Table I.1. Within a site, we focused on centers that served children birth through five years. We did not include centers serving only school-age children. All centers are licensed, as required in Illinois and Tennessee, to participate in QRIS; although not required for Florida, Miami-Dade, currently all participating centers are licensed.

Data sources. Data came from three sources: (1) QRIS rating administrative databases, (2) professional development registries, and (3) observation databases. The primary information source for quality rating levels and components was the QRIS database maintained by each site. However, the coverage of the QRIS database for all quality components leading to the quality rating level assigned varied, such that other sources were often necessary to ensure complete data for the analysis. Two QRIS—Florida, Miami-Dade and Illinois maintain professional development records that were used to determine levels of staff education. Tennessee does not have a linked database on staff professional development. In each of these QRIS, additional data on observed quality measures (ERS data) came from separate databases maintained by an external entity that conducted or scored the observations.

Time period. For Florida, Miami-Dade and Illinois, we obtained historical data on all participants given that ratings are valid for more than one year. In these QRIS databases, centers may have more than one record to document annual maintenance. We selected the most recent record that included a rating level, an ERS score, and data on quality components for the same rating period. For Tennessee, where quality rating assessments are conducted yearly, we analyzed data only from the most recent year of data.

[2] Kentucky was about to launch an evaluation of their own and could not participate at the time.

[3] Additional information on the sample, data sources, and timeframe for analysis is included in Appendix A.

Table I.1. Summary Characteristics of QRIS Included in Secondary Data Analysis

	Miami-Dade, Florida	Illinois	Tennessee
QRIS name	Quality Counts	Quality Counts	Star-Quality Child Care Program
Oversight agency	Early Learning Coalition of Miami-Dade/Monroe	Illinois Department of Human Services	Tennessee Department of Human Services
Year of full implementation	2008	2007	2001
Number of rating levels	5	4	3
Structure of rating levels	Combination	Building Block	Combination
Validity period of rating level	Initial: 2 years Subsequent: 13 months	1 year	1 year
Voluntary	Yes	Yes	Yes; all licensed providers must participate in Report Card program but progression beyond the Report Card is voluntary
Providers included in the analysis	Center-based including participating Head Start programs	Center-based including participating Head Start programs	Center-based, including Head Start programs, that voluntarily participate beyond the Report Card program
Analysis time period	October 2008–September 2010	April 2008–August 2010	October 2009–September 2010
Sample size	253[a]	350	1,369

Source: Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); secondary data analysis conducted as part of the QRS Assessment project.

[a] As the newest QRIS included in the analysis, 137 of the 253 centers had initial baseline ratings that are re-evaluated after 2 years..

C. Analytic Approach, Study Scope, and Limitations

Using the framework of the quality components as defined in the Compendium and the parameters presented by the administrative data available, we specified the quality component variables, quality outcome (observed quality), and a categorization for the rating levels to form the basis for the analysis. We first describe how each was accomplished and then detail the analytic methods used to address the four research topics. The section concludes with a discussion of the study limitations.

Quality rating components. We created common measures across the QRIS, generally one variable for 10 of the 13 quality component categories identified in the Compendium.[4] Where possible, we retained the continuous numeric values for maximum variability in conducting univariate and bivariate analyses and then created binary variables for use in the multivariate analysis. We also established meaningful cut-points to develop categorical variables. Table I.2 lists the quality rating components defined along a common metric based on the data available. Appendix B details key decisions or site-specific considerations used in constructing these cross-QRIS quality component variables. Appendix C includes the standards for child care centers in each of the three QRIS.

Table I.2. QRS Assessment Project Secondary Data Analysis Variables, by Quality Category

	Quality Component Analytic Variable		
Quality Category	Label In Analysis	Description	Values
Child-to-staff ratio for an age group	Ratio for an age group	Children per teacher ratio for an age group	Continuous or: 3-year-old ratio: -8:1 or lower -9:1 or higher 2-year-old ratio: -6:1 or lower -7:1 or higher
Curriculum	Curriculum use	Use of curriculum or standards	Yes/No
	Curriculum staff training	Staff trained on the curriculum or standards	All teaching staff Some teaching staff Not all staff
Child assessment	Assessment use	Use of an assessment tool or screening process	Yes/No
	Assessment guide planning	Use of assessment or screening results to individualize activities	Yes/No
Provisions for special needs	Planning for special needs	Activities or coordination with specialists for children with special needs	Yes/No
Cultural and linguistic diversity	Communicate in family primary language	Staff/resources available to communicate with parents in primary home language other than English	Yes/No
Staff qualifications	Director education	Center director has some college or higher	Yes/No
	Teaching staff education	Percentage of teachers and assistant teachers with some college or higher	Continuous or: Less than 25 percent 25 to 50 percent 50 percent or higher

[4] Data from the three QRIS did not support three component categories: (1) licensing was not included as all centers were licensed; (2) health and safety was not evaluated in any of the three, except for the highest level in Illinois for CPR and first aid certification, which is generally found in licensing regulations; and (3) community involvement standards, such as sharing resources, which were embedded within family partnerships for one QRIS and could not be disentangled as a distinct item.

Table I.2 (continued)

	Quality Component Analytic Variable		
Quality Category	Label In Analysis	Quality Category	Label In Analysis
Family partnerships	Family partnerships	Percentage of evaluated activities present (e.g., parent-provider conferences, program evaluation, education activities)	Continuous or: All Less than all
Administration/ Management	Staff management	At least one management aspect (staff meetings, annual staff performance evaluations, orientation, or staff development plans)	Yes/No
	Salary scale differentiated	Salary differentiation by individual's education, training, or experience	Yes/No
	Staff benefits	Percentage of staff benefits provided of those evaluated (e.g., vacation/sick leave, health insurance, tuition benefit)	Continuous or: All Less than all
	Program management	Percentage of program management strategies present of those evaluated (e.g., risk management plan, business plan, financial records))	Continuous or: All Less than all
Accreditation	Accreditation	Accredited center per QRIS approved list of accrediting bodies	Yes/No
Learning Environment	ERS scores: ITERS-R total ECERS-R total ECERS-R Teaching and Interactions factor ECERS-R Provisions for Learning factor ERS combined score (ITERS-R + ECERS-R)	Center-level averages across all classrooms	Continuous (0 - 7)

Source: Secondary data analysis conducted as part of the QRS Assessment project.

Observed quality. Observed quality (as measured by the ERS) can be found as a quality component in many QRIS. The analysis was limited to existing data collected by QRIS. Lacking an exogenous measure of quality, ERS data collected and maintained by each of the three QRIS as part of the rating process was necessary to use as an outcome. As an outcome, observed quality as represented by the ERS is included to examine the correlation and association of all other quality rating components to observed environment quality scores as an objective measure of quality, commonly used in research (for example, Cost, Quality, and Outcomes Study Team 1995; Early et al. 2005; Love et al. 2003; Mulligan and Flanagan 2006; Ross et al. 2008; West et al. 2007).

We modeled the continuous scale scores, averaged across each center's classrooms. This supported an examination of the full variation in scores that is important given the concern among researchers about using scores as thresholds in determining quality rating levels (Tout et al. 2009). Paralleling the typical approach in QRIS, the analysis focuses on the combined ITERS-R + ECERS-R total scores. The average of these scales provides a single composite of quality at the center level for all ages served. We also used the ITERS-R and ECERS-R total scores. In all three QRIS, the

parent and staff subscale items were not collected; QRIS representatives noted this aspect was generally covered by another component of the rating process. Lastly, we calculated the ECERS-R for the two strongest factors in the National Center for Early Development and Learning (NCEDL) Multi-state Study of Prekindergarten—Teaching and Interactions (the quality of teacher-child interactions) and Provisions for Learning (materials available in the classroom and the arrangement of the classroom space) for use in the analysis (Clifford et al. 2005). These factors reliably assess the areas of classroom quality most proximal to learning.

Quality rating levels. The three sites vary in the number of quality level ratings they assign (as shown in Table I.1). It was not feasible to perfectly map the requirements for each level across the sites into a common rating level rubric for the analysis. For example, the minimum ERS score required for the base level for learning environment varies from 3.0 to 4.0. And, such a common rubric that is system-driven rather than empirically derived may render itself meaningless to any one particular QRIS. Nonetheless, a grouping of levels across the sites was of interest in order to examine patterns in the prevalence of quality components within and between levels across the QRIS. We collapsed the rating levels into three groups—base, middle, and high—mapping the levels from each of the three QRIS as shown in Table I.3. The summary levels indicate where providers are along the spectrum of quality as defined within each site.

Table I.3. Quality Rating Level Categories Used in the Secondary Data Analysis

QRIS	Base Level	Middle Level	High Level
Florida, Miami-Dade	Level 1	Levels 2 and 3	Levels 4 and 5
Illinois	Level 1	Level 2	Levels 3 and 4[a]
Tennessee	Level 1	Level 2	Level 3

Source: Secondary data analysis conducted as part of the QRS Assessment project.

[a]One Level 4 center in Illinois was missing component-level data; therefore, no Illinois Level 4 centers were included in the analysis.

Analysis methods. We used a range of analytic strategies to examine quality measurement across QRIS components and rating levels. To explore variation in quality rating components (the first research topic), we conducted ***univariate analyses*** (means, ranges, percentages) of the individual components using the common measures we developed. We also used univariate descriptives to examine patterns of quality components to develop quality profiles based on select quality components meeting particular thresholds (the fourth research topic). ***Bivariate analyses*** (that is, tests of significance) provide descriptive information on the differences in quality components by quality rating level both across and within sites and between different types of child care centers (the second topic). ***Multivariate analyses*** provide analytic information on the unique contributions of quality rating components in predicting observed quality (the third topic). Here we employed regression models with observed quality as the outcome and the quality rating components as explanatory variables.

Limitations. As stated above, the objective of these analyses is to provide exploratory information that may inform future approaches to analyzing, validating, or refining QRIS. The use of administrative data and the selection of a small number of select QRIS have some inherent limitations:

- Three QRIS were selected given specific operating and administrative features to support data analysis. They differ from each other in longevity, rating levels, and exact quality components. While they provide a variation of approaches, they do not represent all QRIS operating in the U.S.
- The analyses are based on existing data, collected for administrative purposes. The data were not collected as part of a study designed to answer specific research questions relevant to QRIS quality measurement design and refinement. We address research topics from a descriptive perspective with exploratory findings that are correlational in nature. The findings from this analysis are not confirmatory in defining the components to measure to maximize state or county resources and still achieve reliable and valid overall quality ratings. In time, these findings in combination with other current and future work will contribute to a body of research that can be used to make specific programmatic and policy-relevant decisions.
- Observed quality as measured by the ERS is used as an outcome because of its availability within existing data, but in these three and many other QRIS, it is a component contributing toward the overall rating level. We use it here as an objective measure of the setting, separate from provider reports of activities and interactions.[5] The ERS currently represents one of the standard observation quality measures employed in the child care field. While widely used in research, some concerns have emerged about the validity of the ERS in more high-stakes accountability situations such as QRIS ratings tied to funding and financial incentives, when the limited evidence available suggests that the measures of quality currently in use (such as the ERS) are not detecting the dimensions of the care environment most closely associated with children's outcomes (Burchinal et al. 2009). Other or additional measures, such as the Classroom Assessment Scoring System (CLASS™; Pianta et al. 2008), that focus on teacher-child interactions and instructional aspects are included in a small number of QRIS.

Other limitations stem from analytic decisions that were made to accomplish a cross-QRIS analysis:

- These analyses pertain to center-based care only. The observation measures for family child care differed by version (FDCRS versus FCCERS-R) such that they could not be combined across the three QRIS.
- In defining and modeling quality components across QRIS, site-specific components may be muted. That is, the cross-QRIS quality components are often summative in nature, capturing multiple site-specific QRIS indicators within a broader measure. A finer-grained analysis into raw data within one or across multiple sites may find differences in the associations with observed quality as the robustness of the individual component measures increases. In addition, not every quality component indicator within each site could be included.

[5] The ERS includes physical space, materials, health and safety, and social interactions, but also requires that observers follow up with teachers/providers to ask about key items that could not be observed or to fully score an item. This requires item-level self-report from teachers / providers and may introduce bias in the scoring of some items.

Lastly, the scope of the analysis also limited the work to a focus on the rating components. The analysis does not include variables pertaining to the improvement process (such as receipt and nature of funds, coaching, or technical assistance), which could be important in validation of the systems. The regression models also do not control for the variation in characteristics across center-based providers (such as number or ages of children served or the percentage of subsidized children served).

D. Roadmap to the Report

The following chapters (II-V) discuss the four research topics, in turn, presenting the goals of the analysis, analytic strategy, and key findings followed by a set of detailed tables that present the results. The final chapter summarizes the lessons from this work and recommends steps for continued research moving forward.

II. QUALITY COMPONENTS: DESCRIPTIONS ACROSS THREE QRIS

A. Analytic Goals

This descriptive analysis examines the emphasis states place on each quality component in the QRIS rating process and their prevalence across the three QRIS. The analysis addresses the following questions:

- What percentage of center providers obtain the highest quality rating level and what are the average ERS scores across providers?
- What quality rating components are typically included (and, therefore, are measured or evaluated) in the rating process across QRIS?
- What is the prevalence of quality rating components as measured on a common metric across QRIS? For example, what percentage of center providers meet each quality rating component (e.g., curriculum use)?
- How does the prevalence of quality rating components differ between the base and highest quality rating levels across QRIS?

B. Analytic Strategy and Considerations

Certain key decisions in the analytic strategy for this analysis are important to consider in understanding and interpreting the key findings. They include:

- Each quality rating component is structured in line with the common metric across the three QRIS, as described in Chapter I.
- To distinguish between the use of the quality component in the rating process and the known presence of that component, we present descriptive statistics to denote whether each component is:
  - **Demonstrated:** Yes, or other value that indicates the presence and/or level of the quality component (such as "yes" curriculum is used, or "all teaching staff" are trained on the curriculum)
  - **Not-demonstrated:** Yes/no, or other value that indicates the lack of presence and/or low level of the quality component (such as "no" curriculum, or "not all teaching staff" are trained on the curriculum)
  - **Not evaluated:** Indicates that the quality component is not included by the QRIS as a component that contributes to the rating and therefore is not measured or evaluated (at the particular level or within the QRIS as a whole)
- The sites vary in the number of levels with differing requirements for achieving those levels. For cross-QRIS analysis, we collapsed the levels into three categories—base, middle, and high, as described in Chapter I, Table I.3.
- The distribution of centers by the quality rating levels also differ within the three QRIS and are presented here in the aggregate. In particular, the majority of centers comprising the highest QRIS level across sites come from one QRIS—Tennessee, which is the most mature of the three analyzed; the majority of child care centers in this QRIS have achieved the highest level.

C. Key Findings

- Across the three QRIS, just over two-thirds (69 percent) of center-based providers have achieved the highest QRIS rating level. The most mature (and largest) QRIS included in this analysis has a large majority of providers at the highest rating level (81 percent, data not shown).
- Seven quality components are demonstrated by more than half of all center-based providers across the three QRIS. They are:
  - Use of a curriculum based on early learning guidelines or standards
  - Director with some college or higher level of education
  - Majority (50 percent or more) of teaching staff that have some college or higher level of education
  - Implementation of the full range of family partnership activities specified by the QRIS
  - Implementation of staff management practices (such as staff meetings, annual staff performance evaluations, orientation, or staff development plans)
  - Use of a salary scale that differentiates pay by an individual's education, training, or experience
  - Offering of the full range of staff benefits specified by the QRIS
- Across all rating levels, all center-based providers in these QRIS are evaluated on staff qualifications (represented by director and teaching staff education level) (There is no "not evaluated" category for this component.)
- Four components reflecting the degree to which providers individualize services to meet the needs of children and families are not often evaluated—more than 50 percent of centers across the three QRIS are not evaluated on these components. They are (1) conducting child assessment and screening, (2) use of child assessment results to guide planning, (3) communicating with families in their home language, and (4) planning for children with special needs. Program management is a fifth component that is not often evaluated for center-based providers across the three QRIS (the not-evaluated category is over 50 percent).
- Center-based providers at the high rating level have observed quality and the presence of quality components that are higher, on average, than those at the base level. For example, the average combined ERS score (ECERS-R+ITERS-R) for providers at the high level is 5.1 compared to 3.3 for providers at the base level.
- Even when centers have choices for which components to focus on in order to reach the high QRIS level, with only one exception, the substantial majority at the high level demonstrate the required level of evaluated quality components. The one exception is that only 34 percent of centers at the high level indicate that they use child assessment to guide planning.

D. Detailed Findings[6]

1. What percentage of center providers obtain the highest quality rating level and what are average ERS scores across providers? [Figures II.1 and II.2, and Table II.1 for all]

- Across the three QRIS, 69 percent of centers achieve the highest QRIS level(s), 28 percent meet the middle level(s), and 4 percent meet the requirements for the base level (Figure II.1).[7]
- Across the three QRIS, the observed quality approaches the developer-definition of "good" quality[8] on average, ranging from 4.4 to 4.8 points across the center average for the ITERS-R, ECERS-R, and combined ERS (ITERS-R + ECERS-R) total scores. (Figure II.2)
- ECERS-R factor scores (Clifford et al. 2005) averaged 5.5 for Teaching and Interactions and 4.9 for Provisions for Learning (Figure II.2).

2. What quality rating components are typically included (and, therefore, are measured or evaluated) in the rating process across QRIS? [Figures II.3 through II.5 and Tables II.1 and II.2]

- These three QRIS all evaluate director and staff education levels (some college or more) at each rating level (there are no "not evaluated" values in any of the three QRIS for these components of staff qualifications).
- They each also evaluate child-to-staff ratios; Illinois does so as part of licensing only, while Miami-Dade and Tennessee evaluate these ratios specifically in determining the QRIS rating level.
- The three QRIS also commonly evaluate the presence of approaches to **managing family partnerships, staff, and center operations** (Figure II.3 and Table II.1):
  - Approximately 86 percent of centers are rated on a family partnerships component across the three sites.
  - Staff benefits are commonly evaluated for centers across the three QRIS; nearly one in every four centers is evaluated for the provision of staff benefits in the determination of a rating level (26.6 not evaluated).
  - Program management represents a less common quality component, evaluated in 17 percent of centers.
- Two additional administrative and management quality components are commonly evaluated across the three QRIS (Figure II.4 and Table II.1).

---

[6] Refer to Appendix E for additional detailed data tables.

[7] The base level represents the first QRIS level; the middle level includes the second level in Illinois and Tennessee and the second and third levels in Miami); and the highest level includes level three in Illinois and Tennessee and levels four and five in Miami.

[8] In the ERS, a score of 5 is labeled as "good" on a 1 to 7 scale.

  - The presence of staff management practices (such as staff meetings, annual performance evaluations, or professional development plans) and the presence of a differentiated salary scale are evaluated in all but 14 percent of centers each.

- Across the three QRIS, curriculum quality components (use and staff training) are more commonly evaluated than other individualization components, with about 64 and 60 percent of centers evaluated, respectively (Figure II.5 and Table II.1).
- Components that evaluate conducting and using child assessments, planning for children with special needs, and communicating with families in their home language are not widely used across the three QRIS with substantial majorities of centers not evaluated on these components (Figure II.5 and Table II.1).

We see differences in which quality components are evaluated between the highest and base QRIS levels (Table II.2).

- Components that relate to observed quality (ERS scores) and director and staff education levels are evaluated for centers at base, middle, and high levels across the QRIS (and therefore are not included in Table II.2)[9].
- Similar percentages of centers are evaluated on the use of a standards-based curriculum between the base and high QRIS rating level. However, at the high level, emphasis is placed not just on the use of the curriculum, but also the training of staff on the curriculum. Specifically, the percentage of centers not evaluated on curriculum staff training at the high QRIS level is less than the percentage of centers at the base QRIS level. About 17 percent more centers are evaluated on this component at the high level.
- Centers at the high QRIS level are less often evaluated on components for assessment, planning for special needs, and communication in the primary language of families (not evaluated in 94 versus 64 percent of centers). The differences likely reflect that the majority of centers comprising the high level are located in one QRIS (Tennessee) that does not evaluate these components.
- The percentage of high QRIS level centers evaluated on the presence of strategies related to family partnerships, staff benefits, and staff management, as well as the presence of a differentiated salary scale, is higher than the percentage of centers at the base QRIS level (ranging from an 18 to 26 percentage point difference).
- The high QRIS level centers are less often evaluated on the presence of strategies related to program management than centers at the base level (not evaluated in 94 versus 64 percent, respectively).

