
Massachusetts Quality Rating and Improvement System (QRIS) Validation Study Final Report

A project funded by the Massachusetts Department of Early Education and Care with funding from a U.S. Department of Education Race to the Top – Early Learning Challenge Grant



UMass Donahue Institute
Applied Research & Program
Evaluation

Shaping a
Better World
through
Research
& Action
**Wellesley
Centers for
Women**

 MASSACHUSETTS
Department of Early Education and Care

The Wellesley Centers for Women and the UMass Donahue Institute would like to extend their sincere appreciation to the many people who provided critical feedback, perspectives, and helpful suggestions to this study. In particular, we would like to thank the Massachusetts Department of Early Education and Care, members of the study Advisory Board, and the directors, teachers, and support staff from center-based early education and care programs around the state who contributed the information required for this research. We would also like to thank EEC Commissioner Thomas Weber for his ardent support and Gwen Alexander, Anita Moeller and Amy Whitehead-Pleaux who were vital to ensuring the study was carried out with the highest degree of quality and success.

Project Staff



UMass Donahue Institute
Applied Research & Program Evaluation

Joanne Roberts, Senior Research Scientist
Nancy L. Marshall, Associate Director
Allison Tracy, Methodologist

Stefanie Santaniello, Research Manager
Mariana Gerena Melia, Research Manager
Hadley Moore, Research Manager
Steve Ellis, Senior Research Manager
Lonnie Kaufman, Senior Research Manager
**

John Tapper, Consulting Statistician
Sue Leibowitz, Senior Research Manager
Bryan Comer, Project Assistant
Kristen Gibbons, Project Coordinator
Alison Glasgow, Senior Research Analyst
Khaled Khelifi, Report Editor

Report Information

This study was funded by the Massachusetts Department of Early Education and Care with funding from a US Department of Education Race to the Top – Early Learning Challenge Grant. The study was conducted in partnership by the Wellesley Centers for Women and the UMass Donahue Institute.

About the Wellesley Centers for Women

The Wellesley Centers for Women (WCW) is the largest academic, women-and gender-focused, social-change-driven, research-and-action institute in the United States, located at Wellesley College. Scholars at WCW advance gender equality, social justice, and human wellbeing through high-quality research, theory, and action programs. Areas of work include: Education, Economic Security, Mental Health, Youth and Adolescent Development, and Gender-Based Violence. By sharing this work with policymakers, educators, practitioners, advocates, the media, and the public, WCW helps to shape a more just and equitable society.

About the UMass Donahue Institute

The University of Massachusetts Donahue Institute is the public service, outreach, and economic development unit of the University of Massachusetts President's Office. The Institute's Applied Research and Program Evaluation group (ARPE) employs a broad spectrum of applied social science research methods to the study of diverse subject areas. The group's work engages government agencies, municipalities, school districts, community organizations, businesses, philanthropic foundations and non-profit agencies. ARPE studies commonly include formative research to inform program management and decision-making, summative research focused on impact assessment, and process evaluation and monitoring in support of program accountability.

Acknowledgments..... 1

Table of Contents.....2

Table of Tables.....6

Table of Figures.....11

Introduction 12

Quality rating and improvement systems..... 12

Validation of a QRIS 13

Standards and key quality components of the Massachusetts QRIS 14

Support for the MA QRIS and its key components of quality..... 16

Data Collection Design 18

Study sample 18

Program sample 18

Classroom sample 19

Child sample 19

Data collection methods..... 20

Program-level data collected 20

Classroom-level data collected 20

Child-level data collected 21

Study Program Characteristics by MA QRIS level..... 22

Organizational affiliations or type and MA QRIS levels 22

Program size and MA QRIS levels 24

Ages of children served and MA QRIS level..... 25

Participation in subsidy initiatives and MA QRIS levels..... 26

High-needs communities and the MA QRIS 27

Summary and implications 28

Findings 29

Section I: Descriptive analyses of MA QRIS Standards, Criteria, and Verification 30

Overarching considerations..... 30

(1) Better alignment of QRIS criteria and verification and distribution of a single document outlining both criteria and verification 31

(2) Clarification that CEUs or credits are needed for MA QRIS trainings..... 31

(3) Established timeframes for the trainings to reinforce practices..... 31

(4) Reduction of memorandums of understanding as barriers to progresses for programs..... 32

<u>(5) Limitation of the use of overall scales scores for verification and increased focus on specific subscale/items that are more relevant</u>	33
<u>(6) Reduction of compound criteria</u>	33
<u>(7) Incorporation of Continuous Quality Improvement Plans</u>	34
<u>(8) Consolidation of PD requirement into the Professional Development Quality Standard</u>	34
<u>(9) Greater consistency between MA QRIS requirements and other credentialing systems and initiatives</u>	34
<u>Considerations for individual standards and criteria</u>	34
<u>MA QRIS Quality Standard1A: Curriculum, Assessment and Diversity</u>	68
<u>MA QRIS Quality Standard1B: Teacher-Child Relationships and Interactions</u>	73
<u>MA QRIS Quality Standard2: Safe, Healthy Indoor and Outdoor Environments</u>	76
<u>MA QRIS Quality Standard3A: Designated Program Administrator Qualifications and Professional Development</u>	80
<u>MA QRIS Quality Standard 3B: Program Staff Qualifications and Professional Development</u>	87
<u>MA QRIS Quality Standard 4: Family and Community Engagement</u>	91
<u>MA QRIS Quality Standard 5A: Leadership, Management, and Administration</u>	102
<u>MA QRIS Quality Standard 5B: Supervision</u>	109
<u>Summary and implications</u>	117
<u>Section II: Statistical Analyses of MA QRIS Standards and Criteria</u>	118
<u>Criteria differentiation among MA QRIS levels</u>	118
<u>Associations among criteria within a given MA QRIS Quality Standard</u>	120
<u>Summary and implications</u>	121
<u>Section III: Program Quality and the Massachusetts QRIS</u>	122
<u>ERS cut-offs and MA QRIS cut-offs</u>	123
<u>Observed classroom quality and ERS cut-offs</u>	123
<u>Observed quality preschool and ERS cut-offs</u>	123
<u>Observed quality infant and toddler and ERS cut-offs</u>	126
<u>Observed classroom quality and MA QRIS ERS cut-offs</u>	128
<u>Observed preschool quality and MA QRIS cut-offs</u>	129
<u>Observed infant and toddler quality and MA QRIS cut-offs</u>	133
<u>Significant differences in quality by MA QRIS levels</u>	135
<u>Preschool differences in quality by MA QRIS levels</u>	136
<u>Infant and toddler ERS differences by MA QRIS level</u>	140
<u>Item and indicator analyses: Strengths, challenges and licensing</u>	142
<u>Preschool item and indicator analyses</u>	142
<u>Infant and toddler item and indicator analyses</u>	148

<u>The relationship between preschool and infant and toddler classrooms ERS scores</u>	153
<u>Classroom quality and the Arnett CIS</u>	153
<u>Differences in Arnett CIS quality by MA QRIS levels</u>	155
<u>Summary and implications</u>	159
Section IV: Classroom Quality and Key Program Attributes	161
<u>Educator turnover</u>	161
<u>Curriculum</u>	161
<u>Preschool</u>	161
<u>Infant and Toddler</u>	162
<u>NAEYC</u>	162
<u>Preschool</u>	162
<u>Infant and Toddler</u>	163
<u>Head Start</u>	163
<u>Level 4</u>	164
<u>Summary and implications</u>	165
Section V: Education, experience and the Massachusetts QRIS	166
<u>Preschool lead teacher education and classroom quality</u>	166
<u>Infant and toddler lead teacher education and quality</u>	167
<u>Supervisor/administrator education and quality</u>	169
<u>Education, classroom quality, and MA QRIS level</u>	169
<u>Preschool lead teacher education, classroom quality and MA QRIS level</u>	169
<u>Infant and toddler lead teacher education, classroom quality, and MA QRIS level</u>	170
<u>Supervisor/ administrator education, classroom quality and MA QRIS level</u>	171
<u>Experience and quality</u>	171
<u>Experience of the lead teacher and preschool classroom observed quality</u>	171
<u>Experience of the lead teacher and infant toddler classroom observed quality</u>	172
<u>Supervisor/administrator experience and quality</u>	173
<u>Relations among classroom quality and preschool lead teachers teachers' experience and education</u>	173
<u>Relations among classroom quality and infant and toddler lead teachers' teacher experience and education</u>	174
<u>Relations among classroom quality and administrators' experience and education</u>	174
<u>Relations among classroom quality and the education of the lead teacher and support staff</u>	175
<u>Summary and implications</u>	176
Section VI: Child Outcomes and the Massachusetts QRIS	177
<u>Preschool assessment and sample</u>	177