[9] ERS observations were not conducted for all Illinois centers (given an alternate pathway for accredited centers). Similar to ratios, descriptives were calculated for only those centers evaluated.

3. What is the prevalence of quality rating components as measured on a common metric across QRIS?[10] For example, what percentage of center providers meet each quality rating component (e.g., curriculum use)? [Figures 11.3 through II.7; Table II.1 for all]

The quality components involving the staff who work with the children include ***child-to-staff ratios and staff education***:

- For all centers with classrooms for 2-year-olds and 3-year-olds[11], the average center ratio is 6 children per teacher and 8.5 children per teacher, respectively (Figure II.6). As a comparison, the National Association for the Education of Young Children (NAEYC) recommendations identify a ratio no higher than 6:1 for 2-year-olds and 9:1 for 3-year-olds based on group size; to achieve the highest quality rating level for ratios, centers must show ratios of 5:1 and 8:1, respectively in Tennessee and 6:1 and 9:1, respectively, in Florida, Miami-Dade.
- In terms of education, 77 percent of directors have some college or higher education backgrounds (Figure II.7).
- On average, 37 percent of the teaching staff (lead teachers and assistants) in a center have some college or higher. Just over one-half (56 percent) of centers have a teaching staff where 50 percent or more have some college or higher (Figure II.7).

- Approximately 70 percent of centers possess all of the family partnership strategies specified by the QRIS (Figure II.3).
- Administrative quality components include offering staff benefits and implementing program management practices. All of the staff benefits specified by the QRIS are offered by 69 percent of across the three QRIS. Only 8 percent of centers demonstrate that they conduct all the program management strategies that are evaluated by the QRIS (Figure II.3).
- The large majority of centers evaluated on staff management practices and the presence of a differentiated salary scale offer these quality components at 81 and 78 percent, respectively (Figure II.4).
- Twenty-one percent of centers in these three QRIS are accredited (Table II.1).

Individualization practices that a center implements to tailor services for children and families may be evident from quality components rating curriculum, assessment, provisions for children with special needs, and communication in the primary language of families (Figure II.5):

[10] The use of a common metric is described in Chapter I. Refer to Table I.2 for a description of the variables used across the three QRIS.

[11] Illinois does not include child-to-staff ratios as a specific QRIS quality component but relies on existing licensing requirements (i.e., 8:1 for 2-year-olds and 10:1 for 3-year-olds). Among the two states included in these estimates, 656 of 1,622 centers had 2-year-old classroom data available and 1,336 had 3-year-old classroom data available.

- Among the centers evaluated on curriculum quality components (use and staff training), most demonstrate curriculum use and at least some staff trained on the curriculum (at 60.5 and 49 percent, respectively).
- When child assessment is evaluated, we find that centers may conduct assessments but are not necessarily using the assessments to guide planning, with 13 and 4 percent of all centers demonstrating these components, respectively.
- Centers that are evaluated on planning for children with special needs and communicating with families in their home language tend to demonstrate the presence of these components.

4. How does the prevalence of quality rating components differ between the base and highest quality rating levels across QRIS? [Table II.3]

Centers in each of the three QRIS have some choice in which components to focus on to achieve the high QRIS level. Both Florida, Miami-Dade and Tennessee use combination rating structures that assign points by indicator or component that then sum up to an overall rating level (described in Chapter I). In Illinois, providers are assessed using the Program Administration Scale which assesses a series of item-level indicators across multiple quality components.[12] For each rating level, the Illinois QRIS sets cut-points for the PAS overall score, not individual item scores, so a provider may meet the overall cut score even if some individual indicators are not met. Even with this choice (implicit or explicit), in each QRIS there are differences in the prevalence of each quality component between the base and high level QRIS rating, with the substantial majority of centers at the high level demonstrating each component.

Table II.3 presents results of a bivariate analysis of the quality components between the base and highest quality rating levels across the three QRIS. Significance tests indicate whether differences exist in the prevalence of each quality component between the lowest and highest rating levels.

Providers at higher quality rating levels have significantly ***higher observed quality*** scores. Centers achieving the high QRIS level have average scores from 4.9 to 5.2 on the ITERS-R and ECERS-R scores as compared 3.0 to 3.6 among centers receiving the base QRIS level.

Centers, on average, have ***lower (or better) child-to-staff ratios and higher rates of more educated staff*** at the high QRIS level.

- For all centers with classrooms for 2-year-olds, the average child-to-staff ratio is 5.5:1 at the high level, compared to 7.2:1 at the base level. A similar difference exists for 3-year-olds (8.3:1 compared to 10.4:1).

[12] For the analysis, the PAS was used to create variables for the use of developmentally-appropriate curriculum assessment use and its role in planning, family partnerships, staff benefits, staff management practices, differentiated salary scale, program management practices, planning for children with special needs, and communicating with families in their home language.

- With regard to staff education, 86 percent of directors of centers at the high level have some college or higher education, compared to 35 percent at the base QRIS level.
- Additionally, while 74 percent of centers at the high QRIS level have at least one-half of its teaching staff possessing some college education or higher, only one percent of the base QRIS level do. More commonly the latter group have less than one-quarter of its teaching staff having some college or higher.

A greater percentage of high QRIS level centers than those at base levels demonstrate components related to ***family partnerships and administrative and staff management practices***.

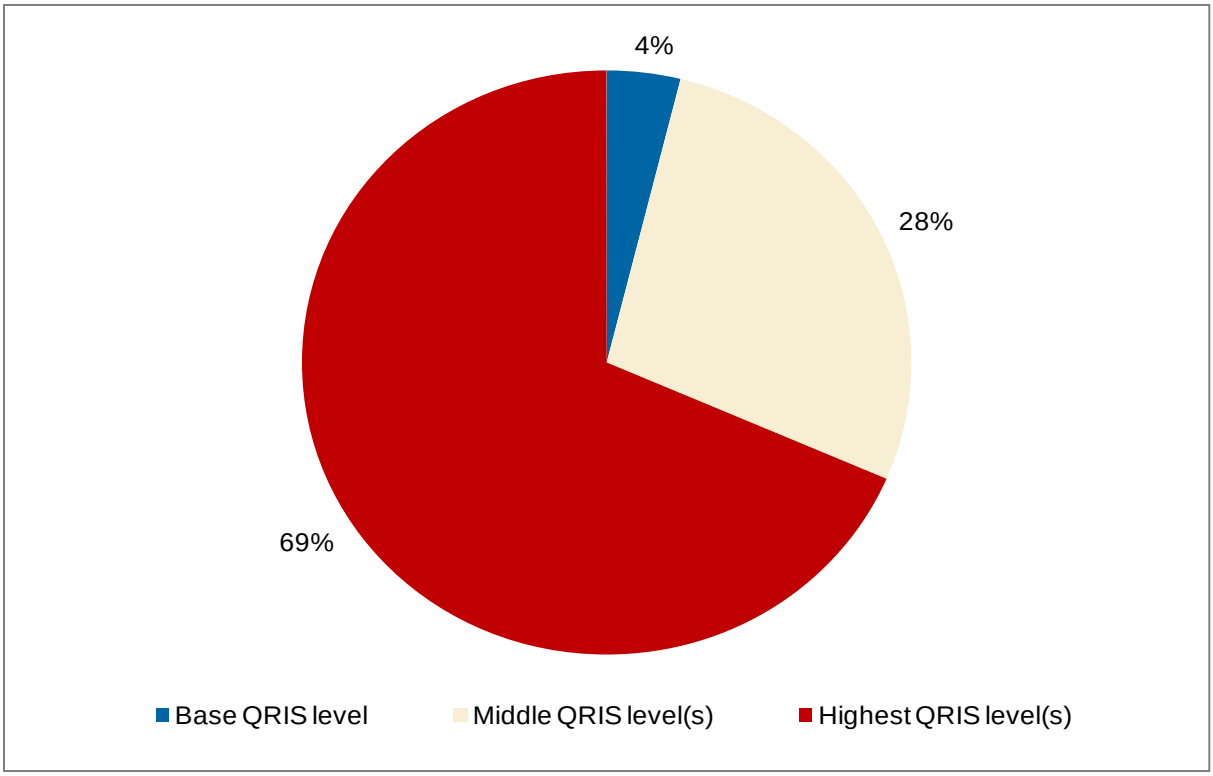
- By the high QRIS level, more than 9 in every 10 centers implement the full range of family partnership strategies specified by the QRIS while less than one in every 10 do so at base level centers.
- At the high level, nearly all centers offer the full range of staff benefits specified by the QRIS, have staff management practices in place (such as professional development plans), and use a differentiated salary scale.
- Whereas the high QRIS level centers are less often evaluated on program management strategies than those at the base level (discussed above), among those evaluated, none of the base level centers demonstrate all of the evaluated strategies compared to 72 percent at the high QRIS level.
- Nearly one-quarter (24 percent) of centers at the high QRIS level are accredited; at the base level, 4 percent are accredited.

A greater percentage of high QRIS level centers than those at the base level demonstrate the ***individualization components***.

- Over 90 percent of centers at the high level indicate that they use a standards-based curriculum and conduct staff training for all staff on the curriculum (99 and 92 percent, respectively). At the base level, a majority of centers use a standards-based curriculum (68 percent) but less than 20 percent of centers at the base level conduct training on the curriculum for some or all staff.[13]
- While a larger proportion of centers at the high level demonstrate the use of child assessments in planning (34 percent as compared to 11 percent), the gap between high and base level centers is smaller for this component than others.
- Among those evaluated, high QRIS level centers are more than twice as likely to communicate in the family's home language and more than three times more likely to demonstrate planning for special needs than centers at the base level.

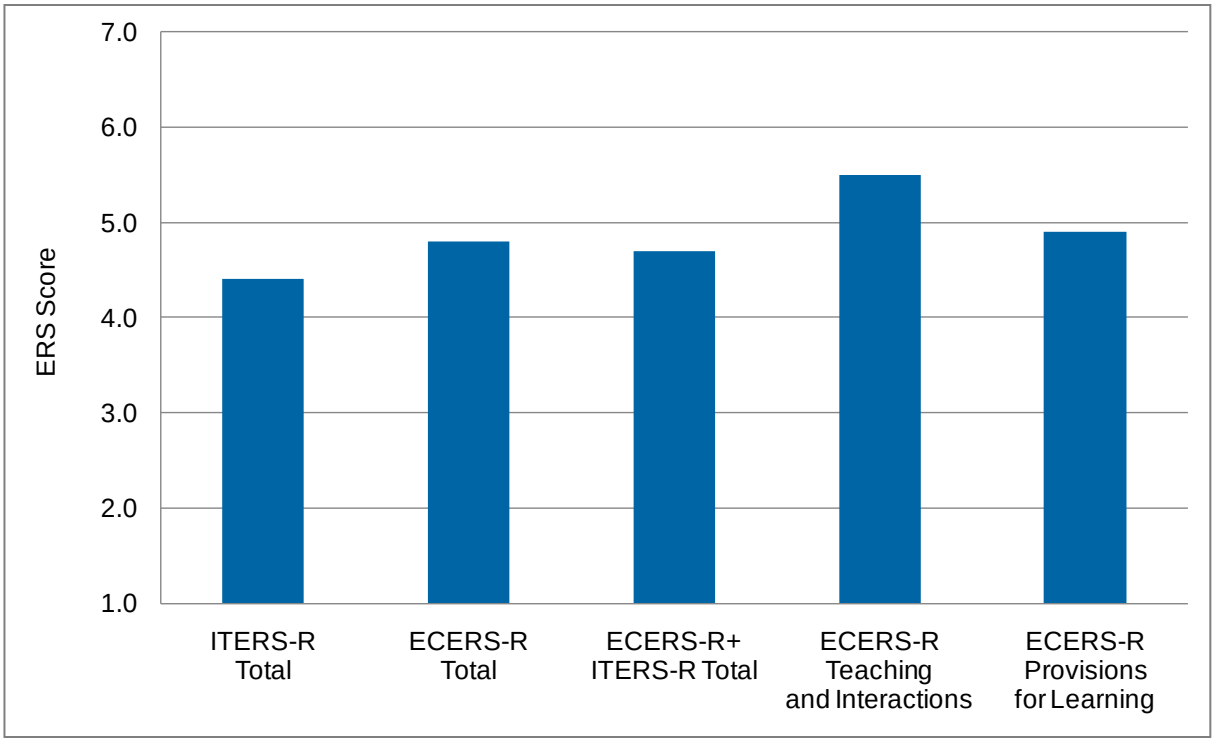
[13] In Miami-Dade, the use of an approved curriculum can be evaluated at any level but does not contribute points in assigning an overall rating level (at any level).

Figure II.1. Percentage of Center- based Providers by Rating Level Across Three QRIS



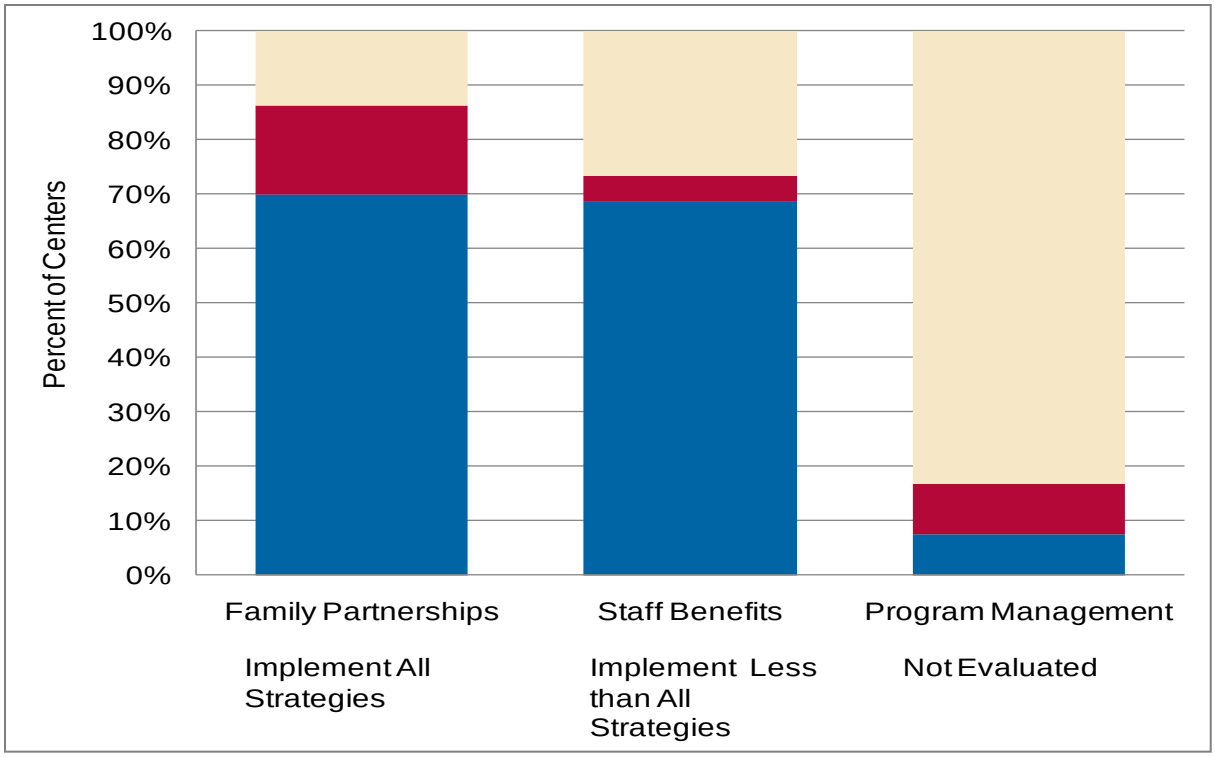
Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

Figure II.2. ERS Scores for Center- based Providers Across Three QRIS



Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

Figure II.3. Family Partnerships, Staff Benefits, and Program Management Quality Components Across Three QRIS



Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part ofthe QRS Assessment project.

Figure II.4. Staff Management and Salary Scale Quality Components Across Three QRIS

Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

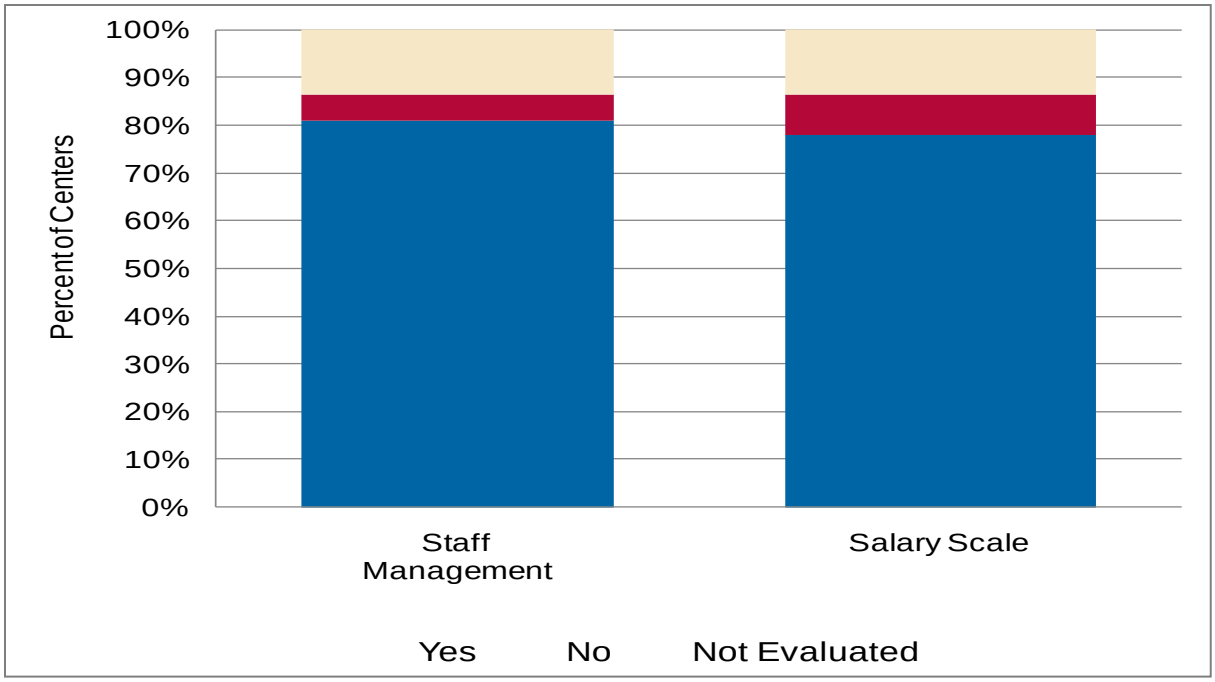
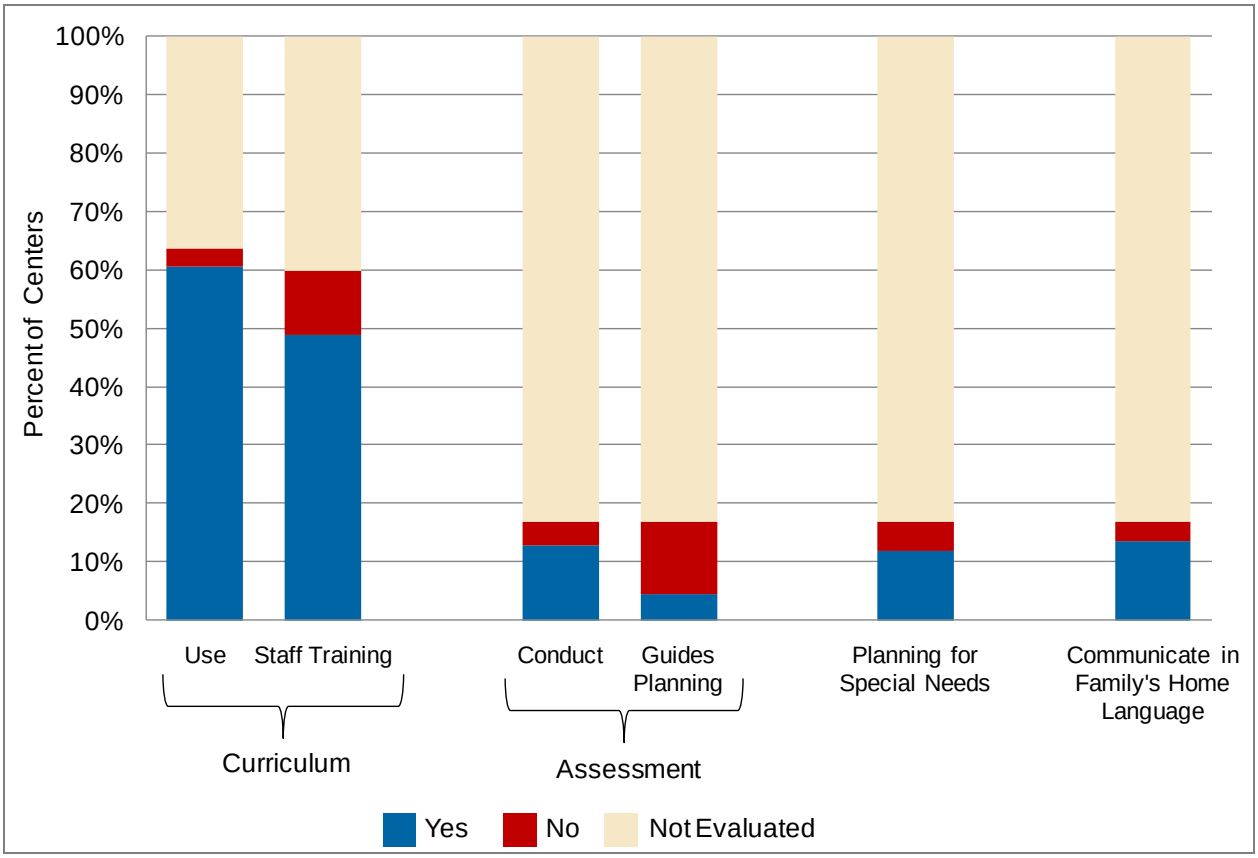
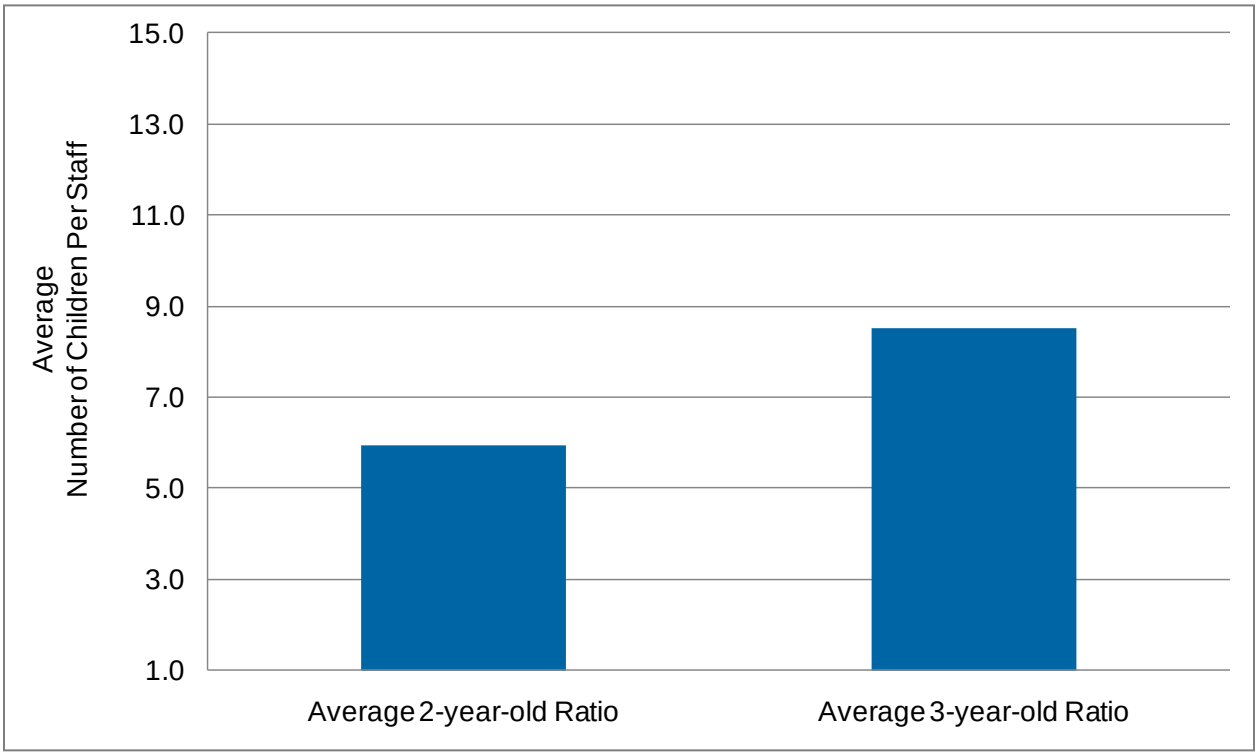


Figure II.5. Individualization Quality Components Across Three QRIS



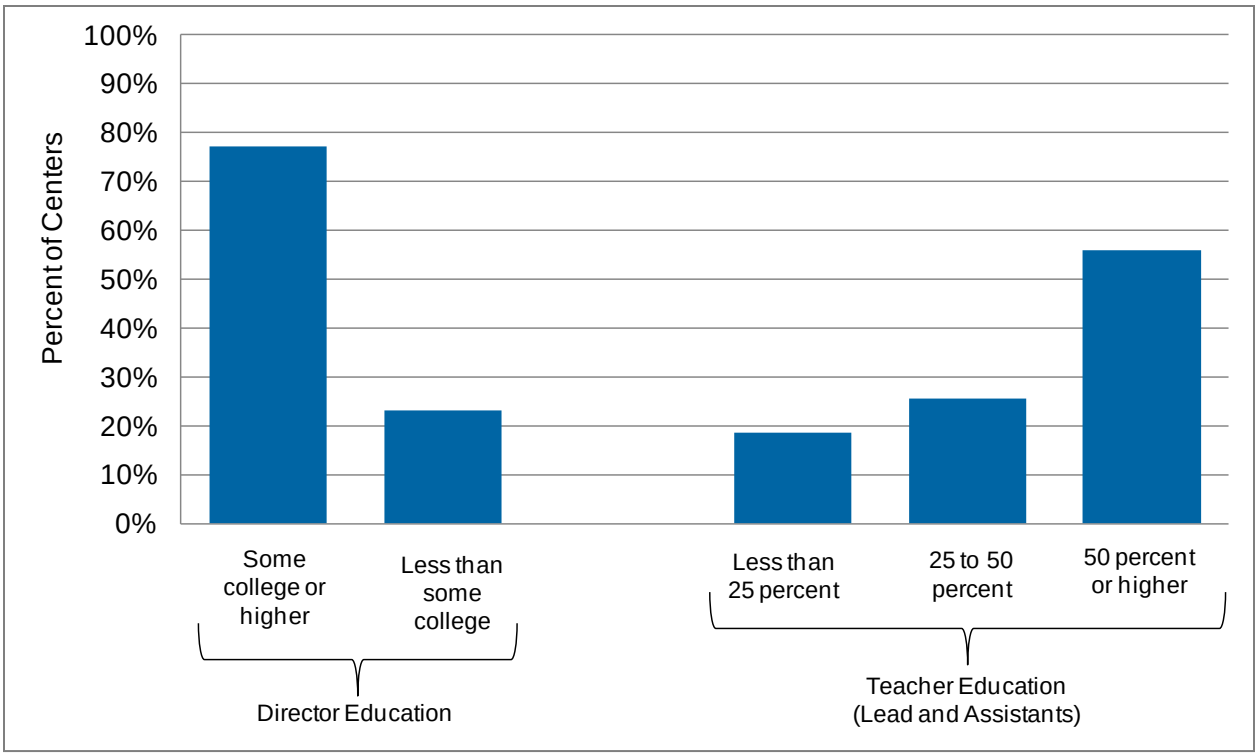
Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

Figure II.6. Child to Staff Ratios for Center- based Providers Across QRIS



Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

Figure II.7. Director and Staff Education in Center- based Providers Across Three QRIS



Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

Table II.1. Quality Rating Levels, Observed Quality, and Quality Components Across Three QRIS

	Percentage of Centers	Mean
QRIS overall rating level		
High level[a]	68.5	
Middle level[b]	27.6	
Base level	3.9	
Environment Rating Scale		
ITERS-R average total score		4.4
ECERS-R average total score		4.8
ECERS-R Teaching and Interactions		5.5
ECERS-R Provisions for Learning		4.9
ERS combined total score (ITERS-R + ECERS-R)		4.7
Center average 2-year-old ratio		5.9
Ratio 6:1 or lower	47.5	
Center average 3-year-old ratio		8.5
Ratio 8:1 or lower	49.2	
Director education		
Some college or higher	77.0	
Less than some college	23.0	
Teacher education (percentage some college or higher)		37.1
Less than 25 percent	18.6	
25 to 50 percent	25.4	
50 percent or higher	55.9	
Family partnerships		
Percentage of available		93.9
All offered	70.0	
Less than all offered	16.3	
Not evaluated	13.7	
Staff benefits		
Percentage of available		97.3
All offered	68.8	
Less than all offered	4.6	
Not evaluated	26.6	
Program management		
Percentage offered on average		75.2
All offered	7.6	
Less than all offered	9.2	
Not evaluated	83.2	
Staff management		
Yes	80.8	
No	5.5	
Not evaluated	13.7	
Salary scale differentiated		
Yes	77.8	
No	8.5	
Not evaluated	13.7	
Accredited center	21.0	
Curriculum/standards used		
Yes	60.5	
No	3.3	
Not evaluated	36.3	

Table II.1*(continued)*

	Percentage of Centers	Mean
Curriculum/standards-staff training		
All teaching staff	45.9	
Some teaching staff	2.9	
Not all teaching staff	10.9	
Not evaluated	40.3	
Child assessment/screening		
Yes	12.9	
No	3.9	
Not evaluated	83.2	
Child assessment guide planning		
Yes	4.4	
No	12.4	
Not evaluated	83.2	
Communicate with families in home language		
Yes	13.6	
No	3.2	
Not evaluated	83.2	
Planning for special needs		
Yes	11.8	
No	5.0	
Not evaluated	83.2	
Number of Centers	1972	

Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

[a]High QRIS level combines Florida, Miami-Dade Quality Counts levels 4 and 5. Illinois level 4 did not have any cases in the sample such that level 3 represents the highest QRIS level.

[b]Middle QRIS level includes Florida, Miami-Dade Quality Counts levels 2 and 3.

Table II.2. Percentage of Centers That Are Not Evaluated on Quality Rating Components At Base and High QRIS Rating Levels Across Three QRIS[a]

	Percentage of Centers[a]	
Quality Component	Base level	High level
Curriculum/standards used Not evaluated	42.9	38.7
Curriculum/standards-staff training Not evaluated	57.1	40.2**
Child assessment/screening Not evaluated	63.6	93.5***
Child assessment guide planning Not evaluated	63.6	93.5***
Communicate with families in home language Not evaluated	63.6	93.5***
Planning for special needs Not evaluated	63.6	93.5***
Family partnerships Not evaluated	29.9	11.3***
Staff management Not evaluated	29.9	11.3***
Staff benefits Not evaluated	52.0	16.3***
Salary scale differentiated Not evaluated	29.9	11.3***
Program management Not evaluated	63.6	93.5***
Number of centers	77	1350

Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

[a]Base QRIS level is the first level for all three QRIS. High QRIS level combines Florida, Miami-Dade Quality Counts levels 4 and 5. Illinois level 4 did not have any cases in the sample such that level 3 represents the highest QRIS level.

[a]The "Not evaluated" category is the percentage of centers who were not evaluated (the numerator) relative to all centers at the base or high rating level (the denominator).

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table II.3. Percentage and Means (SDs) of Centers That Meet Quality Rating Component Criteria, At Base and High QRIS Rating Levels Across Three QRIS[a]

	Percentage of Centers		Mean (SD)			
Quality Component	Base level	High level	Base level		High level	
Environment Rating Scale						
ITERS-R average total score			3.0	(0.8)	4.9	(0.8) ***
ECERS-R average total score			3.6	(0.8)	5.2	(0.7)***
ECERS-R Teaching and Interactions			4.4	(1.0)	5.8	(0.8) ***
ECERS-R Provisions for Learning			3.5	(0.8)	5.3	(0.9)***
ERS average total score (ITERS-R + ECERS-R)			3.3	(0.8)	5.1	(0.7) ***
Center average 2-year-old ratio			7.2	(2.2)	5.5	(1.0) ***
Ratio 6:1 or lower	30.8	86.8***				
Center average 3-year-old ratio			10.4	(2.7)	8.3	(0.7) ***
Ratio 8:1 or lower	19.4	73.3***				
Director education (some college or higher)	35.1	85.8***				
Teacher education (percentage some college or higher)			10.6	(11.3)	43.6	(11.7)***
Less than 25 percent	83.1	2.3***				
25 to 50 percent	15.6	23.7				
50 percent or higher	1.3	74.0***				
Family partnerships						
Percentage of available			51.7	(30.0)	98.6	(6.5) ***
All offered	7.4	93.7***				
Staff benefits						
Percentage of available			68.2	(30.4)	99.3	(5.4) ***
All offered	32.4	98.1***				
Staff management						
Yes vs. No	40.7	99.7***				
Salary scale differentiated						
Yes vs. No	22.2	98.5***				
Program management						
Percentage offered on average			27.5	(25.7)	92.5	(12.7) ***
All offered	0.0	71.6***				
Accredited center	4.1	23.6***				
Curriculum/standards used						
Yes	68.2	99.0***				
Curriculum/standards-staff training						
All teaching staff	12.1	91.7***				
Some teaching staff	6.1	2.9				
Not all teaching staff	81.8	5.5***				
Child assessment/screening						
Yes vs. No	32.1	97.7***				
Child assessment guide planning						
Yes vs. No	10.7	34.1*				

Table II.3*(continued)*

	Percentage of Centers		Mean (SD)	
Quality Component	Base level	High level	Base level	High level
Communicate with families in home language				
Yes vs. No	42.9	97.7***		
Planning for special needs				
Yes vs. No	25.0	93.2***		
Number of centers	77	1350		

Source: Analysis of data from Illinois, Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

[a]Base QRIS level is the first level for all three QRIS. High QRIS level combines Florida, Miami-Dade Quality Counts levels 4 and 5. Illinois level 4 did not have any cases in the sample such that level 3 represents the highest QRIS level.

Note. For the comparison of quality rating components between the high and base QRIS level centers, we conducted t-tests for the presence of quality components (percentage) or level (mean) for centers who are evaluated on the component (that is, "Not evaluated" is treated as missing).

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

III. QUALITY COMPONENTS: DIFFERENCES BETWEEN RATING LEVELS AND TYPES OF CENTERS

A. Analytic Goals

In this chapter, we use bivariate analyses to explore the following:

- How does the prevalence of quality rating components differ between quality rating levels within a site?
- What similarities exist by quality rating level in the prevalence of quality rating components across QRIS?
- What is the prevalence of quality rating components by different types of centers (such as Head Start status or accreditation status)?

A preliminary step in validating a QRIS is to look at the providers' distribution of quality components at different rating levels and assess whether there are differences that make levels distinct. This question is particularly noteworthy for QRIS that rely on points earned for each quality component to determine a rating level given that participating centers may achieve different levels in different ways. To examine the degree to which rating levels may be distinct, we conduct a bivariate analysis using data from Florida, Miami-Dade and Tennessee (both combination systems that use points to determine rating levels) to determine the prevalence of quality rating components among centers within each QRIS site between base, middle, and highest rating levels, and between sites for a given quality rating level. For example, we take each quality rating component, such as director education and test for significant differences in the average percentage of center directors with some college or higher between the base and middle rating levels within each state (and repeat for the middle to high levels). We also test for differences between the two states at each the base, middle, and high rating levels separately.

QRIS developers are faced with numerous design decisions. Among them is whether Head Start programs or accredited centers should be treated differently than others within the QRIS. For example, some QRIS have created alternative pathways for these centers to enter at a particular rating level without having to go through the entire process. Instead, documentation of accreditation or meeting Head Start program performance standards is accepted as evidence of complying with QRIS standards at a specified rating level (typically the highest one or two levels). The cross-walk of these other standards with QRIS standards has not always been well documented. To examine whether there are differences between centers based on Head Start or accreditation status, we conduct a bivariate analysis using data from Florida, Miami-Dade to examine the prevalence of quality components across different types of centers.

B. Analytic Strategy and Considerations

In reviewing the key findings for the differences in quality components across levels within sites and by center type, we must consider the following:

- Bivariate findings only pertain to two QRIS sites—Florida, Miami-Dade and Tennessee. We exclude Illinois centers from this analysis because only a subset of centers are required to receive an ERS and PAS assessment and there were not a sufficient number of cases with complete data. See Appendix D, Methodological Notes, for additional details.

- The analysis uses the common metric for defining the quality components and a common categorization of rating levels that does not necessarily directly map to either QRIS included in the analysis. For example, these two sites vary in the number of levels and requirements for achieving those levels. For comparative purposes, we have defined the middle level to include level two in Tennessee and levels two and three in Florida, Miami-Dade; the highest level includes level three in Tennessee and levels four and five in Florida, Miami-Dade.
- We conducted statistical tests to determine whether estimated differences between centers in base versus middle and middle versus highest QRIS levels are larger than we would expect by chance. We used t-tests to compare means and chi-square tests to compare distributions on categorical variables.
- The analysis of quality rating levels and components by Head Start program status focuses on centers from Florida, Miami-Dade. The analysis is limited to this QRIS because it is the only one in which a reliable designation for Head Start status exists within the administrative data. In addition, all Head Start programs in Florida, Miami-Dade participate in the QRIS, thus producing a sufficient sample for analysis.[14]
- The analysis of quality rating levels and components by accreditation status is also limited to centers within the Florida, Miami-Dade QRIS (53 percent or 135 centers in are accredited). In Illinois, accreditation is an alternate pathway into the QRIS such that most accredited centers are not evaluated on the various quality components. In Tennessee, few centers are accredited (7 percent or 102 centers), but most participating providers throughout Tennessee have achieved the highest quality rating level. The cross-site analysis was, therefore, skewed in important ways; the Miami-only analysis is more equitable in the treatment of accreditation status across all providers in the QRIS.
- The context in Miami-Dade with regard to accreditation is important to understand since the analysis is limited to this QRIS. Miami-Dade's QRIS recognizes accreditation by Florida's Gold Seal program. The Gold Seal program is administered independently of Quality Counts, with its own application and review process for both accrediting agencies and individual child care providers (Florida Department of Children and Families, 2011). The Gold Seal list currently has 13 accrediting entities, including NAEYC and the National Early Childhood Program Accreditation (NECPA) that are commonly accepted across QRIS (Caronongan et al. 2011). Nearly 40 percent of the accredited centers (50 of the 135) included in the analysis for Miami-Dade are accredited by NAEYC.

C. Key Findings

- Centers within each rating level are generally distinct in the prevalence of quality rating components. On average, centers in the middle level demonstrate quality components at higher rates than those at the base, and centers at the highest level demonstrate quality components at higher rates than those in the middle.

[14] Not all Head Start centers are represented in the current data; a few Head Start centers are in a self-study stage such that they have not had a rating assessment. Also noted in Appendix A, centers had to have quality component and ERS data to be included in the analysis sample.

- There are some differences between the two states in these patterns, however. With few exceptions, centers consistently demonstrate increasing rates in each of the quality components from one of the three levels (base, middle, high) to the next in Tennessee. In Miami-Dade, a number of quality components demonstrate a significant difference from only one level to the next, rather than between each of the three levels (for example, the prevalence of a quality component may significantly increase between the base and middle levels, but remain the same between the middle and high levels as is the case with curriculum staff training, the use of child assessments to guide planning, director education, and staff management strategies).
- The prevalence of quality components among centers at the same rating levels but in different QRIS is not entirely the same, but the differences that do exist tend to be smaller at the highest level (with the notable exception of teacher education). A companion report from our in-depth study of quality measurement found that, on the whole, the requirements across five QRIS start out wide-ranging at the base level and seem to come to greater commonality in the requirements for providers at the highest level (Caronongan et al. 2011). This finding is further reflected in the current analysis. However, licensing standards, subsidy reimbursement policies, professional development systems and supports, and simply the longevity of the QRIS, influence the prevalence of quality components between the two QRIS at each rating level such that significant differences exist.
- Head Start programs generally do not differ from other centers participating in the QRIS in Miami-Dade in ERS scores reflecting observed quality. This is likely due to the similarity in the distribution of Head Start and non-Head Start programs across the different rating levels. They are, however, more likely to have in place quality components that include individualization practices (such as those around curriculum, child assessment, planning for children with special needs, and communicating with families in their home language). Head Start programs are also more likely than other centers to have directors and teaching staff with some college or more, to demonstrate all family partnership strategies specified by the QRIS, and to use a differentiated salary scale. These differences are likely a reflection of the emphasis of certain quality components in Head Start performance standards.
- Accredited centers are more likely to demonstrate many of the quality components over non-accredited centers. Similar to the patterns seen among Head Start programs, accredited centers are also more likely than other centers to demonstrate individualization practices, family partnership strategies, staff and program management strategies, and use of a differentiated salary scale. Unlike Head Start programs, accredited centers do not generally differ from other centers in the education levels of directors or teaching staff. Also in contrast to Head Start, accredited centers have higher ERS scores, on average, than other centers and are more likely to be at the highest rating level.

C. Detailed Findings

1. How does the prevalence of quality rating components differ between quality rating levels within Florida, Miami-Dade Quality Counts? [Table III.1]

Table III.1 presents results from two distinct bivariate analysis: (1) we conducted tests of significance between the prevalence and means of each quality component from the base to the

middle rating level, and (2) similarly, we conducted tests of significance between the prevalence and means of each quality component from the middle to the highest rating level. The first three columns include percentages of centers that meet the criteria for each component (the prevalence). The second set of columns presents the means and standard deviations for each component.

- Most centers in Miami-Dade are in the middle QRIS levels (that is, levels two and three): 17 centers are at base, 168 at middle, and 68 at high.
- On average, observed environment quality scores increase for centers at each higher QRIS level:
  - The combined ERS score, (aggregating ITERS-R and ECERS-R at the center level), for example, averages from 2.4 to 3.3 to 4.4 looking across providers rated at the base, middle, and highest QRIS levels, respectively.
  - At 5.4, the average ECERS-R Teaching and Interactions factor score for the highest QRIS level centers crosses the "good" threshold (as defined by instrument authors). The difference between this score and average scores at the base (mean = 3.1) and middle levels (mean = 4.4) is statistically significant.
- In Miami-Dade, the use of an approved curriculum can be evaluated at any level but does not contribute points in assigning an overall rating level (Caronongan et al. 2011). A large majority of centers (88 percent) demonstrate this component from the start. And, while the percentage of centers demonstrating use of the curriculum increases from the base level to the middle and middle to high, the differences are not significant.
- Centers at high and middle QRIS rating levels are more likely to implement individualization practices than centers at the base level. Since Florida, Miami-Dade uses a combination system, centers do not necessarily have to demonstrate individualization practices to achieve the highest rating level; however, we saw large percentages of centers at the two highest levels demonstrating these components.
  - Although curriculum use is common across all levels, there is a significant increase in the rates of curriculum staff training from the base to the middle level. At the base level, none of the centers had all lead teachers or all teaching staff receive such training. At the middle level, 27 percent of centers had trained all teaching staff on the curriculum. Centers at the highest levels have a larger percentage of all staff trained (34 percent) on the curriculum compared to centers at middle levels but the difference is not statistically significant.
  - None of the base QRIS level centers conduct child assessments or demonstrate planning for special needs. Few base level centers (6 percent) have resources to communicate with families in their home language. In comparison, centers in the middle QRIS level are significantly more likely to demonstrate these practices (ranging from 61 to 74 percent). Over 90 percent of centers at the highest QRIS level demonstrate these practices. Centers at the middle and highest levels show similar rates of using child assessments to guide planning (21 versus 25 percent).
- For 2-year-old classrooms, lower child-to-staff ratios emerge at the highest QRIS levels. A substantial majority of centers at the highest level in Miami-Dade are accredited (74 percent); these ratios are in line with NAEYC recommendations (NAEYC 2008). There are statistically significant differences between the middle and highest QRIS levels but not base versus middle. The opposite is evident for centers' average 3-year-old

classroom; centers with the base QRIS rating level differ from the middle level with an average of 12 children per teacher and 9 per teacher, respectively. Middle and high QRIS level centers do not differ; both average 9 children per teacher, matching NAEYC recommendations.

- Increases in the education levels of both directors and teaching staff seem to be greatest between the base and middle level. For example, there are statistically significant differences between centers at the base and middle QRIS levels that have directors with some college or more (24 versus 63 percent) but a marginally significant difference between middle versus high, with nearly three-quarters (74 percent) of directors having some college or more. A similar pattern of decreasing significance in the size of the differences between levels holds true for teacher education as well. Notably, while there is a significant decrease in the percentage of centers that have less than 25 percent of teaching staff having some college education or more between each of the levels, there is not a corresponding significant increase in the percentage of centers with 50 percent or more of the staff achieving this level of education. This could be a reflection of the relative newness of the Miami-Dade QRIS (implemented in 2008) since it takes time for education levels to increase.
- The use of all family partnerships strategies, a differentiated salary scale, and all program management strategies is low in base QRIS level centers, increases dramatically between the base and middle level and then increases steadily between the middle and high levels. Middle QRIS level centers are significantly more likely to offer these components, with one-half to two-thirds of centers doing so. In turn, centers receiving the highest QRIS levels exhibit these at even higher rates reaching 88 percent or higher.
- About half of centers at the base level use staff management strategies compared to 98 percent of centers at the middle levels and 100 percent of centers at the highest levels.
- Higher QRIS level centers are more likely to be accredited, steadily increasing in the percentage of accredited centers from 12 to 49 to 74 percent for base, middle, and highest QRIS levels, respectively.

2. How does the prevalence of quality rating components differ between quality rating levels within the Tennessee Star-Quality child care program? [Table III.2]

Table III.2 presents results from two distinct bivariate analysis: (1) we conducted tests of significance between the prevalence and means of each quality component from the base to the middle rating level, and (2) similarly, we conducted tests of significance between the prevalence and means of each quality component from the middle to the highest rating level. The first three columns include percentages of centers that meet the criteria for each component (the prevalence). The second set of columns presents the means and standard deviations for each component.

- In Tennessee, most centers participating in the QRIS meet the highest rating level (81 percent). Similar to Miami, there are higher observed environment quality scores among centers at higher QRIS levels. There are statistically significant differences in ERS scores of centers at the base vs. middle level, and middle vs. high level. Average scores for the highest level centers meet or cross the "good" threshold as defined by instrument authors.
- Tennessee's QRIS sets few requirements for individualization practices. They began assessing alignment between center curricula and Tennessee's Early Learning standards

in January 2010 so some of the centers in our sample were not evaluated on this component. Of centers that were evaluated, those at middle and high rating levels are more likely to demonstrate compliance with this requirement compared to centers at the base level. Similarly, the percentage of centers in which all teaching staff are trained on the curriculum rises steadily and significantly from the base to the middle and the middle to the high levels (16, 49, and 65 percent respectively).

- There are dramatic increases in the percentage of centers meeting lower child-to-staff ratios for both 2-year-old and 3-year-old classrooms between the middle and high levels (from 57 to 91 percent and from 40 to 76 percent, respectively). The differences in the percentages of centers meeting the lower ratios between the base and middle levels is not significant.
- The education levels of both directors and teaching staff in centers increase steadily across the three QRIS levels. Each higher QRIS level is more likely than the one below to have directors with some college or more and to have 50 percent or higher of teaching staff with a similar level of education. By the highest level, 98 percent of centers have directors with some college education or more and 88 percent of centers have 50 percent of higher of teaching staff with some college or more. Education levels among teaching staff are strikingly different with each higher level. None of the 26 centers at the base level have more than 25 percent of staff with college education or more, while virtually all centers in the highest level have at least 25 percent of teaching staff or more with this level of education.
- The availability of all family partnerships, staff management, a differentiated salary scale, and all staff benefits is low in base QRIS level centers (8 to 39 percent). Middle QRIS level centers are significantly more likely to offer these components (59 to 91 percent). In turn, nearly all of centers receiving the highest QRIS levels exhibit these (95 percent or higher).
- In Tennessee, centers that achieve accreditation receive bonus points toward the final calculation of the QRIS rating level. Overall, most centers do not appear to be accredited. While virtually none of base and middle QRIS level centers are accredited, 9 percent of centers at the highest level achieve this status.

3. What similarities exist by quality rating level in the prevalence of quality rating components across QRIS? [Tables III.3, III.4, and III.5]

In this section, we examine each QRIS level—base, middle, and highest—comparing Florida, Miami-Dade and Tennessee on the cross-QRIS quality rating components that they both evaluate (Tables III.3, III.4, and III.5).[15] These bivariate analyses using tests of significance examine the quality components at each of the rating levels across the two QRIS to understand how different contexts and emphases in components may influence the degree of comparability across systems.

- For observed quality, Tennessee centers have ERS scores that are significantly higher, on average, than those of centers in Miami-Dade. The difference in scores decreases

[15] Appendix C provides the QRIS requirements for each site for reference.

across quality rating levels—starting out at generally over 1.5 points different (across each of the five scores) at the base level to 0.8 points or less different at the highest level. QRIS requirements to receive the highest points for ERS scores begin much higher at the base level in Tennessee than in Miami-Dade (4.49 in Tennessee and 3.49 in Miami-Dade), but narrow at the highest level (5.0 in Tennessee and 5.5 in Miami-Dade).[16].

- Centers in Miami-Dade are more likely than those in Tennessee to use an approved curriculum at the base and middle levels because the use of curriculum is introduced into the QRIS standards at the base level in Miami-Dade and was only introduced in Tennessee halfway through the QRIS year we examined. However, by the highest level virtually all centers in both QRIS use an approved curriculum.
- Differences in the percentage of centers that demonstrate staff management supports, family partnership practices, and use of a differentiated salary scale that are seen at the base and middle levels between the two QRIS disappear at the highest level. Centers in Tennessee and Miami-Dade are essentially equally likely to demonstrate these three quality components at the highest level.
- While child-to-staff ratios are lower in Tennessee centers as compared to those in Miami-Dade across all three levels, the difference between the averages is lower at the highest level. The context of licensing as a foundation for QRIS contributes to these differences.[17] Miami-Dade's QRIS standards for child-to-staff ratio and group size at the base level are identical to Florida licensing standards while requirements at the highest rating levels narrow disparities with other states' licensing standards because they are based on NAEYC recommendations (Caronongan et al. 2011). At the base level the requirements for child-to-staff ratios for 2- year old classrooms in Miami-Dade and Tennessee are 15:1 and 9:1, respectively. At the highest level they are 9:1 and 8:1, respectively. The mean ratios for centers roughly mirror these requirements (except that the mean at the base in Miami-Dade is about 12:1, rather than the 15:1 requirement).
- Differences in the likelihood of centers having directors or greater proportions of staff with some college or more generally do not appear in the base and middle levels, but are substantial at the highest level. QRIS requirements for staff education are quite similar at both the base and highest levels in Miami-Dade and Tennessee. As noted earlier, it is possible that the relative newness of the QRIS in Miami-Dade can account for the lower education levels among center directors and teaching staff than those seen in Tennessee (a 10-year-old QRIS).
- Accreditation plays a similar augmentation role in the QRIS rating in Miami-Dade and Tennessee, such that it is not required but centers can either earn bonus points (in Tennessee) or a "plus" added to their rating (Miami-Dade). In Miami-Dade, centers that care for children receiving child care subsidies receive higher reimbursement rates when

[16] In this analysis, we have grouped the two highest levels for Florida, Miami-Dade. Learning environment level 4 requires a 4.5 ERS score whereas level 5 is the 5.5.

[17] The calculation of ratios differed across QRIS based on the data available which may also contribute to the differences detected. See Appendix B for details on variable construction.

they are accredited. This could contribute to the greater likelihood that a center is accredited in Miami-Dade than in Tennessee at each of the rating levels.

4. What is the prevalence of quality rating components in child care centers by Head Start status in Florida, Miami-Dade? [Table III.6]

In this section, we present results of a bivariate analysis of the rating level and quality components between participating QRIS centers in Miami-Dade that are Head Start programs and those that are not. The analysis examines the degree to which Head Start programs may be different from other centers in the prevalence and composition of quality components.

- A higher percentage of Head Start centers tend to achieve the highest QRIS levels than non-Head Start centers (a trend of 36 versus 24 percent, respectively). They are less likely to be at the base QRIS level; there are no Head Start programs at the base rating level.
- The poverty level in the communities in which centers are located do not differ significantly between Head Start and non-Head Start centers. Head Start centers are less likely to be of medium-size (serving 51-100 children) than other centers, and more likely to be very large (serving 151 children or more). While the data indicate that Head Start programs generally do not care for children that receive child care subsidies, this is a reflection of the funding stream than of the children served by Head Start programs. Head Start programs are targeted to low-income children but federal (and state) Head Start funding rather than child care subsidies supports children in the program.
- Observed quality is similar in Head Start and other centers for preschool classrooms as measured by the ECERS-R. Head Start centers (that include Early Head Start) have small, but significantly higher ITERS-R scores (3.7 versus 3.1 points) as well as combined ERS scores of ITERS-R + ECERS-R (3.7 versus 3.5 points) than those of other centers.
- Child-to-staff ratios differ between Head Start and other centers, but the pattern varies by age. Head Start centers have significantly lower ratios for 2-year-olds, on average 5:1 as compared to 7:1, and 80 percent of Head Start centers have a ratio of 5:1 or lower as compared to 18 percent of other types of centers. For 3-year-old classrooms, the average ratio is similar (between 9:1 and 10:1), but fewer Head Start centers have a ratio of 8:1 or lower (8 versus 39 percent).[18]
- While similar rates of centers, regardless of Head Start status, use an approved curriculum, Head Start centers are more likely to have all teaching staff attend training

[18] Head Start performance standards specify the number of children (that is, group size) and the number of adults rather than ratios. For 3-year-old classrooms, the group size is required to be between 15 and 17 children. Head Start classrooms must have at least two adults, with a third recommended. An approximate ratio for the largest group and required adults would be 8.5:1. As compared to NAEYC standards of 9:1 for a group size of 18, Head Start centers are less likely than non Head Start centers to meet that average ratio (20 versus 59 percent). Closer review of the data shows that over 80 percent of these Head Start centers had a ratio of 10:1 for classrooms labeled as 3-year-olds. The database assigns a specific age group to each classroom. Classrooms, therefore, could be mixed-age.

on the curriculum (62 versus 16 percent) and less likely to have fewer staff (not all lead teachers and/or assistants) trained on the curriculum (31 versus 79 percent).

- Ninety percent or greater of Head Start centers conduct child assessments, demonstrate individualization practices by planning for children with special needs, and communicating information in the family's primary language; these rates are all significantly higher than those of non-Head Start centers (64, 58, and 70 percent, respectively). Head Start centers are also more likely to document using assessments to guide planning.
- Head Start centers have more highly educated staff as compared to non-Head Start centers. Ninety-three percent of Head Start centers have a director with some college or higher as compared to 54 percent of non-Head Start centers. Also, Head Start centers, on average, have a higher percentage of teaching staff with some college or higher (54 versus 22 percent). The distribution shows that Head Start centers are more likely to have 50 percent or more of staff possessing some college or higher (63 versus 9 percent) and less likely to have less than 25 percent of staff with some college (7 percent versus 59 percent).
- Head Start centers demonstrate a higher average percentage of family partnership activities and are more likely than non-Head Start centers to offer families all the activities or strategies evaluated (82 versus 67 percent of centers).
- Head Start centers are more likely to offer a salary scale differentiated by education, training, or experience (79 versus 65 percent). No differences are evident in other quality components that fall in the general administration and management category. That is, staff management with annual performance reviews occur at similar rates in over ninety percent of both Head Start and other types of centers. And, the percentage of program management strategies employed also is not statistically different between Head Start and non-Head Start centers.
- Head Start centers and other centers are equally likely to be accredited (59 and 52 percent, respectively).

5. What is the prevalence of quality rating components in child care centers by accreditation status in Florida, Miami-Dade? [Table III.7]

In this section, we compare the quality rating level assigned and the quality components evaluated between centers that are accredited (53 percent or 135 centers) and those that are not accredited within the Miami-Dade QRIS.

- Accredited centers are more than twice as likely to achieve the highest QRIS levels than non-accredited centers (37 versus 15 percent, respectively). They are less likely to be at the middle or base QRIS levels than non-accredited centers.
- Observed quality is higher in accredited centers than non-accredited centers, with differences of 0.5 to 0.7 points across the different ERS scores. Accredited centers' total scores average between 3.5 and 4.0, with an ECERS-R Teaching and Interactions score approaching "good" (4.8 points with a developer-defined label of a 5.0 as "good").
- Child-to-staff ratios for 2-year-old and 3-year-old classrooms on average do not differ by accreditation status.

- Accredited centers use an approved curriculum at similar rates as non-accredited centers (96 percent each).
- Accredited centers are more likely to demonstrate individualization practices than non-accredited centers. Accredited centers are more likely to conduct assessments (82 versus 59 percent), but there are no differences in using results to guide planning. Accredited centers are more likely to plan for children's special needs (79 versus 52 percent) and to have resources to communicate with families in their primary language (87 versus 63 percent).
- Director and staff education levels are similar between accredited and non-accredited centers, although accredited centers are less likely than non-accredited centers to have less than 25 percent of the teaching staff with some college education or higher.
- Accredited centers demonstrate a greater percentage of family partnership activities and program management strategies than non-accredited centers. At least 90 percent of both accredited and non-accredited centers provide staff management (for example, annual performance reviews), but accredited centers are significantly more likely to do so given that almost all do (99 percent). Accredited centers are also more likely than non-accredited centers to offer a differentiated salary scale (83 versus 52 percent).

Table III.1. Quality Rating Components Among Florida, Miami- Dade Child Care Centers, by Rating Level[a]

	Percentage of Centers			Mean (SD)					
QRIS component	Base level	Middle level	Highest level	Base level		Middle level		Highest level	
Number of Centers	17	168	68						
Environment Rating Scale									
ITERS-R average total score				2.3	(0.5)	2.8	(0.7) **	4.3	(0.7) ***
ECERS-R average total score				2.6	(0.4)	3.6	(0.6) ***	4.5	(0.5)***
ECERS-R Teaching and Interactions				3.1	(0.5)	4.4	(1.0)***	5.4	(0.8) ***
ECERS-R Provisions for Learning				2.6	(0.6)	3.8	(0.7) ***	4.8	(0.7)***
ERS average total score (ECERS-R + ITERS-R)				2.4	(0.4)	3.3	(0.5)***	4.4	(0.4) ***
Curriculum/standards used	88.2	95.8	97.1						
Curriculum/standards-staff training									
All teaching staff	0.0	26.8*	33.8						
Some teaching staff	0.0	4.8	8.8						
Not all teaching staff	100.0	68.5**	57.4						
Child assessment/screening	0.0	67.3***	97.1***						
Child assessment guide planning	0.0	21.4*	25.0						
Communicate with families in home language	5.9	74.4***	97.1***						
Planning for special needs	0.0	60.7***	95.6***						
Center average 2-year-old ratio				8.3	(2.7)	7.3	(2.2)	6.5	(2.2)*
Ratio 6:1 or lower	16.7	38.5	55.8*						
Center average 3-year-old ratio				12.3	(2.9)	9.4	(2.3) ***	8.9	(2.0)
Ratio 8:1 or lower	11.8	31.2+	35.4						
Director education (some college or higher)	23.5	62.5**	74.2+						
Teacher education (percentage some college or higher)				8.5	(14.4)	30.1	(26.7)**	34.9	(20.6)
Less than 25 percent	88.2	48.2**	31.8*						
25 to 50 percent	5.9	29.2*	42.4+						
50 percent or higher	5.9	22.6	25.8						
Family partnerships									
Percentage of available				44.9	(29.3)	90.8	(16.2) ***	98.9	(5.2)***
All offered	5.9	67.3***	94.1***						

Table III.1 *(continued)*

QRIS component	Percentage of Centers			Mean (SD)					
	Base level	Middle level	Highest level	Base level		Middle level		Highest level	
Staff management	52.9	97.6***	100.0						
Salary scale differentiated	5.9	63.1***	97.1***						
Program management									
Percentage of available				26.5	(27.2)	78.6	(27.1) ***	97.1	(8.1)***
All offered	0.0	50.6***	88.2***						
Accredited center	11.8	49.4**	73.5***						

Source: Analysis of data from Florida, Miami-Dade conducted as part of the QRS Assessment project.

Note: Tests of significance were conducted between the base level to the middle level, and highest level to the middle level.