Preschool children’s gains..... 180

Analytic modeling of preschool gains 180

Toddler developmental gains..... 183

Summary and implications 183

Section VII. MA QRIS Program Re-Leveling 184

Re-Leveling and the individual MA QRIS quality standards 186

Re-Leveling and child outcomes 188

Summary and implications 190

Section VIII: Perceptions of the MA QRIS among Programs and Providers 191

Key findings..... 191

Summary and implications 193

Other QRIS Studies..... 194

Recommendations for MA QRIS Revisions and Enhancements..... 196

Recommendations for system-wide levers and supports 196

Recommendations for revisions to system related to classroom quality 197

Recommendations for system revisions related to criteria and verification 197

Recommendations for system revisions to support the workforce 198

Recommendation for system revisions related to the upper tiers of MA QRIS..... 199

Table 1. Key categories of quality in the Massachusetts QRIS	14
Table 2. Number of MA QRIS standards, by quality domain	16
Table 3. Rationale for MA QRIS standards, by quality domain	17
Table 4. Program sample	19
Table 5. Classroom sample by MA QRIS Level	19
Table 6. Child sample: Pre-assessment and post-assessment	20
Table 7. Average number and range of children enrolled by MA QRIS Level	25
Table 8. Timeframe in which most recent training was completed	32
Table 9. Programs delivering services through collaboration and MOU status	33
Table 10. MA QRIS quality standards, criteria and verification process; percent meeting criteria and considerations for refinement	36
Table 11. Programs meeting measured aspects of criterion 1A.2.1, by MA QRIS Level	68
Table 12. Programs meeting measured aspects of criterion 1A.2.2, by MA QRIS Level	69
Table 13. Programs meeting measured aspects of criterion 1A.3.1, by MA QRIS Level	70
Table 14. Programs meeting measured aspects of criterion 1A.3.2, by MA QRIS Level	71
Table 15. Programs meeting measured aspects of criterion 1A.3.3, by MA QRIS Level	71
Table 16. Programs meeting measured aspects of criterion 1A.3.4, by MA QRIS Level	72
Table 17. Programs meeting measured aspects of criterion 1A.4.1, by MA QRIS Level	73
Table 18. Programs meeting measured aspects of criterion 1A.4.2, by MA QRIS Level	73
Table 19. Programs meeting measured aspects of criterion 1B.2.1, by MA QRIS Level	74
Table 20. Programs meeting measured aspects of criterion 1B.3.1, by MA QRIS Level	75
Table 21. Programs meeting measured aspects of criterion 1B.3.2, by MA QRIS Level	75
Table 22. Programs meeting measured aspects of criterion 1B.4.1, by MA QRIS Level	76
Table 23. Programs meeting measured aspects of criterion 2A.2.1, by MA QRIS Level	77
Table 24. Programs meeting measured aspects of criterion 2A.2.2 (Level 2), by MA QRIS Level	78
Table 25. Programs meeting higher level requirements (2A.3.3, Level 3 and 2A.4.1, Level 4) for Personal Care Routines, by QRIS Level	78
Table 26. Programs meeting measured aspects of criterion 2A.3.1, by MA QRIS Level	79
Table 27. Programs that provide or have an agreement in place for screenings, by MA QRIS Level	79
Table 28. Programs meeting measured aspects of criterion 2A.3.2, by MA QRIS Level	79
Table 29. Programs meeting measured aspects of criterion 3A.2.1, by MA QRIS Level	80
Table 30. Reason program meets criterion 3A.2.1	80
Table 31. Programs meeting measured aspects of criterion 3A.2.2, by MA QRIS Level	81
Table 32. Programs meeting measured aspects of criterion 3A.2.3, by MA QRIS Level	81
Table 33. Average years of experience of administrator, by MA QRIS Level	82
Table 34. Programs meeting measured aspects of criterion 3A.2.4, by MA QRIS Level	82
Table 35. Programs meeting measured aspects of criterion 3A.2.5, by MA QRIS Level	83
Table 36. Programs meeting measured aspects of criterion 3A.2.6, by MA QRIS Level	83
Table 37. Programs meeting measured aspects of criterion 3A.3.1, by MA QRIS Level	84
Table 38. Programs meeting measured aspects of criterion 3A.3.2, by MA QRIS Level	84
Table 39. Programs meeting measured aspects of criterion 3A.3.3, by MA QRIS Level	85
Table 40. Reason program meets criterion 3A.3.3, by MA QRIS Level	85
Table 41. Programs meeting measured aspects of criterion 3A.3.4, by MA QRIS Level	85
Table 42. Programs meeting measured aspects of criterion 3A.4.1, by MA QRIS Level	86
Table 43. Programs meeting measured aspects of criterion 3A.4.2, by MA QRIS Level	86
Table 44. Programs meeting measured aspects of criterion 3A.4.3, by MA QRIS Level	87

Table 45. Programs meeting measured aspects of criterion 3A.4.4, by MA QRIS Level	87
Table 46. Programs meeting measured aspects of criterion 3B.2.1, by MA QRIS Level	88
Table 47. Programs meeting measured aspects of criterion 3B.2.2, by MA QRIS Level	88
Table 48. Programs meeting measured aspects of related criteria, by MA QRIS Level	88
Table 49. Programs meeting measured aspects of criterion 3B.2.4, by MA QRIS Level	89
Table 50. Programs meeting measured aspects of criterion 3B.3.1, by MA QRIS Level	90
Table 51. Programs meeting measured aspects of criterion 3B.3.2, by MA QRIS Level	90
Table 52. Programs meeting measured aspects of criterion 3B.4.1 by MA QRIS Level	91
Table 53. Programs meeting measured aspects of criterion 3B.4.3, by MA QRIS Level	91
Table 54. Programs meeting measured aspects of criterion 4A.2.1, by MA QRIS Level	92
Table 55. Programs meeting measured aspects of criterion 4A.2.2, by MA QRIS Level	92
Table 56. Programs meeting measured aspects of criterion 4A.2.3, by Q MA RIS Level	93
Table 57. Reason program meets language access component of criterion 4A.2.3, by MA QRIS Level	93
Table 58. Programs meeting measured aspects of criterion 4A.2.4, by MA QRIS Level	94
Table 59. Programs meeting measured aspects of criterion 4A.2.5, by MA QRIS Level	94
Table 60. Program participation in community event types, by MA QRIS Level	95
Table 61. Programs meeting measured aspects of criterion 4A.3.1, by MA QRIS Level	95
Table 62. Average number of modes of communication used by educators, by MA QRIS Level	95
Table 63. Programs meeting measured aspects of criterion 4A.3.2, by MA QRIS Level	96
Table 64. Programs meeting measured aspects of criterion 4A.3.3, by MA QRIS Level	97
Table 65. Reason program meets criterion 4A.3.3, by MA QRIS Level	97
Table 66. Programs meeting measured aspects of criterion 4A.3.4, by MA QRIS Level	97
Table 67. Reported ways of participating in early education and care organizations, by MA QRIS Level	98
Table 68. Programs meeting measured aspects of criterion 4A.3.5, by MA QRIS Level	98
Table 69. Programs meeting measured aspects of criterion 4A.4.1, by MA QRIS Level	99
Table 70. Programs meeting measured aspects of criterion 4A.4.2, by MA QRIS Level	100
Table 71. Number of listed services programs offered or connected families with, by MA QRIS Level	100
Table 72. Percentage of programs offering or connecting families to services, by MA QRIS Level	100
Table 73. Programs meeting measured aspects of criterion 4A.4.3, by MA QRIS Level	101
Table 74. Programs meeting measured aspects of criterion 5A.2.1, by MA QRIS Level	102
Table 75. Programs using communication strategies with staff at least quarterly, by MA QRIS Level	102
Table 76. Programs using communication strategies with families at least quarterly, by MA QRIS Level	103
Table 77. Programs meeting measured aspects of criterion 5A.2.2, by MA QRIS Level	103
Table 78. Programs meeting measured aspects of criterion 5A.2.3, by MA QRIS Level	104
Table 79. Programs meeting measured aspects of criterion 5A.2.5, by MA QRIS Level	104
Table 80. Reason program meets criterion 5A.2.5, by MA QRIS Level	104
Table 81. Programs meeting measured aspects of criterion 5A.2.5, by MA QRIS Level	105
Table 82. Reason program meets criterion 5A.2.5, by MA QRIS Level	105
Table 83. Programs meeting measured aspects of criterion 5A.3.1, by MA QRIS Level	105
Table 84. Programs meeting measured aspects of criterion 5A.3.2, by MA QRIS Level	106
Table 85. Programs meeting measured aspects of criterion 5A.3.3, by MA QRIS Level	106
Table 86. Programs meeting measured aspects of criterion 5A.3.4 by MA QRIS Level	107
Table 87. Programs meeting measured aspects of criterion 5A.3.5, by MA QRIS Level	107
Table 88. Programs meeting measured aspects of criterion 5A.4.1, by MA QRIS Level	108