[a]Base level is the first level. Middle level includes Florida, Miami-Dade Quality Counts levels 2 and 3. Highest level includes Florida, Miami-Dade Quality Counts levels 4 and 5.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table III.2. Quality Rating Components Among Tennessee Child Care Centers, by Rating Level

	Percentage of Centers			Mean (SD)					
QRIS component	Base level	Middle level	Highest level	Base level		Middle level		Highest level	
Number of Centers	26	233	1110						
Environment Rating Scale									
ITERS-R average total score				4.0	(0.4)	4.5	(0.7) **		
ECERS-R average total score				4.4	(0.4)	4.7	(0.6) *	4.9	(0.8) ***
ECERS-R Teaching and Interactions				5.3	(0.6)	5.5	(0.7)	5.3	(0.7)***
ECERS-R Provisions for Learning				4.2	(0.7)	4.7	(0.8) **	5.8	(0.8) ***
ERS average total score (ECERS-R + ITERS-R)				4.3	(0.4)	4.6	(0.5) **	5.3	(0.9)***
Curriculum/standards used								5.2	(0.7) ***
Yes	23.1	60.1***	66.2+						
No	38.5	12.0***	0.5***						
Not evaluated	38.5	27.9	33.3						
Curriculum/standards-staff training									
All teaching staff	15.4	49.4***	64.7***						
Some teaching staff	7.7	10.7	1.5***						
Not all teaching staff	38.5	12.0***	0.5***						
Not evaluated	38.5	27.9	33.3						
Center average 2-year-old ratio				6.2	(1.0)	6.0	(0.9)		
Ratio 6:1 or lower	42.9	58.5	91.1***					5.4	(0.7) ***
Center average 3-year-old ratio				8.7	(0.5)	8.6	(0.5)		
Ratio 8:1 or lower	26.3	40.0	76.1***					8.2	(0.4)***
Director education	34.6	72.1***	98.1***						
Teacher education (percentage some college or higher)				0.8	(2.7)	22.8	(19.5)***		
Less than 25 percent	100.0	41.6***	0.9***					46.9	(8.7) ***
25 to 50 percent	0.0	30.5***	10.7***						
50 percent or higher	0.0	27.9**	88.4***						
Family partnerships									
Percentage of available				46.2	(29.7)	84.5	(23.6) ***		
All offered	7.7	58.8***	95.2***					98.7	(6.5)***
Staff management	7.7	70.8***	99.6***						

Table III.2 *(continued)*

QRIS component	Percentage of Centers			Mean (SD)					
	Base level	Middle level	Highest level	Base level		Middle level		Highest level	
Salary scale differentiated	38.5	91.0***	99.3***						
Staff benefits									
Percentage of available			.	76.9	(27.3)	96.0	(14.8) ***		
All offered	38.5	91.0***	99.3***					99.8	(3.6) ***
Accredited center	0.0	0.4	9.1***						

Source: Analysis of data from Tennessee conducted as part of the QRS Assessment project.

Note: Tests of significance were conducted between the base level to the middle level, and highest level to the middle level.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table III.3. Quality Rating Components Among Base QRIS Rating Level- Florida, Miami- Dade And Tennessee

	Percentage of Centers		Mean (SD)			
QRIS component	Florida, Miami-Dade	Tennessee	Florida, Miami-Dade		Tennessee	
Number of Centers	17	26				
Environment Rating Scale						
ITERS-R average total score			2.3	(0.5)	4.0	(0.4) ***
ECERS-R average total score			2.6	(0.4)	4.4	(0.4)***
ECERS-R Teaching and Interactions			3.1	(0.5)	5.3	(0.6) ***
ECERS-R Provisions for Learning			2.6	(0.6)	4.2	(0.7)***
ERS average total score (ECERS-R + ITERS-R)			2.4	(0.4)	4.3	(0.4)***
Curriculum/standards used	88.2	37.5**				
Curriculum/standards-staff training						
All teaching staff	0.0	25.0*				
Some teaching staff	0.0	12.5				
Not all teaching staff	100.0	62.5**				
Center average 2-year-old ratio			8.3	(2.7)	6.2	(1.0)*
Ratio 6:1 or lower	16.7	42.9				
Center average 3-year-old ratio			12.3	(2.9)	8.7	(0.5)***
Ratio 8:1 or lower	11.8	26.3				
Director education (some college or higher)	23.5	34.6				
Teacher education (percentage some college or higher)			8.5	(14.4)	0.8	(2.7)*
Less than 25 percent	88.2	100.0+				
25 to 50 percent	5.9	0.0				
50 percent or higher	5.9	0.0				
Family partnerships						
Percentage of available			44.9	(29.3)	46.2	(29.7)
All offered	5.9	7.7				
Staff management	52.9	7.7***				
Salary scale differentiated	5.9	38.5*				
Accredited center	11.8	0.0+				

Source: Analysis of data from Florida, Miami-Dade and Tennessee conducted as part of the QRS Assessment project.

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a] For the comparison of curriculum components, we conducted t-tests among centers who are evaluated on it (that is, "Not evaluated" is treated as missing). Thirty-nine percent of Tennessee providers were not evaluated on this component, introduced in January 2010.

Table III.4. Quality Rating Components Among Middle QRIS Rating Level(s)- Florida, Miami- Dade And Tennessee

	Percentage of Centers		Mean (SD)			
QRIS component	Florida, Miami-Dade	Tennessee	Florida, Miami-Dade		Tennessee	
Number of Centers	168	233				
Environment Rating Scale						
ITERS-R average total score			2.82	(0.70)	4.48	(0.69)***
ECERS-R average total score			3.56	(0.58)	4.69	(0.64)***
ECERS-R Teaching and Interactions			4.35	(0.99)	5.45	(0.73)***
ECERS-R Provisions for Learning			3.75	(0.70)	4.7	(0.82)***
ERS average total score (ECERS-R + ITERS-R)			3.27	(0.51)	4.6	(0.52)***
Curriculum/standards used[a]	95.8	83.3***				
Curriculum/standards-staff training[a]						
All teaching staff	26.8	68.5***				
Some teaching staff	4.8	14.9**				
Not all teaching staff	68.5	16.7***				
Center average 2-year-old ratio			7.29	(2.18)	6.01	(0.91)***
Ratio 6:1 or lower	38.5	58.5**				
Center average 3-year-old ratio			9.38	(2.26)	8.6	(0.49)***
Ratio 8:1 or lower	31.2	40.0+				
Director education (some college or higher)	62.5	72.1*				
Teacher education (percentage some college or higher)			30.1	(26.7)	22.8	(19.5)**
Less than 25 percent	48.2	41.6				
25 to 50 percent	29.2	30.5				
50 percent or higher	22.6	27.9				
Family partnerships						
Percentage of available			90.8	(16.2)	84.5	(23.6)**
All offered	67.3	58.8+				
Staff management	97.6	70.8***				
Salary scale differentiated	63.1	91.0***				
Accredited center	49.4	0.4***				

Source: Analysis of data from Florida, Miami-Dade and Tennessee conducted as part of the QRS Assessment project.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[a] For the comparison of curriculum components, we conducted t-tests among centers who are evaluated on it (that is, "Not evaluated" is treated as missing). Twenty-eight percent of Tennessee providers were not evaluated on this component, introduced in January 2010.

Table III.5. Quality Rating Components Among Highest QRIS Rating Level(s)- Florida, Miami- Dade And Tennessee

	Percentage of Centers		Mean (SD)			
QRIS component	Florida, Miami-Dade	Tennessee	Florida, Miami-Dade		Tennessee	
Number of Centers	68	1110				
Environment Rating Scale						
ITERS-R average total score			4.3	(0.7)	4.9	(0.8) ***
ECERS-R average total score			4.5	(0.5)	5.3	(0.7) ***
ECERS-R Teaching and Interactions			5.4	(0.8)	5.8	(0.8) ***
ECERS-R Provisions for Learning			4.8	(0.7)	5.3	(0.9) ***
ERS average total score (ECERS-R + ITERS-R)			4.4	(0.4)	5.2	(0.7) ***
Curriculum/standards used[a]	97.1	99.3+				
Curriculum/standards-staff training						
All teaching staff	33.8	97.0***				
Some teaching staff	8.8	2.3**				
Not all teaching staff	57.4	0.7***				
Center average 2-year-old ratio			6.5	(2.2)	5.4	(0.7) ***
Ratio 6:1 or lower	55.8	91.1***				
Center average 3-year-old ratio			8.9	(2.0)	8.2	(0.4) ***
Ratio 8:1 or lower	35.4	76.1***				
Director education (some college or higher)	74.2	98.1***				
Teacher education (percentage some college or higher)			34.9	(20.6)	46.9	(8.7) ***
Less than 25 percent	31.8	0.9***				
25 to 50 percent	42.4	10.7***				
50 percent or higher	25.8	88.4***				
Family partnerships						
Percentage of available			98.9	(5.2)	98.7	(6.5)
All offered	94.1	95.2				
Staff management	100.0	99.6				
Salary scale differentiated	97.1	99.3				
Accredited center	73.5	9.1***				

Source: Analysis of data from Florida, Miami-Dade and Tennessee conducted as part of the QRS Assessment project.

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a] For the comparison of curriculum components, we conducted t-tests among centers who are evaluated on it (that is, "Not evaluated" is treated as missing). Thirty-three percent of Tennessee providers were not evaluated on this component, introduced in January 2010.

Table III.6. Quality Rating Levels and Components in Florida, Miami- Dade, by Head Start Status

	Percentage of centers		Mean (SD)			
QRIS component	Head Start Centers	Non-Head Start Centers	Head Start Centers		Non-Head Start Centers	
Number of Centers	61	192				
QRIS overall rating level						
Highest levels (level 4 and 5)	36.1	24.0+				
Middle levels (level 2 and 3)	63.9	67.2				
Base (lowest) level	0.0	8.9*				
Percentage of children receiving subsidies (0-5 years)[a]						
No subsidized children	96.4	12.5[a]				
0.1 to 25 percent	1.8	28.8				
25 to 50 percent	0.0	29.9				
50 percent or higher	1.8	28.8				
Program size (0-5 years)						
Small (50 or fewer)	41.1	34.8				
Medium (51 through 100)	26.8	47.8**				
Large (101 through 150)	8.9	13.0				
Very large (151 or greater)	23.2	4.4**				
Poverty level in community						
0 percent	3.3	9.9				
1 to 49 percent	47.5	51.0				
50 percent or greater	49.2	39.1				
Environment Rating Scale						
ITERS-R average total score			3.7	(0.9)	3.1	(1.0)*
ECERS-R average total score			3.7	(0.6)	3.8	(0.8)
ECERS-R Teaching and Interactions			4.5	(1.0)	4.6	(1.1)
ECERS-R Provisions for Learning			3.8	(0.7)	4.0	(1.0)
ERS average total score (ECERS-R + ITERS-R)			3.7	(0.6)	3.5	(0.8)*
Curriculum/standards used	96.7	95.3				
Curriculum/standards-staff training						
All teaching staff	62.3	15.6***				
Lead teachers	6.6	5.2				
Not all staff	31.2	79.2***				
Child assessment screening	93.4	63.5***				
Child assessment guides planning	63.9	7.3***				
Planning for special needs	91.8	57.8***				
Communicate with families in home language	95.1	69.8***				
Center average 2-year-old ratio[b]			4.9	(1.3)	7.3	(2.2)
Ratio 6:1 or lower	90.0	39.2**				
Center average 3-year-old ratio			9.7	(1.2)	9.4	(2.7)
Ratio 8:1 or lower	8.3	38.6***				
Director education (some college or higher)	93.1	53.5***				
Teacher education (percentage of staff with some college or higher)			54.3	(23.7)	22.3	(20.5)***
Less than 25 percent	6.7	59.2***				
25 to 50 percent	30.0	31.4				
50 percent or higher	63.3	9.4***				

Table III.6 *(continued)*

QRIS component	Percentage of centers		Mean (SD)			
	Head Start Centers	Non-Head Start Centers	Head Start Centers		Non-Head Start Centers	
Family partnerships						
Percentage of available			95.1	(15.0)	88.2	(21.0)*
All offered	82.0	66.7*				
Staff management	98.4	94.3				
Salary scale differentiated	78.7	65.1*				
Program management						
Percentage of available			84.4	(23.4)	78.7	(30.2)
All offered	59.0	56.8				
Accredited center	59.0	51.6				

Source: Analysis of data from Florida, Miami-Dade conducted as part of the QRS Assessment project.

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a]The percentage of subsidized children was calculated by the number of children receiving child care subsidies divided by the total enrollment. Given Head Start itself is a subsidized program, it would appear that children enrolled are not being counted. However, for the non-Head Start centers this information provides us information on the families served.

[b] Two-year-old ratio data was available for 10 Head Start centers and 158 non-Head Start centers.

Table III.7. Quality Rating Components Among Florida, Miami Dade Child Care Centers, By Accreditation Status

	Percentage of centers		Mean (SD)			
QRIS component	Accredited Centers	Non-accredited Centers	Accredited Centers		Non-accredited Centers	
Number of Centers	135	118				
QRIS overall rating level						
Highest levels (level 4 and 5)	37.0	15.3***				
Middle levels (level 2 and 3)	61.5	72.0+				
Base (lowest) level	1.5	12.7***				
Program size (0-5 years)						
Small (50 or fewer)	25.4	48.3***				
Medium (51 through 100)	45.2	40.4				
Large (101 through 150)	15.9	7.9+				
Very large (151 or greater)	13.5	3.5**				
Poverty level in community						
0 percent	8.2	8.5				
1 to 49 percent	46.7	54.2				
50 percent or greater	45.2	37.3				
Environment Rating Scale						
ITERS-R average total score			3.5	(1.0)	2.8	(0.9)***
ECERS-R average total score			4.0	(0.6)	3.5	(0.8)***
ECERS-R Teaching and Interactions			4.8	(1.0)	4.3	(1.2)***
ECERS-R Provisions for Learning			4.2	(0.8)	3.7	(1.0)***
ERS average total score (ECERS-R + ITERS-R)			3.8	(0.7)	3.2	(0.7)***
Center average 2-year-old ratio			7.0	(2.0)	7.3	(2.5)
Ratio 6:1 or lower	43.5	40.8				
Center average 3-year-old ratio			9.5	(2.2)	9.4	(2.6)
Ratio 8:1 or lower	28.7	33.6				
Curriculum/standards used	95.6	95.8				
Curriculum/standards-staff training						
All	31.1	22.0				
Some	7.4	3.4				
None	61.5	74.6*				
Child assessment/screening	81.5	58.5***				
Child assessment guide planning	23.0	18.6				
Communicate with families in home language	87.4	62.7***				
Planning for special needs	78.5	51.7***				
Director education (some college or higher)	65.4	60.2				
Teacher education (percentage of staff with some college or higher)			32.7	(23.7)	26.8	(26.7)
Less than 25 percent	40.6	53.4*				
25 to 50 percent	34.6	27.1				
50 percent or higher	24.8	19.5				
Family partnerships						
Percentage of available			95.2	(13.2)	83.8	(24.1)***
All offered	82.2	56.8***				
Staff management	98.5	91.5**				
Salary scale differentiated	83.0	51.7***				
Program management						
Percentage of available			89.1	(20.6)	69.7	(33.0)***
All offered	70.4	42.4***				

Source: Analysis of data from Florida, Miami-Dade conducted as part of the QRS Assessment project.

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

IV. UNIQUE CONTRIBUTIONS OF QUALITY COMPONENTS IN PREDICTING OBSERVED QUALITY

A. Analytic Goals

As QRIS continue to develop, there remain many unanswered questions about which components to include in a rating system and in what combination. To date, QRIS developers have made these decisions based on a limited body of research to identify important components of quality, combined with state or local goals and contexts, to specify the composition of components and the number of rating levels (Caronongan et al. 2011). Additional research is still needed to help states and localities in shaping QRIS to create a meaningful framework for identifying quality and differentiating between levels of quality, particularly in ways that connect with positive outcomes for children.

The goal of this analysis is to further build the knowledge base about what contributes to quality by examining the association of each quality component with observed quality. A few QRIS evaluations to date have employed multivariate approaches to examine the association of particular quality components with ERS scores as an available proxy for quality (Norris et al. 2003; Zellman et al. 2008). Their focus tended to be on staff qualifications, with additional components (such as ratios and family partnerships) noted as controls. While staff qualifications did not make a significant unique contribution in predicting observed quality in these evaluations, the exact component metrics included in the models varied, limiting the ability to draw conclusions across QRIS models. Norris and colleagues, however, did find that other variables, including master teacher ratio, family partnerships, and presence of interest centers (a learning environment component), were significant predictors of observed quality (as measured by the ERS).

Defining and measuring quality is a complex process and has many variations across different QRIS—there is unlikely to be just one correct approach. Nonetheless, looking across different QRIS to conduct analysis using a common metric can answer questions about shared concepts on which systems are built.

Through multivariate analyses, we examine the unique contribution of each quality rating component in predicting observed environment quality, as measured by the ERS, to answer the following questions:

- What is the association between each quality rating component-with observed quality when all other components are held constant?
- What patterns in these associations are present across different ERS scales and different QRIS?

B. Analytic Strategy and Considerations

Certain features of our analytic strategy have implications for understanding and interpreting findings. These include:

1. Specifying a proxy for quality (the outcome variable)

- These data were not collected for the purpose of answering specific research questions relevant to QRIS quality measurement, design, and refinement. Existing administrative

data can be a rich resource for examining questions about the design and influence of systems, but the analysis must rely on measures and outcomes that are available in the data. For example, observed quality as measured by the ERS is used as an outcome, but it is also a component that contributes toward the overall rating level in the QRIS we studied (and numerous others). The ERS score is the only component assessed using an independent, standardized measure, separate from provider self-reports and documentation. As a widely used measure of global quality, it provides a common reference point across QRIS as well as across various research studies. For the purpose of this analysis, we pull out the ERS score as an independent measure of quality and treat it as distinct from the other components even though it was collected as part of the rating process.

- We specify ERS scores in several ways: (1) a combined ERS score that is the average across ECERS-R and ITERS-R scores for a particular center and most closely represents a facility-level score; (2) separate ECERS-R and ITERS-R average scores that represent scores across classrooms serving children of particular ages (children ages 2 through 5, and children from birth to 2-1/2 years, respectively); and (3) scores on the two ECERS-R factors—Teaching and Interactions and Provisions for Learning. The first specification mirrors how Tennessee calculates facility-level ERS scores for use in quality ratings. The second mirrors the approach of Florida, Miami-Dade to calculating facility-level ERS scores. The third was used as a means of focusing on child-related items within the broader construct of global environment quality. These dimensions may have a closer association with changes in child outcomes.

2. Specifying the sample for analysis

- These data are specific to center-based providers in two QRIS—Miami-Dade and Tennessee. The particular characteristics of these two sites as well as the specific design of each system, precludes generalizing the results to other systems or geographical areas.
- These analyses included centers from Florida, Miami-Dade and Tennessee that had ERS data and child-to-staff ratio information. Sample sizes vary by outcome, especially between ITERS-R and ECERS-R given the ages of the children served by a center. Each table notes the analytic sample size and the methodological notes for Chapter IV included in Appendix D provide further details about how each sample was derived.

3. Specifying the quality components (the explanatory variables)

- The analysis again builds on the common metric for defining quality components across QRIS. Therefore, the analysis is not an assessment of any particular system; the variables representing quality components are not specified in exactly the same way as Florida, Miami-Dade or Tennessee define them within their own QRIS. The variables were derived to best capture the concept or intent of the component across QRIS while making the most use of the data at the greatest level of specificity possible in a cross-site analysis. For example, staff qualifications is a component of quality that is included in at least 26 QRIS (Tout et al. 2010). However, the exact specification of this component in terms of the break down by specific staff roles (director, lead teacher, assistant teacher), the level of education, and specific training in early care and education varies widely (Caronongan et al. 2011). Based on the data available from the QRIS included in this analysis, we specify a teacher education variable that distinguishes centers by the

percentage of staff that have at least some college education. The percentage of staff is a common approach employed across QRIS in setting and assessing rating level requirements (Caronongan et al. 2011).