Table 89. Programs meeting measured aspects of criterion 5A.4.2, by MA QRIS Level	109
Table 90. Programs meeting measured aspects of criterion 5A.4.3, by MA QRIS Level	109
Table 91. Programs meeting measured aspects of criterion 5B.2.1, by MA QRIS Level	110
Table 92. Programs meeting measured aspects of criterion 5B.2.2, by MA QRIS Level	111
Table 93. Programs meeting measured aspects of criterion 5B.3.1, by MA QRIS Level	111
Table 94. Communication methods used at least monthly, by MA QRIS Level	111
Table 95. Programs meeting measured aspects of criterion 5B.3.2, by MA QRIS Level	112
Table 96. Benefits offered to full and/or part time staff, by MA QRIS Level	112
Table 97. Programs meeting measured aspects of criterion 5B.3.3, by MA QRIS Level	113
Table 98. Programs meeting measured aspects of criterion 5B.3.4, by MA QRIS Level	113
Table 99. Career ladder supports, by MA QRIS Level	114
Table 100. Programs meeting measured aspects of criterion 5B.3.5, by MA QRIS Level	114
Table 101. Programs meeting measured aspects of criterion 5B.4.1, by MA QRIS Level	115
Table 102. Programs meeting measured aspects of criterion 5B.4.2, by MA QRIS Level	115
Table 103. Programs meeting measured aspects of criterion 5B.4.4, by MA QRIS Level	116
Table 104. Career ladder incentives in place at programs, by MA QRIS Level	116
Table 105. QRIS Criteria that do not empirically and significantly differentiate among MA QRIS Levels	118
Table 106. Criteria with inconsistent relationships within each quality standard	120
Table 107. ERS scores in relation to MA QRIS ERS criteria by MA QRIS Level	123
Table 108. Average ECERS-R Subscale ranges, means, and standard deviations	124
Table 109. Percent of Preschool Programs at Meeting ERS Benchmarks	125
Table 110. Average ITERS-R Subscale ranges, means, and standard deviations	126
Table 111. Percent of infant and toddler programs at meeting ERS benchmarks	127
Table 112. MA QRIS ERS cut-offs for each MA QRIS Level	129
Table 113. Percent of preschool programs at MA QRIS cut-off points	131
Table 114. Percent of infant and toddler programs at MA QRIS cut-off points	134
Table 115. Means and standard deviations for ERS subscales for preschool by MA QRIS Level	137
Table 116. ANOVA table of significant differences for preschool classrooms by MA QRIS Level	138
Table 117. Post hoc analyses of significant differences among levels for preschool classrooms	139
Table 118. Means and standard deviations on the ITERS-R by MA QRIS Level	140
Table 119. ANOVA table of significant differences for infant and toddler on the ITERS-R by MA QRIS Levels	141
Table 120. Post hoc analyses of significant differences among infant and toddler classrooms by MA QRIS Levels	142
Table 121. Items of strength on ECERS-R for preschool programs	143
Table 122. Items of challenge on ECERS-R for preschool programs	144
Table 123. ECERS-R challenge item indicator analysis	145
Table 124. Percent of programs with safety issues related to gross motor space	146
Table 125. Percent of preschool programs with issues related to care routines	146
Table 126. ITERS-R items of strengths for infant and toddler classrooms	148
Table 127. ITERS-R items of challenge for infant and toddler classrooms	149
Table 128. ITERS-R challenge indicator analyses	149
Table 129. Percent of programs not meeting specific routine care requirements	151
Table 130. Arnett CIS Scores for preschool classrooms, descriptive statistics	154
Table 131. Arnett CIS scores for infant and toddler classrooms, descriptive statistics	154
Table 132. Mean and standard deviation for Arnett CIS by QRIS level for preschool classrooms	156
Table 133. ANOVA table of significance differences by Arnett CIS Subscale for preschool classrooms	157

Table 134. Post hoc analysis of significant differences among levels for preschool classrooms	157
Table 135. Mean and standard deviation for Arnett CIS by QRIS level for infant and toddler Classrooms	159
Table 136. Percent of educator turnover by MA QRIS Level	161
Table 137. NAEYC descriptive statistics for preschool classrooms by MA QRIS Level	162
Table 138. NAEYC descriptive statistics for infant and toddler classrooms by MA QRIS Level	163
Table 139. Means for Level 3 preschool Head Start and Non-Head Start classrooms	164
Table 140. Level 4 ERS Descriptive Statistics for Preschool Classrooms by MA QRIS Level	165
Table 141. Preschool classroom quality by teacher education	166
Table 142. Infant/toddler classroom quality by teachers' education	167
Table 143. Classroom quality by education of the supervisor/administrator	169
Table 144. Quality among programs in which lead teacher had a degree (AA, BA or higher) in ECE or related field by MA QRIS Level	170
Table 145. Quality among programs in which the lead teachers has a BA or higher ECE (excluding related fields) by MA QRIS Level	170
Table 146. Teacher had a degree (AA or BA or higher than a BA) in ECE or related field	170
Table 147. Significant differences in classroom quality by MA QRIS Level for programs with a supervisor/administrator that had BA or higher in ECE or related field	170
Table 148. Correlations between preschool lead teachers' experience and quality measures	172
Table 149. Correlations among infant and toddler classroom quality and lead teachers' years of experience	173
Table 150. Correlations among quality and years of experience of administrator/supervisor	173
Table 151. Regression model predicting classroom quality by preschool lead teachers' experience and education	174
Table 152. Regression model for predicting classroom quality by infant and toddler lead teachers' experience and education	174
Table 153. Regression model for predicting preschool classroom Overall Average Item Score by administrators' experience and education	175
Table 154. Regression model for predicting infant and toddler classroom Overall Average Item Score by administrators' experience and education	175
Table 155. Regression model for predicting the combined overall average item score (for infant/toddler and preschool classrooms) within a program by administrators' experience and education	175
Table 156. Regression model for predicting classroom quality by preschool teachers' experience and education	176
Table 157. Child assessment measures and domains	178
Table 158. Descriptive characteristics of participating children by QRIS level	179
Table 159. Developmental pre/post net gain by measure	180
Table 160. Distribution of programs, by Re-Level using both QRIS criteria and ERS scores	184
Table 161. Shifts in centers' levels as a result of Re-Leveling including ERS criteria	185
Table 162. Most commonly unmet criteria for programs granted Level 2 that did not meet Level 2 after Re-Leveling, including ERS	185
Table 163. Most commonly unmet criteria for programs granted Level 3 that did not meet Level 3 after Re-Leveling, including ERS	186
Table 164. Most commonly unmet criteria for programs granted Level 4 that did not meet Level 4 after Re-Leveling, including ERS	186
Table 165. Programs' Re-Level within individual quality standards	187
Table 166. Programs' Re-Level within individual quality standards compared with EEC granted level	187

Table 167. Gains from pre-test to post-test on individual measures by Re-Level of program (Covariate-adjusted gain scores)	188
Table 168. Tests of differences among Levels in magnitude of change scores on individual measures (Wald test coefficient)	189

Figure 1. Status of QRIS in the United States	13
Figure 2. The levels of the Massachusetts QRIS	15
Figure 3. Programs in multi-site umbrella organization, by QRIS Level	22
Figure 4. Programs that are part of multi-service organizations, by QRIS Level	23
Figure 5. Head Start programs, by MA QRIS Level	23
Figure 6. Percent of non-profit programs, by QRIS Level	24
Figure 7. Program size, based on number of classrooms, by MA QRIS Level	25
Figure 8. Age groups served by programs, by QRIS Level	26
Figure 9. Children receiving subsidized tuition by MA QRIS Level	27
Figure 10. Programs located in MA high-needs communities by MA QRIS Level	28
Figure 11. Percent of programs meeting ECERS-R benchmarks	124
Figure 12. Percent of programs meeting ITERS-R benchmarks	127
Figure 13. Preschool programs meeting MA QRIS cut-offs regardless of assigned MA QRIS level	130
Figure 14. Infant and toddler programs meeting MA QRIS cut-offs regardless of assigned level	133
Figure 15. Overall Average item Score for ECERS-R by MA QRIS Levels	136
Figure 16. Average ECERS-R subscale scores by MA QRIS Levels	136
Figure 17. Infant and toddler classrooms ITERS-R overall average item score by MA QRIS level	140
Figure 18. Overall Arnett CIS score by MA QRIS level for preschool classrooms	155
Figure 19. Average Arnett CIS subscale scores by QRIS level for preschool classrooms	156
Figure 20. Average overall Arnett CIS score by MA QRIS Level for infant and toddler classrooms	158
Figure 21. Average Arnett CIS subscale scores by MA QRIS Level for infant and toddler classrooms	158
Figure 22. Comparing adjusted mean post-test scores for PPVT and WJ-III among QRIS Levels	181
Figure 23. Pre/post outcomes gains for preschoolers and differences in change scores among MA QRIS Levels	182
Figure 24. Pre/post outcomes gains for toddlers and differences in change scores between MA QRIS Level 1 and Level 2	183

The availability of high-quality early education and care in Massachusetts is of critical importance for everyone, including policymakers, parents, young children, and the public at large. As part of its ongoing commitment to ensure the highest quality of early education and care for all children in the Commonwealth, the Massachusetts Department of Early Education and Care (EEC) launched its Quality Rating and Improvement System (QRIS) in March 2011. The state's QRIS functions as an important tool for assessing, communicating, and, ultimately, improving the quality of early education and care providers across the state.

A federal Race to the Top – Early Learning Challenge (RTTT-ELC) grant award has significantly supported the statewide implementation of Massachusetts' QRIS. As part of the RTTT-ELC award, grantees are required to conduct a validation study of the state's QRIS. The Wellesley Centers for Women of Wellesley College and the University of Massachusetts Donahue Institute collaborated with MA EEC to conduct a validation study to evaluate the extent to which key components of the MA QRIS quality levels differentiate program quality; assess whether MA QRIS Levels are associated with differences in the developmental outcomes of children; and determine whether and how the system can be improved.

Quality rating and improvement systems

Quality Rating and Improvement Systems are systems by which to assess, communicate, and improve the quality of early education and care. A QRIS may take one of several forms, the three most common are building block systems, point systems, and hybrid systems. These approaches differ in terms of how ratings are determined. In a block system — the approach currently used by the Massachusetts QRIS — programs are required to meet a full set of standards or requirements at each level in order to advance to the next level. In a point system, programs accrue points after attaining an individual quality standard and are assigned a rating that corresponds to the total number of points earned. A hybrid system combines aspects of block and points systems, although the specific way in which these approaches are combined can vary.¹

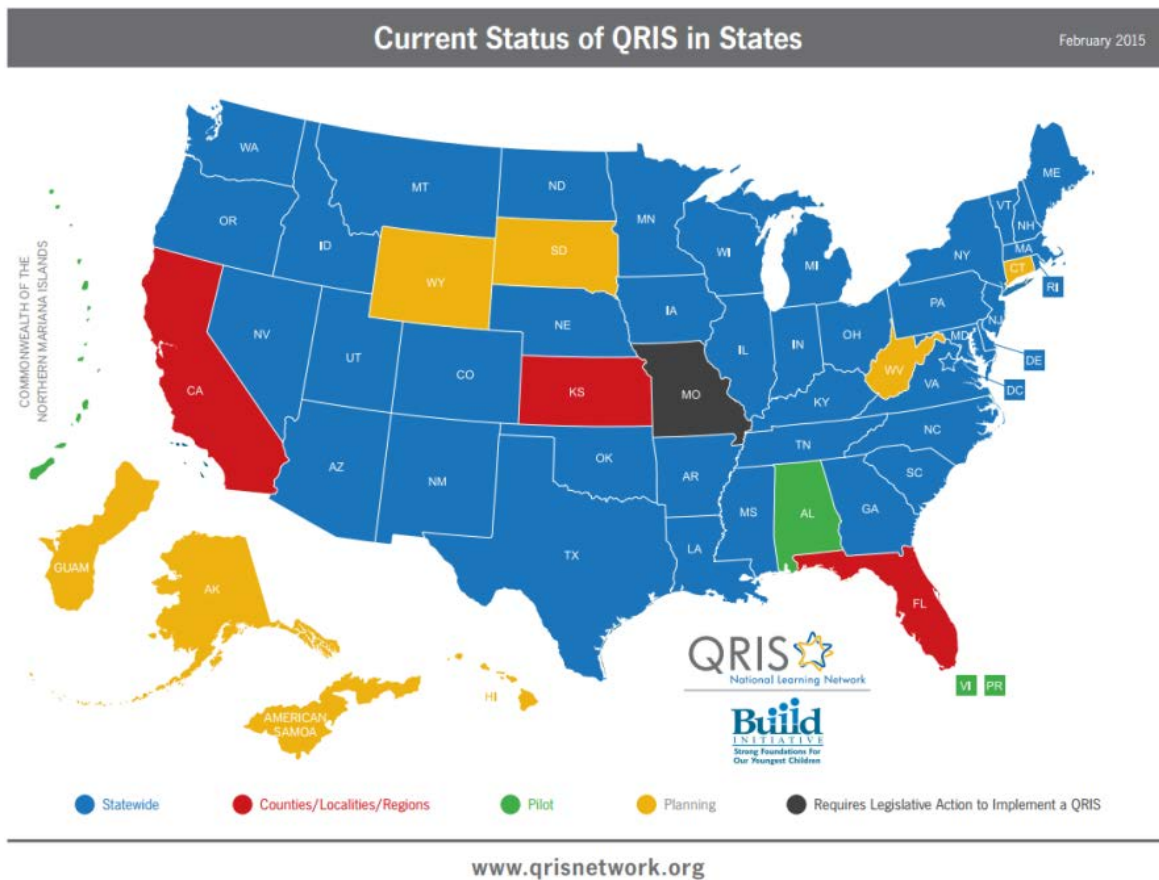
A QRIS collects information about programs in order to produce program-level ratings and typically have five components: (1) Standards that define program quality; (2) a rating system that measures and designates quality levels; (3) support for program improvement, such as technical assistance and/or training; (4) financial incentives for engagement or improvement; and (5) communication of quality ratings to caregivers and consumers.² Nationally, the development of QRIS has been accelerated by grant programs like the RTTT-ELC, and as of February 2015, 43 states across the country have a QRIS (three of which are organized by county, locality, or region), with more states currently developing systems.³

¹ Tout, K., Chien, N., Rothenberg, L. & Li, W. (2014). Implications of QRIS design for the distribution of program ratings and linkages between ratings and observed quality. OPRE Research Brief #2014-33. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

² Tout, Chien, Rothenberg & Li. (2014).

³ What is the pace of QRIS growth over the past 18 years? *QRIS Compendium*. Retrieved November 2015 from <http://qriscompendium.org/top-ten/question-1/>

Figure 1. Status of QRIS in the United States



Validation of a QRIS

Given that the central aspect of any QRIS involves defining and measuring quality, it is important to establish whether an individual system and its ratings, as implemented, are working to meaningfully distinguish the quality of early education and care programs. Unlike program evaluation and related research, validation studies look primarily at a QRIS’s efficacy as a measure of quality, focusing more on the process of — and criteria for — rating a program. Massachusetts is currently one of 24 states that have completed or are in the process of completing a validation study.⁴

As described by Zellman and Fiene, validation is a multistep process that does not necessarily provide a “yes” or “no” answer about QRIS, but instead provides information that can help support a QRIS and its ongoing refinement.⁵ The specific validation approach or approaches used often reflect the stage of a system’s development. Validation of programs in earlier stages of implementation often necessitates approaches that focus more directly on key underlying concepts and measurement strategies, while approaches related to assessing system outputs and examining associations between ratings and children’s development may be more appropriate to more mature systems.