- This cross-QRIS approach allows for the inherent differences across systems yet begins to address key questions that are common about which components are most closely linked to observed quality. By measuring the concepts that are often common across sites (and can likely be captured in similar ways as structured here), the findings can have broader applicability than finely tuned analysis that is specific to a particular QRIS (although such an analysis is valuable for that QRIS). An added advantage is that the common metric could be similarly employed by other QRIS evaluation teams to conduct a parallel analysis to further amass information and inform policy and practice. The limitation in the use of common metrics is that the variables tend to be more summative in nature than the finer-grained measures that may be examined for any one particular QRIS. This analysis may mask effects that might otherwise be detected through use of more robust measures.
- In the cross-site model, we introduce an "individualization practices" variable, which denotes the presence of four quality components: (1) conduct of child assessment screening, (2) use of information from child assessments to guide planning, (3) planning for children with special needs, and (4) communicating with families in their home language.
- The information we have on individualization practices comes only from Florida, Miami-Dade; Tennessee does not include requirements for these four components in its QRIS. In order to fit a cross-site model that accounts for site differences, we created three groups of centers: (1) Miami-Dade centers that demonstrated two or fewer individualization practices; (2) Miami-Dade centers that demonstrated three or more individualization practices; and (3) Tennessee centers that were not evaluated on these components. Each group is represented with dichotomous indicators in the cross-site model, with the first group as the referent category. These groups allows us to accomplish two analytic goals: (1) control for site differences (a site effect) since all the Tennessee centers fall into one group, and (2) assess the effect of the presence of numerous individualization practices by breaking up the Florida, Miami-Dade centers into two groups. Rather than drop these new and developing components of QRIS, this approach enabled an analysis, albeit limited, of their association with observed quality.
- The regression coefficients for each component represent the estimated average difference in ERS scores between centers demonstrating that component and centers that do not. The ERS scores are from a single time point and do not provide information on change or increasing quality.
- Categorical variables have been defined for use in the regression models such that one category is omitted as the referent, or comparison group. The referent category represents the low end or absence of the quality component for conceptual ease in describing differences. For example, curriculum staff training has categories for all teaching staff trained, some teaching staff trained, not all teaching staff trained, and not evaluated. The referent category is not all teaching staff, and the comparisons for each of the other three categories are relative to not all teaching staff are trained.

4. Specifying the models

- The regression models employed are straightforward ordinary least squares (OLS) regressions that provide information about the unique contribution of each quality component (as explanatory variables) in predicting the ERS score (as the outcome variable) when all other components are held constant. As the exploratory analyses focused on the quality components, the intent was not to develop a best fit model for predicting the ERS outcome, but to include all the quality components that could be commonly defined across the QRIS. Control variables based on other characteristics of the center-based provider (such as size or percentage of subsidized children served) or the locality (such as neighborhood income level) were not included. The tables for this chapter present results from the regression models as follows:
    - **Table IV.1** presents results from fitting a cross-site model using data from two QRIS. Common quality components are included in the model as predictors of the average combined ERS score. To examine whether associations vary by site, we also developed site-specific models. In Florida, Miami-Dade, two models are shown: one that mirrors the model for Tennessee and one that includes the individualization practices variable, which cannot be included in the Tennessee model.
    - **Tables IV.2 through IV.5** present results from models that use this same approach but use the ITERS-R score, ECERS-R score, ECERS-R Teaching and Interactions factor score, and ECERS-R Provisions for Learning factor score, as outcomes respectively.
    - **Table IV.6** presents site-specific models that include quality components that cannot be included in a cross-site model.

C. Key Findings

- Across sites and different specifications of the ERS outcome, the use of a differentiated salary scale and accreditation status were consistently associated with higher scores. Centers with differentiated salary scales were found to have higher ERS scores compared to centers that do not offer a similar benefit. Accredited centers were consistently found to have higher ERS scores than those not accredited.
    - These associations do not indicate that certain quality components cause higher ERS scores or vice versa. Rather, it suggests that in our sample, these are the components most strongly correlated with ERS scores. For example, if we were to compare an accredited provider with a non-accredited provider with identical values on all other observed cross-site components, on average the accredited provider will tend to have higher ERS scores.
    - The unique contribution of these two components may reflect characteristics not observed or not directly measured. For example, accredited centers may have the motivation to achieve a high rating level as well as time and resources that can also influence the presence of other quality initiatives overall. Also, accreditation may reflect other quality components that could be brought to a center's focus that are not evaluated. Similarly, the contribution of salary scale may encompass other center or staff characteristics not included, such as retention rates, which could contribute to ERS scores.

- Both Florida, Miami-Dade and Tennessee assign points for individual components to derive rating levels from total points across components. As such, providers can achieve a higher level in a number of ways. The consistent association of the use of differentiated salary scale and accreditation status signifies that these are components that typically accompany high ERS scores and likely higher QRIS ratings in the two sites. Given that both QRIS are voluntary (and the majority of the centers in the sites are rated at the highest level), it is possible that there is some self-selection that occurs and that centers that have each of these components are motivated to participate knowing that they can achieve a high rating.
- Three other components—teacher education, child-to-staff ratios, and family partnerships—demonstrated unique contributions across models, but not consistently and not at the level of significance seen for the salary scale and accreditation components.
    - QRIS, beyond those included here, consider staff qualifications as important elements in their standards. This is reflected in resources often provided through professional development systems to assist providers in meeting requirements at higher levels, as well as in the time and resources they are willing to expend to assess this component given its complexities. This analysis, while not confirmatory, suggests that a measure of the level of teacher education can make a unique contribution to predicting observed quality.
    - Even as a newer and less refined component adopted by QRIS (i.e. a simple count of various activities), the presence of family partnership strategies demonstrates a unique contribution in predicting observed quality. Although often just marginally significant, this suggests that family partnerships could be a useful emerging area to further define and measure.
- We found some unique patterns of associations for specific ECERS-R factor scores:
    - The Teaching and Interactions factor predominantly includes items on greeting/departing, language-reasoning, and interactions. Centers with lower ratios and that employ more individualization practices (components that demonstrated unique contributions on this factor score) may be better able to facilitate activities tapped by this subscale.
    - The Provisions for Learning factor consists of items including space and furnishings, activities offered, and program structure subscales. Family partnerships (which demonstrated only marginal significance) often include the provision of family activities and modes of communication like bulletin boards—practices that may reflect organization to design experiences and arrangement of space.
- The differences in ERS scores associated with individual quality components were all less than half a point in magnitude. Half a point is equivalent to the difference in ERS scores between a level one and level two provider in Florida, Miami-Dade and Tennessee. Within the context of QRIS as currently designed, the largest estimated difference in ERS scores associated with any of the components is smaller than what the systems would require to move up one quality rating level.

D. Detailed Findings.[19]

1. What is the unique contribution of each quality rating component in predicting the combined ERS score? [Table IV.1]

- Four quality components demonstrated a statistically significant and unique contribution in predicting combined ERS scores (average ITERS-R + ECERS-R) in the cross-site model. Lower child-to-staff ratios, having 50 percent or more of teaching staff with at least some college, utilizing a differentiated salary scale, and being accredited are each associated with having a higher combined ERS score, on average, controlling for other components. These associations signify that, for example, centers in which 50 percent or more of teaching staff have some college or more have a combined ERS score that is 0.25 points higher, on average, than centers in which less than 25 percent of the teaching staff have that level of education but identical values on all other components.
- The magnitude of differences associated with the four components range from 0.12 points on the combined ERS score (for child-staff ratio) to 0.41 points (for differentiated salary scale).
- Estimates for the individualization practices variable reflect both site effects and a unique contribution for such practices:
  - The coefficient for the "not evaluated" category indicates that Tennessee centers have a combined score that is 1.49 points higher, on average, than Florida, Miami-Dade centers that implement few individualization strategies—this is the site effect.
  - We find a marginally significant difference between Florida, Miami-Dade centers by the level of individualization practices demonstrated. Centers demonstrating more than half of these practices have slightly higher combined ERS scores than those that have half or less of the individualization components (0.18, $p < .10$). This difference is a significant association in the Florida, Miami-Dade model.
- Of the four components with a unique contribution toward explaining differences in observed quality in the cross-site model, both salary scale and accreditation status demonstrated unique contributions on combined scores in each of the site-specific models.
  - The magnitude of differences (for higher observed quality) between accredited and non-accredited centers is similar in the two sites (0.32 to 0.33 points in Florida, Miami-Dade, 0.36 points in Tennessee).
  - We also find similar differences in combined ERS scores between centers that use a differentiated salary scale and those that do not across the two sites (0.39 to 0.44 in Miami, 0.36 in Tennessee).
- The association of having 50 percent or more of teaching staff with some college or more to observed quality is statistically significant in the Tennessee model but only marginally significant in the Florida, Miami-Dade model.

[19] Refer to Appendix F for additional detailed data tables.

- The association of child-to-staff ratios to observed quality is statistically significant and family partnerships strategies is marginally significant in the Tennessee model but not in the Florida, Miami-Dade models.

2. What patterns in the associations between quality rating components and observed quality are present for ITERS-R and ECERS-R scores separately? [Tables IV.2 and IV.3]

- Overall, the pattern of unique contributions of quality components in predicting the ECERS-R score more closely mirrors those found in the combined ERS score model than those in the ITERS-R model. This could be expected given the relative sample sizes that make for a larger influence of ECERS-R in the combined score.
- There is a slight decrease in the number of quality components that demonstrated a significant association with observed quality in the separate ITERS-R and ECERS-R models in comparison with the combined ERS score model. The combined score may be a closer reflection of the facility level quality outcome and, therefore, closer in relevancy to the quality components that are also measured at the facility level (rather than classroom-level variables).
- Salary scale and accreditation demonstrate unique contributions in predicting higher observed quality across each ERS outcome.
- The child-to-staff ratio continues to demonstrate a unique contribution in predicting the ECERS-R score, as it did on the combined ERS score. However, it has no association with the ITERS-R score. The opposite case is reflected in teacher education: similar to the model predicting combined ERS scores, we find higher ITERS-R scores among centers with a more educated teaching staff (50 percent or more with some college education or more), but teacher education does not demonstrate a significant association with the ECERS-R score.
- The family partnerships component has a marginally statistically significant association in the combined ERS model but is significant in the ECERS-R model.
- Site effects are significant across the three models (ERS combined, ECERS-R, and ITERS-R), as reflected in the "not evaluated" category of the individualization practices variable. The size of the coefficient reflects that Tennessee centers have scores that are 1.3 to 1.8 points higher, on average, than centers in Florida, Miami-Dade that do not have multiple individualization practices in place.
- There is a significant association of the "more than half" individualization practices variable with the ITERS-R score, but not the ECERS-R score. This indicates that centers in Florida, Miami-Dade that have numerous individualization practices in place have ITERS-R scores that are 0.45 points higher, on average, than those that have few individualization practices in place.

3. What patterns in the associations between quality rating components and observed quality are present for ECERS-R factor scores: Teaching and Interactions and Provisions for Learning? [Tables IV.4 and IV.5]

- While the factor scores may potentially capture dimensions of quality that have greater relevance in predicting child outcomes than global quality, these scores seem to get

further away from a connection to the facility-level quality component variables in that we see a further decrease in the number of components that demonstrated unique contributions in predicting scores.

- As with the previous models, the use of a differentiated salary scale and accreditation status demonstrated unique contributions in predicting each of the factor scores.
- Child-to-staff ratios demonstrated a significant association with the Teaching and Interactions factor score.
- A curriculum variable demonstrated a unique, but counter-intuitive, association with the Provisions for Learning 399999score. There is a small, but significant, association that suggests that centers in which some teaching staff are trained on a standards-based curriculum have a score that is 0.32 points lower, on average, than centers in which none of the staff are trained.
- Site effects are again significant in each of the models. The level of individualization practices in Florida, Miami-Dade centers demonstrated a significant association with the Teaching and Interactions score, but not the Provisions for Learning score.

4. What patterns in the associations between quality rating components and observed quality are present across the different QRIS? [Table IV.6]

The site-specific models examined associations of site-specific quality components that could not be captured in a common approach. The intent was to explore how the associations might change when each site model includes an increased number and/or specificity to the components. The Florida, Miami-Dade model includes each of the individualization practices as separate predictors and adds program management as a quality component evaluated only in this site. The Tennessee model includes a quality component not evaluated in Florida, Miami-Dade—the number of staff benefits offered. We examine the combined ERS score as the primary outcome of interest.[20]

- Two of the four quality components that demonstrated a unique contribution in predicting the combined ERS score in the cross-site model maintained these significant associations in each of the site-specific models. Accreditation status has a statistically significant association with higher combined ERS scores in both site models. Higher educated teaching staff (50 percent or more with at least some college education) has a marginally significant association with the combined ERS score in the Florida, Miami-Dade model, and a significant association in the Tennessee model.
- The two other components that demonstrated a unique contribution in predicting the combined ERS score in the cross-site model—child-to-staff ratio and the use of a differentiated salary scale—maintain a unique contribution in the Tennessee model.
- The addition of staff benefits to the Tennessee model actually changes very little from the model shown in Table IV.1. The staff benefits component does not have a statistically significant association and the parameter estimates for other variables are

[20] Refer to Appendix F for tables that present the model steps for each site on the combined ERS score, as well as the site-specific model steps for the other ERS outcomes examined.

largely unchanged when it is added to the model. This is likely due to collinearity between staff benefits and the salary scale variable.

- In the Florida, Miami-Dade model, the unique contribution of the use of a differentiated salary scale in predicting the combined ERS score disappears with the addition of the program management variable. The program management component demonstrated a trend toward a positive association with the combined score (only marginally significant). None of the individualization practice variables (child assessment screening, child assessment guides planning, planning for children with special needs, and communicate with families in their home language) demonstrated a significant association with the combined ERS score when they are included separately.

Table IV.1. Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ECERS-R+ITERS-R Total Scores, by Cross-Site and Site-Specific Models

	Cross-site model		Florida, Miami-Dade				Tennessee	
			Cross-site common components		Individualization components			
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	2.63	0.11***	2.54	0.20***	2.52	0.20***	4.04	0.15***
Curriculum-staff training (vs. not all staff)								
All teaching staff	0.05	0.08	-0.05	0.11	-0.08	0.11	0.17	0.12
Some teaching staff	-0.14	0.11	0.01	0.19	0.02	0.19	-0.10	0.15
Not evaluated	0.13	0.08					0.23	0.12+
Child-to-staff ratio (3-year-old - 8:1 or lower)	0.12	0.04**	0.08	0.10	0.08	0.10	0.12	0.04**
Teacher education (vs. less than 25 percent)								
50 percent or greater with some college or higher	0.25	0.08**	0.21	0.12+	0.20	0.12+	0.35	0.12**
25 to 50 percent with some college or higher	0.12	0.08	0.09	0.10	0.10	0.10	0.23	0.13+
Director education (some college or higher vs. less)	0.03	0.06	0.05	0.10	0.03	0.10	0.04	0.08
Family partnerships (all vs. less than all)	0.12	0.06+	0.19	0.12	-0.01	0.16	0.12	0.07+
Staff management	0.06	0.10	0.25	0.22	0.25	0.22	-0.10	0.14
Salary scale	0.41	0.09***	0.44	0.12***	0.39	0.13**	0.43	0.12***
Accredited center	0.34	0.06***	0.33	0.09***	0.32	0.09***	0.36	0.07***
Individualization; assessment, planning for special needs, family communication (vs. two or less)								
More than half	0.18	0.11+			0.31	0.15*		
Not evaluated	1.49	0.12***						
R-square	0.51		0.30		0.31		0.12	
Adjusted R-square	0.50		0.27		0.28		0.11	
Number of centers	1327		227		227		1100	

Source: Analysis of data from Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table IV.2. Results of OLS Regressions of Quality Rating Components Among Center- based Providers with Average ITERS- R Total Scores, by Cross- Site and Site- Specific Models

	Cross-site model		Florida, Miami-Dade				Tennessee	
			Cross-site common components		Individualization components			
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	2.07	0.17***	1.95	0.35***	1.96	0.35***	3.94	0.23***
Curriculum-staff training (vs. not all staff)								
All teaching staff	0.10	0.13	0.11	0.19	0.10	0.19	0.07	0.19
Some teaching staff	0.03	0.18	0.19	0.28	0.19	0.28	-0.14	0.27
Not evaluated	0.05	0.14					0.02	0.20
Child-to-staff ratio (2-year-old - 6:1 or lower)	0.05	0.08	0.09	0.15	0.08	0.15	0.06	0.09
Teacher education (vs. less than 25 percent)								
50 percent or greater with some college or higher	0.26	0.13*	0.32	0.27	0.33	0.27	0.05	0.23
25 to 50 percent with some college or higher	0.21	0.12+	0.29	0.15+	0.30	0.15+	-0.03	0.24
Director education (some college or higher vs. less)	0.02	0.10	-0.02	0.15	-0.02	0.15	0.08	0.14
Family partnerships (all vs. less than all)	0.04	0.10	0.59	0.20**	0.50	0.28+	-0.03	0.11
Staff management	0.10	0.17	0.11	0.38	0.11	0.39	0.30	0.26
Salary scale	0.50	0.13***	0.46	0.20*	0.44	0.21*	0.50	0.19**
Accredited center	0.26	0.10**	0.25	0.15	0.24	0.15	0.24	0.15
Individualization; assessment, planning for special needs, family communication (vs. two or less)								
More than half	0.45	0.16**			0.13	0.26		
Not evaluated	1.82	0.18***						
R-square	0.52		0.31		0.31		0.07	
Adjusted R-square	0.51		0.26		0.26		0.04	
Number of centers	632		152		152		480	

Source: Analysis of data from Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table IV.3. Results of OLS Regressions of Quality Rating Components Among Center- based Providers with Average ECERS- R Total Scores, by Cross- Site and Site- Specific Models

	Cross-site model		Florida, Miami-Dade				Tennessee	
			Cross-site common components		Individualization components			
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	2.95	0.12***	2.80	0.22***	2.79	0.21***	4.15	0.17***
Curriculum-staff training (vs. not all staff)								
All teaching staff	0.02	0.09	-0.11	0.12	-0.13	0.12	0.19	0.13
Some teaching staff	-0.19	0.12	0.00	0.20	0.01	0.20	-0.10	0.17
Not evaluated	0.09	0.09					0.25	0.13+
Child-to-staff ratio (3-year-old - 8:1 or lower)	0.10	0.04*	0.11	0.10	0.12	0.10	0.08	0.05+
Teacher education (vs. less than 25 percent)								
50 percent or greater with some college or higher	0.08	0.08	-0.02	0.13	-0.02	0.13	0.22	0.14
25 to 50 percent with some college or higher	-0.01	0.09	0.01	0.11	0.01	0.11	0.11	0.14
Director education (some college or higher vs. less)	0.07	0.07	0.06	0.10	0.04	0.10	0.11	0.09
Family partnerships (all vs. less than all)	0.14	0.07*	0.10	0.13	-0.07	0.17	0.15	0.07*
Staff management	0.12	0.11	0.39	0.23+	0.39	0.23+	-0.08	0.15
Salary scale	0.41	0.09***	0.46	0.13***	0.42	0.13**	0.39	0.13**
Accredited center	0.29	0.06***	0.27	0.10**	0.26	0.10**	0.31	0.08***
Individualization; assessment, planning for special needs, family communication (vs. two or less)								
More than half	0.07	0.12			0.26	0.16		
Not evaluated	1.30	0.13***						
R-square	0.40		0.23		0.24		0.08	
Adjusted R-square	0.39		0.19		0.20		0.07	
Number of centers	1327		227		227		1100	

Source: Analysis of data from Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table IV.4. Results of OLS Regressions of Quality Rating Components Among Center- based Providers with Average ECERS- R Teaching and Interactions Scores, by Cross- Site and Site- Specific Models

	Cross-site model		Florida, Miami-Dade				Tennessee	
			Cross-site common components		Individualization components			
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	3.86	0.14***	3.38	0.33***	3.35	0.33***	5.16	0.18
Curriculum-staff training (vs. not all staff)								
All teaching staff	-0.07	0.10	-0.17	0.18	-0.20	0.18	0.06	0.14
Some teaching staff	-0.17	0.14	0.19	0.31	0.21	0.31	-0.20	0.18
Not evaluated	-0.04	0.11					0.08	0.14
Child-to-staff ratio (3-year-old - 8:1 or lower)	0.10	0.05*	0.21	0.16	0.22	0.16	0.08	0.05
Teacher education (vs. less than 25 percent)								
50 percent or greater with some college or higher	0.09	0.10	0.08	0.20	0.07	0.20	0.26	0.14+
25 to 50 percent with some college or higher	0.05	0.10	0.04	0.17	0.05	0.17	0.22	0.15
Director education (some college or higher vs. less)	0.07	0.08	0.08	0.16	0.05	0.16	0.10	0.09
Family partnerships (all vs. less than all)	0.03	0.08	0.14	0.20	-0.12	0.26	0.03	0.08
Staff management	-0.02	0.12	0.57	0.36	0.58	0.35	-0.30	0.16+
Salary scale	0.39	0.11***	0.38	0.20+	0.32	0.21	0.39	0.14**
Accredited center	0.23	0.07**	0.28	0.15+	0.26	0.15+	0.21	0.09*
Individualization; assessment, planning for special needs, family communication (vs. two or less)								
More than half	0.30	0.14*			0.41	0.25		
Not evaluated	1.32	0.15***						
R-square	0.27		0.13		0.14		0.03	
Adjusted R-square	0.26		0.09		0.10		0.02	
Number of centers	1327		227		227		1100	