⁴ How many QRIS have conducted or are in the process of conducting a QRIS evaluation? *QRIS Compendium*. Retrieved November 2015 from <http://qriscompendium.org/top-ten/question-8/>

⁵ Zellman, G. L. & Fiene, R. (2012). Validation of Quality Rating and Improvement Systems for early care and education and school-age care. Research-to-Policy, Research-to-Practice Brief OPRE 2012-29. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Conducting validation studies may provide several important benefits to states. Validation may, for example, lend credibility to the system, help direct quality improvement efforts to programs and practices most in-need of improvement, and support continuous improvement.⁶

Potential benefits of validation study

Lend credibility to the system by building trust in the ratings and increasing support for the system among parents, education providers, and other key stakeholders.

Focus quality improvement efforts on programs and practices most in need of improvement.

Support continuous improvement of a QRIS by providing additional information that could help revise the rating process and increase efficiency, such as removing unnecessary standards.

Standards and key quality components of the Massachusetts QRIS

The Massachusetts QRIS was developed for programs of all types serving children across the Commonwealth's mixed-delivery system. The system encompasses quality standards — customized for center-based providers, family child care providers, and afterschool and school-age providers — across five key domains of quality, three of which include subdomains (Table 1). In its current form, for center-based providers, which are the focus of this study, the Massachusetts QRIS encompasses a total of 79 quality standards across five domains. The standards have been in place since the system was implemented at scale in January 2011, although the way in which some standards are measured and verified has been modified with approval of EEC leadership.

Table 1. Key categories of quality in the Massachusetts QRIS

<p>1. Curriculum and Learning 1a. Curriculum, Assessment, and Diversity 1b. Teacher-Child Relationships and Interactions</p>
<p>2. Safe, Healthy Indoor and Outdoor Environments</p>
<p>3. Workforce Development and Professional Qualifications 3a. Designated Program Administrator Qualifications and Professional Development 3b. Program Staff Qualifications and Professional Development</p>
<p>4. Family and Community Engagement</p>
<p>5. Leadership, Management, and Administration 5a. Leadership, Management, and Administration 5b. Supervision</p>

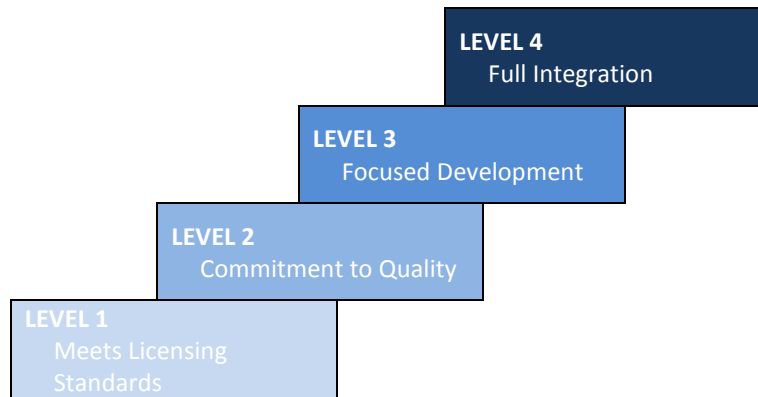
Programs participating in the Massachusetts QRIS voluntarily apply to receive a quality rating, or level, that may then be granted on the basis of EEC’s assessment of whether or not the required standards for a particular level have been met. These are conducted by Program Quality Specialists (PQS), which are assigned to each of Massachusetts’s five EEC Regions.

At Level 1 — the entry point to the system — programs must meet state licensing standards, and no other documentation or ratings are required. For Levels 2, 3, and 4, programs must complete required self-assessments; meet administrator and program staff qualifications; and provide documentation that the standards for that level have been met (Figure 2). Prior to entry to the system’s two higher levels, program classroom quality must be externally verified by either an EEC program quality specialist (Level 3) or a reliable rater (Level 4). As currently construed, programs at Level 2 are engaged in a process of self-assessment and

⁶ Zellman & Fiene. (2012).

continuous improvement, those at Level 3 show adherence to structural indicators and provide data demonstrating focused development, and those at Level 4 demonstrate full integration of the indicators of program quality.⁷

Figure 2. The levels of the Massachusetts QRIS



Source: Massachusetts Department of Education and Care

For the MA QRIS, all programs enter at Level 1, regardless of other credentials that a program might have. Because Massachusetts uses a building block system, programs must meet all criteria for a level before advancing to that or subsequent levels. However, programs may apply for and receive up to four exemptions. Barring exemptions, progression to Level 2 requires attaining 27 quality standards, including 3 standards related to Curriculum and Learning, 2 related to Safe and Healthy Indoor and Outdoor Environments, 10 related to Workforce Qualifications and Professional Development, 5 related to Family and Community Engagement, and 7 related to Leadership, Management, and Administration (Table 2). To achieve Level 3, a program must meet those 27 standards, plus an additional 30 standards, or a total of 57. To achieve Level 4, all of the system's 79 standards must be met. Entry into and progression with the system is voluntary, although programs serving children with subsidies, or participating in some specific initiatives or programs such as Massachusetts' UPK initiative, are required to enter and/or advance in the system.

⁷ Schilder, D., Young, J., Anastopoulos, L., Kimura, S. & Rivera, B. (2011). Massachusetts Quality Rating and Improvement System provisional standards study: Final report. *Massachusetts Department of Early Education and Care*. Retrieved from http://www.eec.state.ma.us/docs1/qris/20110307_ma_qris_provisional_study_final_rpt.pdf

Table 2. Number of QRIS standards, by quality domain

	Level 1	Level 2	Level 3	Level 4	Overall
1. Curriculum and Learning	n/a	3	6	3	12
1a. Curriculum, Assessment, and Diversity	n/a	2	4	2	8
1b. Teacher-Child Relationships and Interactions	n/a	1	2	1	4
2. Safe, Healthy Indoor and Outdoor Environments	n/a	2	3	1	6
3. Workforce Qualifications and PD	n/a	10	6	7	23
3a. Designated Program Administrator Qualifications and PD	n/a	6	4	4	14
3b. Program Staff Qualifications and PD	n/a	4	2	3	9
4. Family and Community Engagement	n/a	5	5	3	13
5. Leadership, Management, and Administration	n/a	7	10	8	25
5a. Leadership, Management, and Administration	n/a	5	5	4	14
5b. Supervision	n/a	2	5	4	11
All domains	n/a	27	30	22	79

Source: Analysis of Massachusetts Department of Early Education and Care Center and School-Based Standards, Revised November 16, 2012, available at <http://www.mass.gov/edu/docs/eec/qris/20121116-qris-center-based-standards.pdf>

Support for the MA QRIS and its key components of quality

The QRIS and its standards were developed through multiple iterations and with significant engagement with a variety of stakeholders, as well as a detailed assessment of the empirical support for proposed standards. The process, which began in 2008, involved the development of provisional standards, which were piloted with programs in 2010. For center- and school-based providers, the provisional standards comprised 149 proposed quality indicators, which were tested and reviewed as part of a study conducted on behalf of EEC by the Education Development Center (EDC). Following a rigorous review of the rationale for each standard, the 149 provisional standards were reduced to the 79 standards that comprise the current system.

The standards that were retained were those that were supported by the research, aligned with a research-based observational tool, able to be objectively documented or assessed, included in other states' QRIS, and/or articulated by stakeholders as reflecting best practice.⁸ Standards were also moved, reordered, or combined through this process. In total, 39 of the system's 79 standards were determined to have an empirical basis through a review of the literature (Table 3). This includes all or nearly all the standards within the subdomains of "Teacher-Child Relationships and Interactions" (for which 4 of 4 standards were determined to have an empirical basis) and "Supervision" (for which 10 of 11 standards were determined to have an empirical basis). Two subdomains, "Designated Program Administrator Qualifications and Professional Development" and "Leadership, Management, and Administration," were mostly comprised of standards aligned with either best practice or aligned with research-based observational tools.