Source: Analysis of data from Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Table IV.5. Results of OLS Regressions of Quality Rating Components Among Center- based Providers with Average ECERS- R Provisions for Learning Scores, by Cross- Site and Site- Specific Models

	Cross-site model		Florida, Miami-Dade				Tennessee	
			Cross-site common components		Individualization components			
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	3.02	0.15***	2.87	0.26***	2.87	0.26***	4.07	0.21***
Curriculum-staff training (vs. not all staff)								
All teaching staff	-0.02	0.11	-0.20	0.14	-0.21	0.14	0.23	0.16
Some teaching staff	-0.32	0.15*	-0.08	0.25	-0.08	0.25	-0.18	0.21
Not evaluated	0.04	0.11					0.28	0.16+
Child-to-staff ratio (3-year-old - 8:1 or lower)	0.07	0.05	0.16	0.12	0.16	0.12	0.03	0.06
Teacher education (vs. less than 25 percent)								
50 percent or greater with some college or higher	0.01	0.10	-0.13	0.16	-0.13	0.16	0.16	0.17
25 to 50 percent with some college or higher	-0.04	0.10	0.00	0.13	0.00	0.13	0.07	0.18
Director education (some college or higher vs. less)	0.09	0.08	0.10	0.12	0.09	0.13	0.12	0.11
Family partnerships (all vs. less than all)	0.15	0.08+	0.22	0.16	0.15	0.20	0.12	0.09
Staff management	0.22	0.13	0.37	0.28	0.37	0.28	0.03	0.19
Salary scale	0.54	0.12***	0.53	0.16**	0.51	0.16**	0.48	0.17**
Accredited center	0.31	0.08***	0.34	0.12**	0.33	0.12*	0.29	0.10**
Individualization; assessment, planning for special needs, family communication (vs. two or less)								
More than half	0.06	0.15			0.11	0.20		
Not evaluated	1.15	0.16***						
R-square	0.27		0.24		0.24		0.06	
Adjusted R-square	0.26		0.20		0.20		0.05	
Number of centers	1327		227		227		1100	

Source: Analysis of data from Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table IV.6. Results of OLS Regressions of Quality Rating Components Among Center-based Providers with Average ECERS-R+ITERS-R Total Scores, by QRIS

	Florida, Miami-Dade		Tennessee	
	Coefficient	SE	Coefficient	SE
Constant	2.49	0.20***	4.09	0.32***
Curriculum-staff training (vs. not all staff)				
All	-0.04	0.15	0.17	0.12
Some	0.08	0.19	-0.10	0.15
Not evaluated			0.24	0.12+
Child-to-staff ratio (3-year-old - 8:1 or lower)	0.09	0.10	0.11	0.04**
Teacher education (vs. less than 25 percent)				
50 percent or greater with some college or higher	0.25	0.13+	0.35	0.12**
25 to 50 percent with some college or higher	0.10	0.10	0.23	0.13+
Director education (some college or higher vs. less)	0.03	0.10	0.04	0.08
Family partnerships (all vs. less than all)	0.04	0.20	0.12	0.07+
Staff management	0.07	0.24	-0.10	0.14
Salary scale	0.25	0.15	0.46	0.21*
Accredited center	0.30	0.09**	0.36	0.07***
Child assessment screening	0.18	0.23		
Child assessment guides planning	-0.05	0.18		
Planning for special needs	0.27	0.21		
Communicate with families in home language	-0.30	0.23		
Program management	0.50	0.29+		
Staff benefits			-0.08	0.46
R-square	0.33		0.12	
Adjusted R-square	0.29		0.11	
Number of centers	227		1100	

Source: Analysis of data from Miami-Dade, Florida and Tennessee conducted as part of the QRS Assessment project.

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

V. QUALITY PROFILES: PATTERNS IN QUALITY COMPONENTS

A. Analytic Goals

The previous chapters provided information on the prevalence of individual quality rating components and their unique contributions in predicting observed quality. However, the influence on quality and, in turn, child outcomes may best be accounted for by considering more than one component at a time. We begin to explore a composite view, describing the patterns for a subset of cross-QRIS quality components across two QRIS (Florida, Miami-Dade and Tennessee), by addressing the following questions:

- What is the prevalence of centers demonstrating select components on the high end of quality?
- How might centers be categorized into profiles of quality based on the number and combination of quality components that demonstrate significant associations with observed quality?
- Do the profiles map to existing rating levels in each of the two QRIS in ways that create distinctions in quality?

Using the information from the analysis presented in prior chapters, we selected five quality components from which to build profiles of providers in the two QRIS. Four components—use of a differentiated salary scale, family partnership strategies, teacher education levels, and child-to-staff ratios—consistently or frequently demonstrated unique contributions in predicting observed quality in the models discussed in Chapter IV. We add observed quality (the ERS score) back as a component, rather than an outcome, to contribute to the profiles since the learning environment is often included as a rating component in QRIS. In general, these five components are among those that are commonly included in rating systems (Caronongan et al. 2011) and therefore are relevant to a broader set of QRIS.

From the profiles, we then connect back to the current overall quality rating levels in each of the two QRIS. This allows us to bring the knowledge gained through the cross-site analysis and use of the common metric for quality components to bear on the actual rating systems in each site. Specifically, this connection examines the extent to which a small set of components that are associated with observed quality may also be drivers in defining distinctions in quality across varied QRIS.

B. Analytic Strategy and Considerations

1. Selecting cross-QRIS quality profile components

Each QRIS includes numerous indicators that fall across several quality components, as described in Chapter I and throughout the preceding chapters. With this analysis, we focused on a subset of five quality components commonly evaluated in QRIS as follows:[21]

[21] The latter four components demonstrated unique effects with observed quality, as described in Chapter IV.

- *Learning environment* as measured by the ERS combined score (5.0 or higher)
- *Structural quality* as measured by the center–average ratio for 3-year-old classrooms of 8:1 or lower
- *Staff qualifications* as measured by at least 50 percent of the teaching staff having some college (such as enrollment toward a Child Development Associate)
- *Administration/management* as measured by offering a salary scale differentiated by staff education, training, and experience
- *Family partnerships* as measured by a center offering all activities evaluated by the QRIS

For each component, we used a threshold for maximizing quality. Four of these matched our definition in previous chapters. For the fifth component—the learning environment—we used the cut-point of 5.0, which is defined by the ERS instrument developers as indicating "good" quality (Harms et al. 2005) on the combined ECERS-R and ITERS-R facility-level score. Such a score may best approach what several researchers describe as an active zone, above which there could be a greater ability to influence child outcomes (Burchinal et. al 2011). This ERS cut-point also reflects the lowest threshold requirement for center-based providers to meet the highest rating level for the learning environment in the two QRIS included in the analysis.

In reviewing the key findings for patterns of quality components, our analytic strategy included these important considerations:

- As in Chapters III and IV, we explored quality profiles based on centers in Florida, Miami-Dade and Tennessee QRIS. This analysis included a combined 1,334 centers with all quality components examined.
- We defined profiles based on a select 5 of the full 17 cross-QRIS quality components, limiting to those that are common to all centers in the two QRIS and that demonstrated unique contributions in predicting observed quality in the models described in Chapter IV.
- Our method was akin to an index based on the number of components present, looking at the exact pattern or combination of those components in the data for this sample. Alternative methods include theoretical constructs aligning with optimal quality or empirically driven techniques such as cluster or latent class analyses, which require stronger assumptions on variable distribution. (See the methodological notes in Appendix D for additional discussion of our method.)
- While we present these profiles across two QRIS, the subset of centers that meet all five quality component thresholds came from just one of the two.

C. Key Findings

Based on five selected quality components and associated thresholds, centers in two QRIS fall into seven distinct quality profiles. The profiles are differentiated by the (1) number of quality components that meet the thresholds (that is, one to five), (2) ability of the center to meet the learning environment threshold of a combined ERS score of 5.0 or higher, and (3) ability to meet the pairing of the family partnership component and use of a differentiated salary scale.

- The most prevalent profile, in which 27 percent of centers fall, is the one in which all five quality component thresholds are met. This suggests that for nearly three-quarters of the centers, these five quality components and associated thresholds are sufficiently demanding, demonstrated by the variation in centers meeting different numbers and combinations of the thresholds.
- The next most common profile demonstrates mixed components. The 24 percent of centers in this profile did not meet the learning environment threshold, but have the pairing of the family partnership and salary scale components. They may also meet the threshold for one other component (either teacher education or child-to-staff ratio, but not both).
- Very few centers—only three percent—do not meet any of the five quality component thresholds.
- The overall quality rating levels assigned to centers in each of the two QRIS map to the profiles such that providers assigned to the higher quality profiles have higher rating levels, and vice versa. This suggests that the five components and the associated thresholds may indeed be similar drivers in creating distinctions in quality across different QRIS.

D. Detailed Findings

1. What is the prevalence of centers demonstrating select components, individually or in combination, on the high end of quality?

- Based on the definitions described above, a substantial majority of centers in the two QRIS use a differentiated salary scale (92 percent) and implement all of the family partnership strategies that are specified by the QRIS (84 percent), as shown in Table V.1.
- Approximately two-thirds of centers in the two QRIS have child-to-staff ratios for classrooms that serve 3-year-old children that are 8:1 or lower (62 percent), and have 50 percent or more of the teaching staff with some college education or higher (67 percent).
- Close to half (44 percent) of the centers in the two QRIS have an ERS combined score of 5.0 or higher.

When examining quality components, it is also useful to understand patterns of scoring and the proportion of programs that are able to meet the various thresholds set for the components. If the majority of centers can meet all of the thresholds, we might assume that the thresholds are not rigorous enough. Likewise, if no centers can meet all of the thresholds, we might assume that they are unrealistic or too difficult. We first examined the patterns among centers for all components rated as low (that is, not meeting the threshold) and all rated high (that is, all thresholds met) to capture the upper and lower quality profiles, and to assess what proportion of centers have a mixed combination of components rated high and low (Table V.2).

- Three percent of the centers do not meet any of the five quality component thresholds.
- About 71 percent centers have a mixed pattern ranging from having one to four of the quality component threshold met—6 percent with one, 14 percent with two, 22 percent

with three, and 29 percent with four. This mixed count comprises 24 unique pattern combinations.

- Across the two QRIS centers, 27 percent of centers met the thresholds for all five quality components.[22] However, those centers came from only one QRIS.

2. How might centers be categorized into profiles of quality based on the number and combination of quality components that demonstrate significant associations with observed quality?

As a first step in developing profiles, reviewing patterns of correlations among quality components provides information about the degree to which components are tapping unique dimensions of quality. High correlations among components would indicate that they are assessing overlapping constructs and that QRIS data collection may be redundant for some components. Overall, the correlations among these five components (as categorized by our thresholds) range from weak to moderate (Table V.3), indicating that the components measured in the QRIS are relatively distinct and unique.

- The correlations between the ERS 5.0 or higher score and other components range from 0.18 (with family partnerships) to 0.30 (with teacher education).
- The ratio component is correlated at 0.15 with use of a differentiated salary scale and family partnerships; it is most strongly correlated with teacher education at 0.26.
- Teacher education (50 percent or more of staff with some college or above) has the highest correlations, though still not more than moderate, with each of the other components ranging from 0.26 with the ratio component to 0.35 with the family partnerships component.
- The strongest correlation detected—0.44—was between family partnerships and the use of a differentiated salary scale.

Next, to better understand the various combinations of quality components that centers are meeting, we created quality profiles. These profiles provide insights into what dimensions co-occur in the production of quality. Identifying these profiles may assist with making advancements in the development of needs assessments, alignment of quality improvement strategies with the diverse needs of centers, and the tracking of center improvements over time. We began with the two upper and lower bounds—centers that met all five component thresholds as defined, and those that did not meet any. We then examined patterns that were the most frequent (occurring for at least five percent of providers) and found that meeting the thresholds for family partnerships and salary scale were generally co-occurring; providers tended to either meet them both, or meet neither. Finally, we categorized centers by whether they met the ERS threshold. These combined steps, produced seven distinct quality profiles across providers in the two QRIS. The profiles are roughly organized by the total number of quality components met in descending order (Table V.4). Additional details are presented in Table G.1 in Appendix G.

[22] Percentage may not equal 100 percent due to rounding.

- **Maximizing quality (Profile A).** The most frequent pattern, prevalent among 27 percent of centers, is that in which all five quality component thresholds are met.
- **High operational quality (Profile B).** The learning environment component (combined ERS score of 5.0 or higher) may be one of the more difficult to meet. This profile includes centers (17 percent) that meet the threshold on four components but do not meet the ERS score threshold.
- **High center environment (Profile C).** Another set of centers (13 percent) met the ERS threshold as well as the family partnerships/differentiated salary scale combination. These three components were met among all centers in this profile and some also met either the threshold for teacher education or the ratio component, but not both.
- **High observed quality plus (Profile D).** While not particularly common (only 4 percent of centers), this profile is still distinct in that these centers met the ERS threshold and also met either the threshold for teacher education or the ratio component, or both. The centers in this profile did not meet the family partnerships/differentiated salary scale combination, but they may meet just one of the two.
- **Mixed center environment (Profile E).** The second largest profile, this includes centers that did not meet the ERS threshold but implemented all family partnership activities specified by the QRIS *and* provided staff with a differentiated salary scale, sometimes meeting one other component as well.
- **Lower observed quality plus (Profile F).** About 11 percent of centers did *not* meet the ERS threshold but had other components. However, they did not possess the family partnerships/differentiated salary scale combination observed in other centers; in some cases only one of the thresholds for these two components were met, but not both. In this profile, centers had one, two, or three component thresholds met, but most commonly only one.
- **Base quality (Profile G).** Few centers (3 percent) had a base level of quality in that they did not meet the thresholds on any of the five components

3. Do the profiles map to existing rating levels in each of the two QRIS in ways that create distinctions in quality?

To assess how meaningful the select group of components from this cross-QRIS analysis may be in defining distinctions in quality within individual QRIS, we examine the patterns of the profiles to the actual rating levels in each Miami-Dade and Tennessee (Tables V.5 and V.6).

- In general, the concentration of providers across the rating levels maps to the quality profiles such that what we have defined as higher quality profiles (based on the components) tend to receive higher rating levels, and lower quality profiles receive low rating levels.
- Because both Miami-Dade and Tennessee rely on points in the determination of ratings, there are multiple ways in which providers may achieve a particular rating. This is particularly noticeable in Tennessee where providers across each of the profiles except base quality achieve the highest rating level, though in decreasing percentages as the profiles descend down the definitions.

- That no provider has achieved the profile of "maximizing quality" by meeting all five component thresholds in Miami-Dade is likely a function of the relative newness of the QRIS. Relatively small percentages of providers have achieved level four and five ratings in Miami-Dade (25 percent combined), but even fewer (less than 5 percent) have met the criteria for the top four quality profiles as we have defined them.

Table V.1. Percentage of Centers Demonstrating Quality Profile Components Across Two QRIS

Quality profile component	Percentage of Centers
ERS combined score of 5.0 or higher	44.0
Ratio of 8:1 or lower for 3-year-old classrooms	62.1
Teacher education: 50% or more of teaching staff with some college or higher	66.6
All family partnership strategies implemented	84.3
Use of a differentiated salary scale	91.8
Number of centers	1,334

Source: Analysis of data from Florida, Miami-Dade and Tennessee conducted as part of the QRS Assessment project.

Table V.2. Percent of Centers Meeting a Particular Count of Quality Profile Components Across Two QRIS

Number of Quality Profile Components Met	Percentage of Centers
0	3.2
1	5.6
2	13.5
3	21.7
4	28.9
5	27.1
Number of centers	1,334

Source: Analysis of data from Florida, Miami-Dade and Tennessee conducted as part of the QRS Assessment project.

Table V.3. Correlations Among Quality Profile Components Across Two QRIS

	1	2	3	4	5
ERS combined score 5.0 or higher	1.00				
Ratio of 8:1 or lower for 3-year-old classrooms	0.21	1.00			
Teacher education 50% or greater with some college or higher	0.30	0.26	1.00		
All family partnerships offered	0.18	0.15	0.35	1.00	
Differentiated salary scale	0.25	0.15	0.30	0.44	1.00
Number of centers					1,334

Source: Analysis of data from Florida, Miami-Dade and Tennessee conducted as part of the QRS Assessment project.

Note: All correlations are significant at $p < .001$.

Table V.4. Quality Profiles Across Two QRIS

Profile		Description	Number of Threshold Components	ERS Meet Threshold	Percentage of Centers
A	Maximizing quality	All components meet threshold	5	Yes	27.1
B	High operational quality	ERS not meet threshold, all others do	4	No	17.0
C	High center environment	ERS meet threshold as well as at least family partnerships and salary scale	3 - 4	Yes	13.2
D	High observed quality plus	ERS meet threshold and have at least one other component, but lack the family partnership-salary scale combination	2 - 4	Yes	3.7
E	Mixed center environment	ERS not meet threshold but have at least family partnerships and salary scale	2 - 3	No	24.4
F	Lower observed quality plus	ERS not meet threshold but have at least one component, but lack the family partnership-salary scale combination	1 - 3	No	11.4
G	Base quality	No components meet thresholds	0	No	3.2
Number of centers					1,334

Source: Analysis of data from Florida, Miami-Dade and Tennessee conducted as part of the QRS Assessment project.

Table V.5. Percentage of Centers at Each Quality Profile by Rating Level, Miami- Dade

		Level 1	Level 2	Level 3	Level 4	Level 5	Number of Centers
A	Maximizing quality	0.0	0.0	0.0	0.0	0.0	0
B	High operational quality	0.0	0.0	0.9	1.3	0.4	6
C	High center environment	0.0	0.0	0.0	0.9	1.3	5
D	High observed quality plus	0.0	0.0	0.0	0.4	0.0	1
E	Mixed center environment	0.0	8.1	26.9	17.5	3.4	131
F	Lower observed quality plus	2.1	11.5	9.8	1.3	0.4	59
G	Base quality	5.1	6.4	2.1	0.0	0.0	32
Number of centers		17	61	93	50	13	234

Source: Analysis of data from Miami-Dade, Florida as part of the QRS Assessment project.

Note: Shaded cells indicate the highest concentration of providers for each quality profile.

Table V.6. Percentage of Centers at Each Quality Profile by Rating Level, Tennessee

		Level 1	Level 2	Level 3	Number of Centers
A	Maximizing quality	0.0	0.1	32.8	362
B	High operational quality	0.0	0.8	19.3	221
C	High center environment	0.0	1.4	14.2	171
D	High observed quality plus	0.0	2.5	1.9	48
E	Mixed center environment	0.1	7.4	10.2	194
F	Lower observed quality plus	0.8	5.5	2.2	93
G	Base quality	0.8	0.2	0.0	11
Number of centers		19	195	886	1100

Source: Analysis of data from Tennessee as part of the QRS Assessment project.

Note: Shaded cells indicate the highest concentration of providers for each quality profile.

VI. SUMMARY AND FUTURE DIRECTIONS FOR RESEARCH

This analysis covered a number of topics related to quality measurement through use of existing administrative data from a small and select number of QRIS. As a descriptive, exploratory analysis it provides useful and interesting information, but it also identifies areas in need of further research to add both context and evidence. In this last chapter, we discuss a number of key topics as they relate to the findings of this work and needed next steps in efforts to understand how QRIS functions to improve quality among child care programs and, ultimately, to improve outcomes for children.

A. Programs in the QRIS

An important issue in the evaluation of QRIS as well as the interpretation of any research and evaluation is which centers select into voluntary QRIS. Specifically, do centers that are high quality based on the definition of the QRIS choose to participate at a higher rate than those that may be rated as low quality? QRIS stakeholders and evaluators need to further explore and understand the characteristics of programs that participate in QRIS and those that do not in order to fully examine changes in quality and, ultimately, changes in child outcomes that may be associated with QRIS.