⁸ Schilder, D., Canada, D., Paulk, S., & Smith Leavell, A. (2012) Review of research on QRIS standards. *Massachusetts Department of Early Education and Care*. Retrieved from <http://www.eec.state.ma.us/docs1/NewsUpdates/20121015-review-qris.pdf>

Table 3. Rationale for QRIS standards, by quality domain

	Number of standards that are...					
	Total #	Evidence-based	Measures-aligned	Best practice	Articulated by stakeholders	In other QRISs
1. Curriculum and Learning	12	7	2	4	—	1
1a. Curriculum, Assessment, and Diversity	8	3	2	3	—	1
1b. Teacher-Child Relationships and Interactions	4	4	—	1	—	—
2. Safe, Healthy Indoor and Outdoor Environment	6	3	4	—	—	—
3. Workforce Qualifications and PD	23	9	6	8	3	4
3a. Designated Program Administrator Qualifications and PD	14	4	5	8	—	1
3b. Program Staff Qualifications and PD	9	5	1	—	3	3
4. Family and Community Engagement	13	8	7	2	2	1
5. Leadership, Management, and Admin.	25	12	6	8	—	3
5a. Leadership, Management, and Admin.	14	2	6	8	—	3
5b. Supervision	11	10	—	—	—	—
All domains	79	39	25	22	5	9

Source: Analysis of Massachusetts Department of Early Education and Care Center and School-Based Standards, revised November 16, 2012, which is available at <http://www.mass.gov/edu/docs/eec/qr/20121116-qr-center-based-standards.pdf>. For each standard, the document lists the rationale for the standard, informed largely by Schilder, et al. (2011).

The MA QRIS validation study addresses a range of approaches and research questions, with a particular emphasis on assessing the outputs of the rating process and examining how ratings are associated with children's development. These are required components of validation studies conducted in states participating in the RTTT-ELC. Specifically, the study assesses whether components of quality and the quality levels can be relied on to accurately differentiate program quality, and the extent to which QRIS quality levels are associated with expected differences in children's development and growth.

Study sample

The study focused exclusively on center-based programs in Massachusetts and does not include family-based, school-based, or afterschool /out-of-school-time programs. This was done due to the population and distribution of programs within the MA QRIS that did not support the study of programs at Level 3 or above at the time of the study. The primary goal of the sample design was to provide reliable estimates that represent center-based programs in the Massachusetts QRIS. To meet the study objectives, researchers conducted a multistage sampling design stratified by MA QRIS Level.

Program sample

The study included 126 programs, selected from programs participating in the Massachusetts QRIS in summer 2014 (Table 4). The sample was stratified by MA QRIS Level to ensure sufficient power to detect differences in quality and child outcomes across MA QRIS Levels. Comparable numbers of programs were selected from Levels 1, 2, and 3.⁹ For Levels 1 and 2, programs were randomly selected and prescreened to identify and remove those that had active applications for the next QRIS level and were likely to transition within the next six months. Because there were relatively few programs at Level 3, all programs at that level were invited to participate in the study. A total of 128 programs were successfully recruited into the study, and all but two completed the study.¹⁰

Comparative analyses by level show that programs in the study did not differ significantly from the statewide population of programs in the MA QRIS in several measurable characteristics, including region, ages served, accreditation status, and percentage of children receiving subsidized tuition. The only significant within-level difference between study participants and non-participants was with regard to Head Start status, as few Head Start programs at Level 3 chose not to participate, so the sample for that level comprised more Head Starts at Level 3 than was typical of the QRIS (58 percent vs. 46 percent). Because the sample comprises roughly equal numbers of programs in each level, but the distribution of programs within the QRIS is heavily weighted towards the lower levels, sample participants are representative of the MA QRIS Levels, but not the MA QRIS as a whole without weighting. As such, caution should be exercised when making comparisons between the aggregate sample and the MA QRIS as a whole.

Despite efforts to prescreen programs to limit movement between levels during the course of the study, several programs did move — including six programs that moved from Level 3 to Level 4. For analytic purposes, programs were grouped using their level at the end of data collection (June 2015) in order to ensure that analyses properly represented quality at the time of the study. Due to the small sample size for Level 4 throughout this report, this level is presented in descriptive analyses, but is not included in statistical testing.

⁹ Programs in Level 1 and 2 were chosen in July 2015, and programs in Level 3 were chosen in August 2015 to provide EEC with time to complete a review of applications for Level 3 status. At the time the sample was selected, no programs had been rated at Level 4.

¹⁰ One program at Level 2 closed and another at Level 1 opted not to finish the study.

Table 4. Program sample

	Level 1	Level 2	Level 3	Level 4	Total
Number of programs per level: July 2014	41	44	43	—	128
Number of programs per level: June 2015	39	40	41	6	126

Classroom sample

In each selected program, up to two classrooms — one serving preschoolers and one serving either infants or toddlers — were randomly selected from lists of classrooms provided by the program. This sampling approach was used to examine the state's QRIS model of assigning a level to a program as a whole — that is, every classroom in the program should be in line with their assigned level's requirements. Because relatively few programs have infant rooms, any program with both infant and toddler rooms had an infant room randomly selected until the quota of infant classrooms was reached. After this, toddler rooms were selected to maximize the number of toddlers in the child outcomes samples (Table 5).

Table 5. Classroom sample by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4	Total
Number of preschool rooms	39	39	41	5	124
Number of toddler rooms	15	16	7	1	39
Number of infant rooms	12	15	8	—	35

Child sample

In preschool classrooms, trained assessors randomly selected six children for participation in direct child assessments and rating scales. Assessors selected three boys and three girls in attendance on the day of fall pre-assessments whose parents had not opted-out through the consent process. An oversample of two children per classroom was employed in the fall to account for attrition and absences at the spring assessment period, and to allow sample goals to be met. For spring post-assessments, four children from the original pool that were administered pre-assessment and who were still enrolled and in attendance were reassessed. If fewer than four children were present in the spring, all pre-assessed children were selected for the post-assessments. Complete pre-post direct child assessments were done with 481 preschoolers, while complete pre-post teacher ratings were done with 469 children.

In toddler classrooms, teachers completed rating scales on all participating toddlers in the classroom at pre-test. At post-test, classroom teachers completed rating scales on all toddlers selected at pre-test still enrolled in the program and who had not exceeded the age requirements for the toddler rating scale at post-test (36 months). Complete pre-post rating scales were done with 190 toddlers (Table 6).

Table 6. Child sample: Pre-assessment and post-assessment

	Pre-Assessment	Post-Assessment
Direct preschool child assessments	737	481
Preschool social emotional teacher-rating scales	726	469
Toddler social emotional teacher-rating scales	294	190

Data collection methods

A holistic approach to data collection was undertaken in the study of the MA QRIS. This strategy included the collection of data on all criteria of the MA QRIS through surveys and interviews with a program administrator, classroom observations, and direct child assessments.

Program-level data collected

Program-level data were collected directly by the research team during the winter, between child assessment periods. Director interviews were conducted over the phone, while program information questionnaires, staff rosters, and supervisor supplements were paper forms that were delivered and returned via mail. The data collection instruments are described in full below.

- **Director interview:** Structured interviews captured key program information focused on measures of quality that extend from program characteristics and policies, including curriculum, screening, and assessment practices; services for children and families; staff development practices; and program self-assessment and self-improvement.
- **Program information questionnaire:** Programs completed questionnaires with basic information about their program that observations cannot capture, including educator professional development, staff compensation policies and turnover, communication with families and staff, business practices, and participation in community and professional organizations.
- **Staff roster:** Programs completed a staff roster with information about early education experience and education qualifications for each of the educators at their program. Example data collected included experience in early education and care settings, teacher education, and teacher certification.
- **Supervisor supplement:** This brief supplement was completed by the person primarily responsible for supervising staff at each study program. Data collected related to the experience and credentials of the supervisor and included education, experience, tenure, professional development, and training.

Classroom-level data collected

Classroom observations were conducted by reliable raters using the Environment Rating Scales (ERS,) at one preschool classroom and one infant or toddler classroom (if applicable) at each study program. Classroom observations took place during winter between child assessment periods. To reduce any potential bias, observers were blind to the MA QRIS Level ratings of the programs and did not participate in other data collection aspects of the study. Observations lasted approximately 3–4 hours and were scheduled on a typical morning for the program. All indicators on the ECERS-R and ITERS-R were rated as “yes” or “no,” regardless of the final score on each item, to allow for analyses of specific indicators. During the classroom observations, ERS supplemental teacher interviews and Arnett Caregiver Interaction Scales were also completed.

- **The Early Childhood Environment Rating Scale – Revised (ECERS-R)¹¹:** This scale includes 43 items organized into seven subscales: Space and Furnishings, Personal Care Routines, Language-Reasoning, Activities, Interactions, Program Structure, and Parents and Staff.
- **The Infant/Toddler Environment Rating Scale – Revised (ITERS-R)¹²:** This scale includes 39 items organized into seven subscales: Space and Furnishings, Personal Care Routines, Listening and Talking, Activities, Interaction, Program Structure, and Parents and Staff.
- **ERS supplemental teacher interview:** As part of the ECERS-R and ITERS-R protocol, teachers were interviewed to complete the scoring of subscale items and to capture additional data. Example data collected included teacher professional development, classroom curriculum and assessment practices, and classroom-level parent involvement.
- **Arnett Caregiver Interaction Scale¹³:** This measure of caregiver involvement and interaction style with children was completed for two educators within each observed classroom. The scale contains 26 items with four subscales: Sensitivity, Harshness, Detachment, and Permissiveness.

Child-level data collected

The direct child assessment took between 25–40 minutes per child and was administered one on one in the child’s program, in a quiet area with a familiar adult present. Direct child assessors were well-trained and experienced assessors who achieved reliability standards during training. Preschool teachers were asked to complete rating scales following the assessments while toddler teachers received rating scales in the mail.

- **Direct Child Assessments:** Assessment instruments included:
 - Peabody Picture Vocabulary Test, 4th edition¹⁴ – Assesses children’s receptive vocabulary
 - Woodcock-Johnson Tests of Achievement, 3rd edition: Letter Word Identification¹⁵ – Assesses children’s recognition of letters and words
 - Woodcock-Johnson Tests of Achievement, 3rd edition: Applied Problems¹⁶ – Assesses children’s understanding of early mathematical concepts
- **Teacher Rating Scales:** Assessment instruments included:
 - The Devereux Early Childhood Assessment for Preschoolers, 2nd edition¹⁷ – Assesses social-emotional development
 - The Preschool Learning Behavior Scale¹⁸ – Assesses behaviors related to learning
 - The Devereux Early Childhood Assessment for Toddlers¹⁹ – Assesses social-emotional development
- **Child roster:** Programs completed child roster forms for all children directly participating in the study. Data collected included gender, ethnicity, limited English proficiency / English language learner, primary language spoken at home, how tuition is funded, whether child receives early intervention or special education services and type, average hours per week of program attendance, and whether child regularly attends another program.

11 Harms T, Clifford RM, Cryer D. (1998) Early Childhood Environment Rating Scale, Revised Edition.; New York: Teachers College Press. .
12 Harms, T., Cryer, D., & Clifford, R. M. (2003). Infant/toddler environment rating scale. Revised. Edition. New York: Teachers College Press.

13 Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology*, 10, 541- 522.

14 Dunn,L,M. & Dunn, D. M. (2007). Peabody Picture Vocabulary Test (4th Edition). Minneapolis: MN: NCS Pearson.

15 Woodcock, R., McGrew, K. S., & Mather, N. (2001). *The woodcock-Johnson III*. Itasca, IL: Riverside Publishing.

16 Woodcock, R., McGrew, K. S., & Mather, N. (2001).

17 LeBuffe, P.A. & Naglieri, J.A. (2012). Devereux early childhood assessment for preschoolers, second edition. Lewisville, NC: Kaplan Early Learning Company

18 McDermott, P. A., Green, L. F., Francis, J. M., & Stott, D. H. (2000). Preschool Learning Behaviors Scale. Philadelphia: Edumetric and Clinical Science.

19 Mackrain, M., LeBuffe, P., & Powell, G. (2007). Devereux early childhood assessment for infants and toddlers. Lewisville, NC: Kaplan Early Learning Company

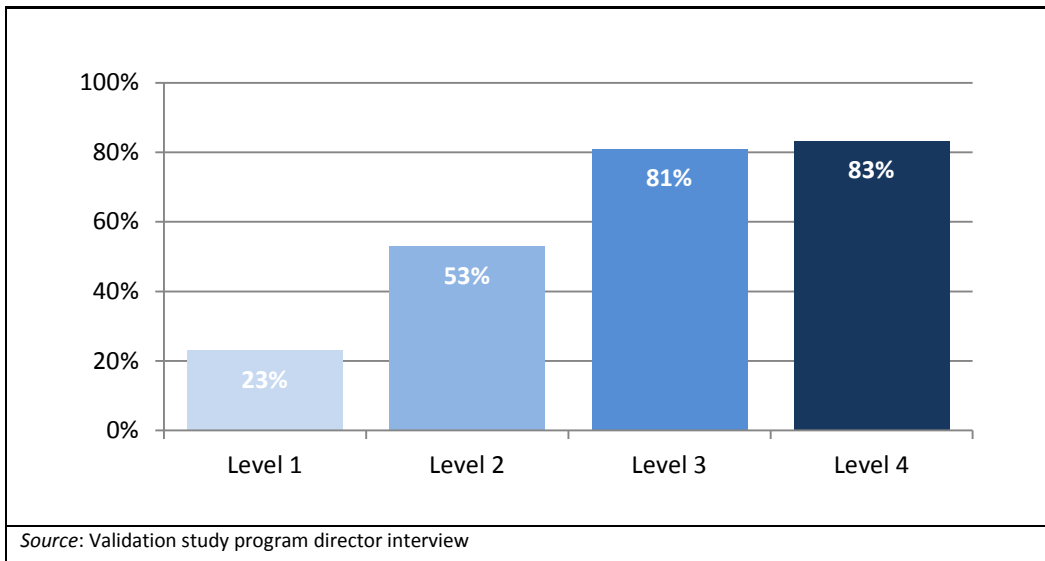
Using the sample, the validation study looked at the representation of programs across MA QRIS Levels by a number of key program characteristics, including organizational affiliation, ages served, size, and whether or not the program served children with subsidized enrollment. These analyses are intended to describe the study sample and also identify program types that may be over- or underrepresented at particular levels. Because the sample is designed to be representative of each MA QRIS Level more broadly, these findings may be useful in highlighting program types that may benefit from targeted support or outreach.

In considering potential implications of these findings, it should be noted that the distribution of program types across levels is likely to be influenced by a number of factors related to system implementation, including different groups’ level engagement with the system; external requirements, such as those tied to grants or contracts, or other incentives to participate or progress in the system; and EEC outreach strategies in the early phases of QRIS implementation. This is particularly true given the nature of the MA QRIS as a voluntary system in which programs choose to engage with the system as they pursue and demonstrate attainment of standards for higher levels.

Organizational affiliations or type and MA QRIS Levels

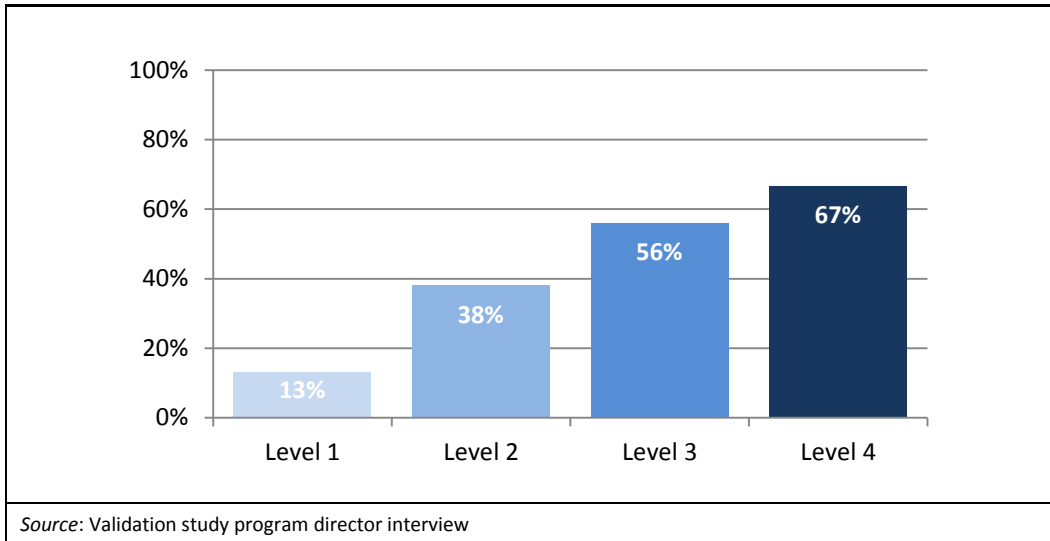
Within the validation study sample, programs at higher levels were much more likely to be part of larger organizations, such as umbrella organizations that operate multiple sites or multi-service organizations which provide services in a broad domain of social and human service areas in addition to early education and care. For example, in Levels 3 and 4, the vast majority of study programs (81 and 83 percent, respectively) were affiliated with an umbrella organization operating multiple early education and care sites (Figure 3). By comparison, just over half of Level 2 programs and about a quarter of Level 1 programs were part of a multi-site umbrella organization.