The analysis of child care centers participating in QRIS, by Head Start and accreditation status in Miami-Dade (presented in Chapter III), provides findings that undergird the importance of understanding the patterns of participation in the QRIS and the potential differences between programs that participate and those that do not. A number of QRIS tend to view accreditation as the highest level and build a lattice work through the rating levels to support programs in achieving accreditation, or at least standards that are similar (Caronongan et al. 2011). On its surface, the analysis of data from Miami-Dade could be supportive of this approach, but it requires further context and exploration based on the following:

- Accredited centers in Miami-Dade are more than twice as likely as non-accredited centers to be at the highest rating levels and have at least a half-point higher ERS score, on average (across all five specifications of ERS scores examined) than centers that are not accredited. While less than 25 percent of participating child care centers have achieved levels 4 or 5 in Miami-Dade, accredited programs comprise 74 percent of the centers that have. As a voluntary QRIS, accredited centers that believe they can demonstrate the components to meet the highest levels may be the ones that pursue the rating. From this analysis, we are unable to determine whether the accredited programs that chose not to participate in the QRIS are similar to those that do.

- Similar to accredited programs, Head Start programs in Miami-Dade are more likely than other programs to demonstrate quality components such as those reflecting individualization practices, family partnership strategies, and higher education levels among directors and teaching staff. Yet, the differences between Head Start centers and other centers in the distribution across rating levels and on ERS scores are minimal. In Miami-Dade, QRIS participation is essentially not voluntary for Head Start programs; all programs are strongly encouraged to participate in the QRIS based on their contracts with the Community Action Agency, an agency that oversees Head Start programs in Miami-Dade county. In fact, 85 percent of Head Start programs currently have a QRIS rating and those that do not are generally in a self-study, pre-rating phase. The near full participation of Head Start programs in the QRIS compared to only partial participation among accredited centers may limit generalizability of our findings in ways we cannot fully understand. Similarly, we urge caution in interpreting the estimated contribution of

accreditation in predicting observed quality based on the data from Miami-Dade. As suggested in Chapter IV, accredited providers that participate in the QRIS may have unobserved characteristics (such as motivation and resources) that are influencing this result.

B. Patterns in Quality Improvement

Based on data collected from a cross-section of child care centers in a few, select QRIS it appears that there are distinctions between the base, middle, and high rating levels in that providers at the higher rating levels are more likely to demonstrate the quality components examined than those at the lower levels. What we cannot determine from this analysis, however, is whether and how these components have changed for providers over time. That is, we cannot assess the degree to which this snapshot of providers by level may be an indication of patterns that have persisted or evolved over time.

This analysis raises the possibility of an accumulation effect—that even in combination or points systems providers may continue to improve on a broader set of quality components even after the highest rating has been achieved. Although we cannot substantiate this effect with certainty from a cross-sectional analysis of providers, we see differences in components across the rating levels that are more pronounced (or consistent) in one QRIS than in another. For example, the comparison of quality components between each of the rating levels (base to middle, middle to high) in Tennessee indicates a fairly consistent pattern that providers at each higher rating level are more likely than providers at the level below to demonstrate each of the quality components. In Miami-Dade, the differences between the base to middle rating levels are generally more pronounced in magnitude (and many were statistically significant) compared to differences between centers at the middle versus highest rating levels which were smaller (and fewer were statistically significant). The difference in patterns observed in the two sites may be due to the maturity of the Tennessee QRIS as compared to that of Miami-Dade. While providers are beginning to achieve the highest levels in Miami-Dade, they may do so by gaining points for components that take less time to meet. For example, significantly higher percentages of centers at the highest rating level in Miami-Dade conduct child assessment screening, communicate with families in their home language, plan for children with special needs, and implement all family partnership strategies specified by the QRIS compared with those at the middle rating level. However, we did not find statistically significant differences in the percentages of centers at the middle and highest rating levels on training all teaching staff on the curriculum, using child assessments to guide planning, and having 50 percent or higher of all teaching staff with some college education or more. It is possible that even if they have already reached the minimum points required for the highest level, these providers in Miami-Dade will eventually demonstrate more of the components associated with the highest level.

This potential accumulation effect might also be reflected in the analysis of quality profiles and rating levels for Tennessee (Table V.6). Results from that analysis show that while some centers at the highest QRIS rating level are at the middle quality profiles we have defined, substantial proportions of providers are in the higher quality profiles. Again, we cannot assess whether these providers improved over time or entered into the QRIS with these characteristics.

The patterns in the quality profiles and their connections with the actual rating levels in both Miami-Dade and Tennessee also suggest that the maturity of a QRIS may influence the distribution of providers on the quality spectrum, although again this cannot be confirmed without examining providers' ratings over time. Specifically, only a handful of providers in Miami-Dade have achieved the thresholds for the five quality components that make up the top quality profiles and similarly,

few providers have gained a level 5, the highest QRIS rating. In Miami-Dade, the highest concentration of providers is in the lower quality profiles and lower QRIS rating levels (Table V.5). In contrast, in Tennessee, the concentration of providers is generally at the other end of the spectrum, both in the profiles and the actual rating levels (Table V.6). As a relatively new QRIS, providers in Miami-Dade may still be working their way up through the levels and quality profiles whereas in Tennessee providers appear to be increasingly moving toward the top rating and higher profiles.

C. Rating Levels

The fact that we find more similar proportions of centers meeting quality components when we compare centers at the *highest* rating level in the two sites (Table III.5) (except for teacher education) builds on the findings in our companion qualitative report that QRIS requirements are more similar at the highest level (Caronongan et al. 2011). This may provide some confidence in the face validity of QRIS indicators and suggest that they are implemented with some consistency across QRIS. (Though we stress that these findings are based on just two QRIS.)

Nonetheless, this supports the importance of continuing to conduct validation studies of QRIS with a focus on the highest level, at least to start, and an examination of child outcomes as well as other indicators such as parent satisfaction and service to low-income children. As mentioned above, QRIS can vary considerably in how quality components enter into ratings. The path to the highest level may be very different based on each state's licensing and subsidy policies, adoption of early learning guidelines to guide curriculum selection and use, and strength of the supporting professional development system.

Validation studies also gain importance as a number of QRIS have been in place for close to 10 years and are reaching a point where a majority of providers are already at the highest level. For example, in Tennessee, it will be important to examine if the highest rating level is associated with positive outcomes for children given that most providers are already at that level. Alternatively, are the requirements at the highest level too easily met by programs such that quality is not sufficiently strong to show linkages to children's outcomes? If so, should QRIS standards be refined to include additional levels? As systems mature, the availability of data on providers' ratings over time can also inform the design of systems. Examination of changes in provider ratings over time may help uncover requirements or indicators that are particularly challenging for providers to achieve and act as barriers to progressing to higher levels. These patterns of ratings can be important motivations for evaluation of QRIS to determine the design features related to being "stuck" at particular levels and understanding the outcomes that relate to different levels.

D. Quality Components

Critical to decisions of design and refinement of QRIS are issues related to which quality components to measure and in what combination. This analysis provides some suggestive findings from two QRIS. The larger contribution of this work may be the introduction of a framework for using a common metric across QRIS that can be used in future evaluations to continue building knowledge in this area.

The findings from this analysis on the unique predictors of observed quality build on existing knowledge by using data collected in QRIS practice and not in traditional research contexts. The analytic approach and research questions had to be refined based on the data available. Specifically, we predicted variation in ERS scores as a function of other quality rating components. Overall, the

patterns of findings in this analysis confirm the importance of domains long considered predictors of quality (such as teacher education, child-to-staff ratios, and salary). The findings also support the inclusion of other domains with a smaller evidence base in early childhood, such as the use of family partnership strategies and of practices that promote the individualization of services for children and families.

Five quality components consistently demonstrated unique contributions in predicting observed quality across the various models tested (different ERS scores, cross-site, and site specific): (1) accreditation status, (2) use of a salary scale that is differentiated by staff education, training, and experience, (3) the percentage of teaching staff with some college education or more, (4) child-to-staff ratios, and (5) the use of the full range of family partnership strategies specified by the QRIS. Based on these findings, we honed the analysis in on five components—excluding accreditation because it is often equated with the highest rating level in QRIS, but including the ERS as a component to build quality profiles. The profiles were developed using thresholds for the quality components as defined on a common metric.

Evidence from the correlations among the five components used to build the quality profiles supports the assumption that the quality components common to QRIS are measuring related but distinct aspects of program quality. This pattern of findings is similar to that noted in the evaluation of Colorado's Qualistar QRIS (Zellman et al. 2008). The actual rating levels in each of the two QRIS map to the profiles such that providers mapped to the higher quality profiles have higher rating levels, and vice versa. This suggests that the five components and the associated thresholds may indeed be similar drivers in creating distinctions in quality across different QRIS. These findings are correlational and not confirmatory; they are not intended to advise that only these five components are critical in a QRIS. It is possible that the increase in the presence of these five components also leads to or is associated with other components (that are measured in these two QRIS or not) in important ways that could not be identified.

These findings do help structure an agenda for next steps in both research and, potentially practice, through further examination of the resources necessary for quality improvement, measurement refinement of particular components, and definition to emerging components. For example, two of the five components depend on a strong professional development system to support quality at a system level—to enhance the level of education among teaching staff and to possibly promote a more stable workforce through salary scales that support retention and advancement. Further research into the level and types of resources it takes to move a provider from one rating to the next, or more broadly from the lowest to the highest rating level, is needed to contribute to an analysis of costs and benefits associated with these types of quality improvements.

We also need to better formulate what is important to implement around family partnerships that can be meaningful for children's outcomes. The emphasis on family partnerships is gaining an important role in many QRIS, but it is difficult to know what is being captured through the range of strategies and general checklist nature of the current measures in QRIS requirements. Similarly, this analysis also suggests that the growing emphasis on strategies toward individualization practices around child assessment, special needs, and accommodating children from culturally diverse backgrounds is worth further pursuit. The cross-site models indicated that the presence of multiple practices may be a factor in predicting observed quality. These individualization practices are not currently evaluated by many QRIS, or when they are, it is generally only at the higher rating levels.

E. Quality Outcome

Lastly, but most importantly, the field needs to continue building knowledge to better understand how the ERS are related to the other components of quality that are measured in QRIS, as well as to continue to seek other outcome measures that capture the dimensions of quality most closely linked with child outcomes.

Previous studies have documented associations between ECERS-R total and factor scores and other measures of quality such as the CLASS (Pianta et al. 2005) and Early Language and Literacy Classroom Observation (ELLCO; Clifford et al. 2009). Researchers have also found correlations between the ECERS-R and child language, cognitive, and socio-emotional outcomes (Burchinal et al. 2008). These findings suggest that variation in ERS associated with QRIS participation (as evidenced by higher scores for centers with higher ratings) signifies that QRIS standards are consistent with what extant research suggests is appropriate.

That said, there are outstanding issues that require further inquiry before the full effects of QRIS can be quantified. First, determining whether QRIS is improving quality requires a longitudinal study of centers. As discussed, a cross-sectional study provides a snapshot of current distributions of quality but does not provide information on whether providers at the lower end of the spectrum have improved over time. Second, there is limited research on the use of ERS in high-stakes contexts. It remains to be seen whether previous findings regarding ERS and other measures of quality as well as child outcomes would persist in the QRIS environment where scores are tied to ratings and incentives (Perlman, Zellman, and Le 2004; Hamre and Maxwell 2011). This question must be investigated directly through validation and evaluation studies of QRIS. Third, few studies have accounted for selection bias in studying relationships between ERS and child outcomes which makes it difficult to determine whether differences between children in high- and low-quality settings (or QRIS participating and nonparticipating providers) is due to pre-existing differences or to quality itself.

REFERENCES

Administration for Children and Families. "Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start." Final Technical Report. Washington, DC: U.S. Department of Health and Human Services, 2002.

Burchinal, Margaret, Yange Xue, Hsiao-Chuan Tien, Anamarie Auger, and Andrew Mashburn. "Testing for threshold in associations between child care quality and child outcomes." Presentation at the Society for Research in Child Development Biennial Meeting, Montreal, March 31, 2011.

Burchinal, Peg, Kirsten Kainz, Karen Cai, Kathryn Tout, Martha Zaslow, Ivelisse Martinez-Beck, and Colleen Rathgeb. "Early Care and Education Quality and Child Outcomes." Washington, DC: Child Trends, 2009.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., et al.. "Predicting Child Outcomes at the End of Kindergarten from the Quality of Pre-Kindergarten Teacher-Child Interactions and Instruction." *Applied Developmental Science*, vol. 12, no. 3, 2008, pp. 140–153.

Clifford, R. M., Reszka, S. S., & Rossbach, H-G. . "Reliability and Validity of the Early Childhood Environment Rating Scale." Draft version of a working paper. Chapel Hill, NC: University of North Carolina at Chapel Hill, Frank Porter Graham Child Development Institute, 2009.

Clifford, R., O. Barbarin, F. Chang, D. M. Early, D. Bryant, C. Howes, M. Burchinal, and R. Pianta. "What Is Pre-Kindergarten? Characteristics of Public Pre-Kindergarten Programs." *Applied Developmental Science*, vol. 9, no. 3, 2005, pp. 126–143.

Cost, Quality and Outcomes Study Team. "Cost, Quality, and Child Outcomes in Child Care Centers." Denver, Colorado: Economics Dept., University of Colorado at Denver, 1995.

Caronongan, Pia, Gretchen Kirby, Lizabeth Malone, and Kimberly Boller. "Defining and Measuring Quality: An In-Depth Study of Five Quality Rating and Improvement Systems." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation. August 2011.

Early, Diane, Oscar Barbarin, Donna Bryant, Margaret Burchinal, Florence Chang, Richard Clifford, Gisele Crawford, Wanda Weaver, Carollee Howes, Sharon Ritchie, Marcia Kraft-Sayre, Robert Pianta, and W. Steven Barnett. "Pre-Kindergarten in Eleven States: NCEDL's Multi-State Study of Pre-Kindergarten and Study of State-Wide Early Education Programs (SWEEP)." Preliminary Descriptive Report. NCEDL Working Paper. Chapel Hill, NC: University of North Carolina, Chapel Hill, May 24, 2005.

Elicker, James, Carolyn C. Langill, Karen Ruprecht, and Kyong-Ah Kwon. "Paths to Quality: A Child Care Quality Rating System for Indiana. What is its Scientific Basis?" West Lafayette, IN: Purdue University. 2007.

Florida Department of Children and Families. "List of Approved Accrediting Agencies." Accessed May 10, 2011 from [http://www.dcf.state.fl.us/programs/childcare/goldseal.shtml].

Halle, Tamara, Nicole Forry, Elizabeth Kair, Kate Perper, Laura Wandner, Julia Wessel, and Jessica Vick. "Disparities in Early Learning and Development: Lessons from the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B), Executive Summary." Washington, DC: Child Trends, 2009.

Hamre, Bridget K., and Kelly L. Maxwell. "Best Practices for Conducting Program Observations as Part of Quality Rating and Improvement Systems." Research-to-Policy, Research-to-Practice Brief OPRE 2011-11b. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. June 2011.

Harms, Thelma, Richard M. Clifford, and Debby Cryer. "Early Childhood Environment Rating Scale, Revised Edition." New York, NY: Teachers College Press, 2005.

Harms, Thelma, Debby Cryer, and Richard M. Clifford. "Infant/Toddler Environment Rating Scale, Revised Edition." New York, NY: Teachers College Press, 2006.

Harms, Thelma, Debby Cryer, and Richard M. Clifford. "Family Child Care Environment Rating Scale, Revised Edition." New York, NY: Teachers College Press, 2007.

Harms, Thelma, Ellen V. Jacobs, and Donna Romano. "The School-Age Care Environment Rating Scale." New York, NY: Teachers College Press, 1995.

Love, John M., Ellen E. Kisker, Christine Ross, Helen Raikes, Jill Constantine, Kimberly Boller, Jeanne Brooks-Gunn, Rachel Chazan-Cohen, Louisa Banks Tarullo, Christy Brady-Smith, Allison S. Fuligni, Peter Z. Schochet, Diane Paulsell, and Cheri Vogel. "The Effectiveness of Early Head Start for 3-Year-Old Children and Their Parents: Lessons for Policy and Programs." *Developmental Psychology,* vol. 41, no. 6, 2005, pp. 885–901.

Love, John M., Linda Harrison, Abraham Sagi-Schwartz, Marinus H. van IJzendoorn, Christine M. Ross, Judy A. Ungerer, Helen H. Raikes, Christy Brady-Smith, Kimberly Boller, Jeanne Brooks-Gunn, Ellen E. Kisker, Diane C. Paulsell, and Rachel Chazan-Cohen. "Child Care Quality Matters: How Conclusions May Vary with Context." *Child Development*, vol. 74, no. 4, 2003, pp. 1021–1033.

Mulligan, Gail M., and Kristin Denton Flanagan. "Age 2: Findings From the 2-Year-Old Follow-up of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)." (NCES 2006–043). Washington, DC: National Center for Education Statistics, 2006.

Norris, Deborah J., and Loraine Dunn. "'Reaching for the Stars: Family Child Care Home Validation Study Final Report." Stillwater, OK and Norman, OK: Early Childhood Collaborative of Oklahoma, 2004.

Norris, Deborah J., Loraine Dunn, and Lisa Eckert. "'Reaching for the Stars: Center Validation Study Final Report." Stillwater, OK and Norman, OK: Early Childhood Collaborative of Oklahoma, 2003.

Perlman, Michael, Gail L. Zellman, and Vi-Nhuan Le. "Examining the Psychometric Properties of the Early Childhood Environment Rating Scale-Revised (ECERS-R)." *Early Childhood Research Quarterly*, vol. 19, 2004, pp. 398-412.

Pianta, R., K. LaParo, and B. Hamre. *The Classroom Assessment Scoring System Pre-K Manual*, Charlottesville, VA: University of Virginia, 2008.

Pianta, R., Howes, C., Burchinal, M. R., Bryant, D., Clifford, R., Early, D., et al. "Features of Pre-Kindergarten Programs, Classrooms, and Teachers: Do They Predict Observed Classroom Quality and Child-Teacher Interactions?" *Applied Developmental Science*, vol. 9, no. 3, 2005, pp. 144-159.

Ross, Christine, Emily Moiduddin, Cassandra Meagher, and Barbara Carlson. "The Chicago Program Evaluation Project: A Picture of Early Childhood Programs, Teachers, and Preschool-Age Children in Chicago." Princeton, NJ: Mathematica Policy Research, December 2008.

Smart Start Evaluation Team. "Validating North Carolina's 5-Star Child Care Licensing System." Chapel Hill, NC: Frank Porter Graham Child Development Institute, University of North Carolina, Chapel Hill, 2001.

Talan, Teri N. and Paula Jorde Bloom. *Business Administration Scale for Family Child Care*. New York: Teachers College Press, 2009.

Talan, Teri N. and Paula Jorde Bloom. *Program Administration Scale: Measuring Early Childhood Leadership and Management*. New York: Teachers College Press, 2004.

Tout, Kathryn, Rebecca Starr, Margaret Soli, Shannon Moodie, Gretchen Kirby, and Kimberly Boller. "Compendium of Quality Rating Systems and Evaluations." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation. 2010.

Tout, Kathryn, Martha Zaslow, Tamara Halle, and Nicole Forry. "Issues for the Next Decade of Quality Rating and Improvement Systems." Washington, DC: Child Trends, 2009.

West, Jerry, Louisa B. Tarullo, Nikki L. Aikens, Susan C. Sprachman, Christine M. Ross, and Barbara L. Carlson. "FACES 2006: Study Design." Washington, DC: U.S. Department of Health and Human Services, 2007. Available at: [http://www.acf.hhs.gov/programs/opre/hs/faces/reports/faces_studydesign/faces_studydesign.pdf].

Zaslow, M., R. Anderson, Z. Redd, J. Wessel, L. Tarullo, and M. Burchinal. "Quality Dosage, Thresholds, and Features in Early Childhood Settings: A Review of the Literature, OPRE 2011-5." Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2010.

Zellman, G. L., M. Perlman, V.-N. Le, and L. M. Setodji. "Assessing the Validity of Qualistar Early Learning Quality Rating and Improvement System as a Tool for Improving Child Care Quality." Santa Monica, CA: RAND Corporation, 2008.