Figure 3. Programs in multi-site umbrella organization, by MA QRIS Level



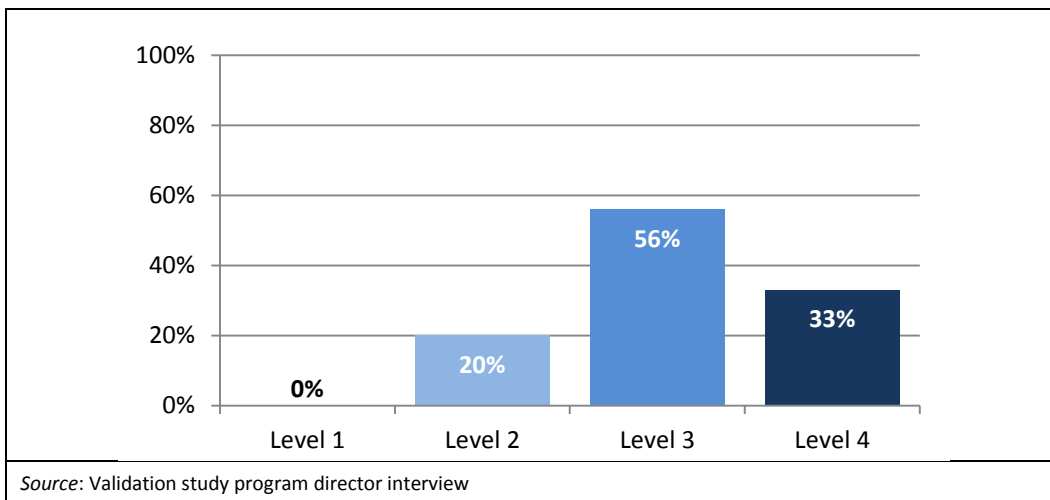
Similarly, 56 percent of study programs in Level 3 were part of multi-service organization, as were 67 percent of those in Level 4. In the lower tiers, fewer programs (13 percent at Level 1 and 38 percent at Level 2) operated as part of such an agency.

Figure 4. Programs that are part of multi-service organizations, by MA QRIS Level



There were also notable differences in the representation of Head Start programs across levels (Figure 5). More than half of study programs at Level 3 were Head Start (56 percent), compared to 20 percent at Level 2. There were no Head Start programs in the Level 1 sample (0 percent).

Figure 5. Head Start programs, by MA QRIS Level

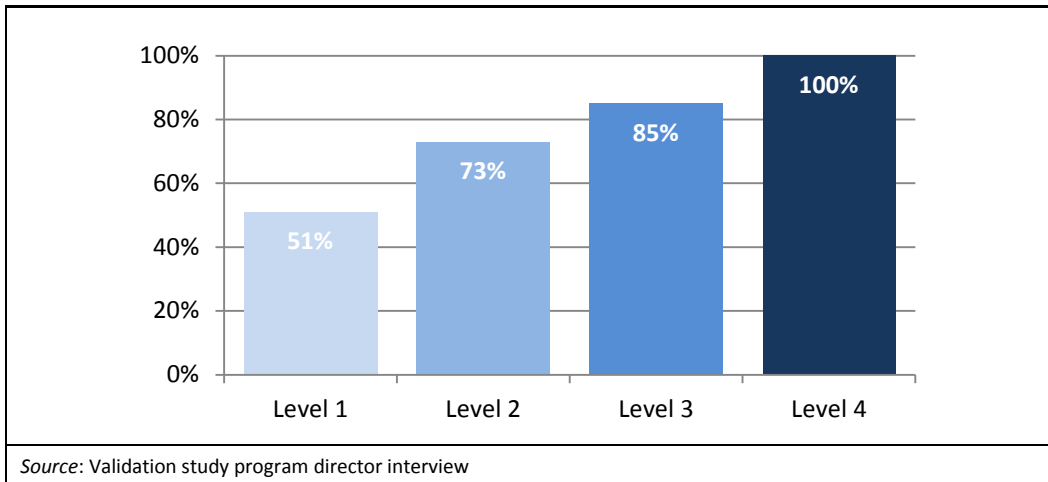


These findings are not surprising, as it was noted that during early implementation of the MA QRIS, EEC conducted outreach to umbrella organizations, Head Starts, and other agencies, which in turn played a substantial role in encouraging and supporting their programs’ participation in and progression through the system. Additionally, Head Starts were encouraged to participate in the MA QRIS by other Head Start support networks outside of MA EEC. Further, in regional focus groups conducted prior to the start of the validation study, participants noted that larger organizations may have more administrative capacity to understand the requirements and complete the self-assessment than independent providers, which may also have contributed to programs connected to larger organizations being overrepresented in the higher levels within the sample. One potential implication of these findings is that, as the MA QRIS matures, small, independent community providers not affiliated with larger organizations may require intentional outreach to promote their full engagement with the system and encourage progression through the levels. Additionally, findings support

continual outreach to other groups or support networks as a means of promoting participation and movement within the MA QRIS.

There were also differences in the proportion of programs operating as non-profit or public agencies at higher levels. At Level 1, approximately half of study programs (51 percent) were non-profit or public agencies, with the remainder of programs (49 percent) organized as for-profit programs. At Level 4, all study programs were public or non-profit (Figure 6). This suggests that outreach may be needed to encourage for-profit programs to engage further with MA QRIS and for MA EEC to emphasize that the system is relevant to all providers in MA.

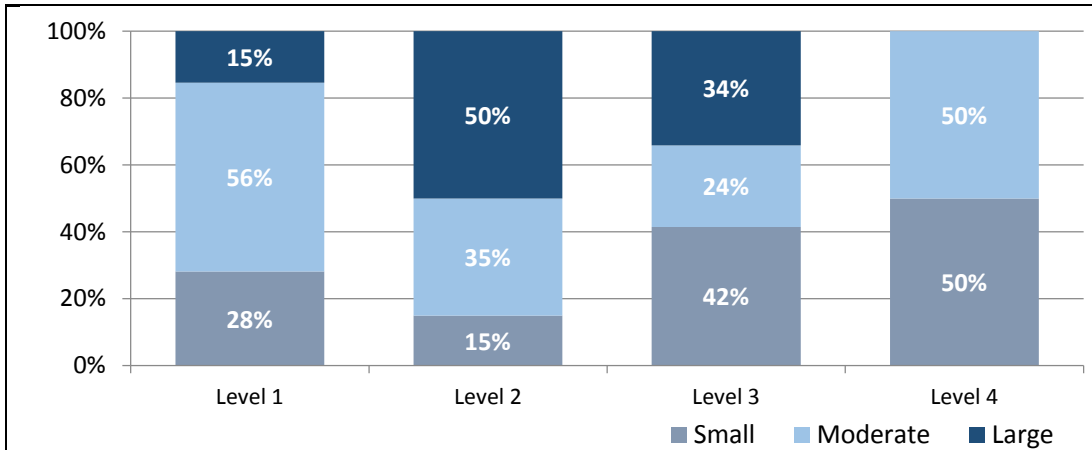
Figure 6. Percent of non-profit programs, by MA QRIS Level



Program size and MA QRIS Levels

The validation study also looked at the distribution across levels of programs of various sizes, as measured by number of classrooms and overall enrollment. With respect to number of classrooms, centers with two or fewer classrooms were coded as small, those with three to five classrooms were coded as being of moderate size, and those with six or more classrooms were considered large. Data shows differences in the distributions by program size across QRIS levels. From Levels 2 through 4, there appears to be a decreasing percentage of small programs and an increasing percentage of large programs at progressively higher levels. Level 1, on the other hand, was primarily comprised of moderately-sized programs (56 percent).

Figure 7. Program size, based on number of classrooms, by MA QRIS Level



There were also differences in program size, as measured by overall enrollment, across levels within the sample. As shown in Table 7, programs at Levels 2 and 3 tended to be larger, on average, than those in either Levels 1 or 2, and comprised a broader range of enrollment sizes. Although, in each level, the smallest programs tended to be similarly sized, with enrollment sizes between 12 (Level 2) and 16 (Level 4), in Levels 2 and 3, the largest programs were more than twice as large as those at Level 1. Programs at Level 4, on the other hand, tended to be more similar in size, ranging from 16 children to 76 children.

Table 7. Average number and range of children enrolled by level

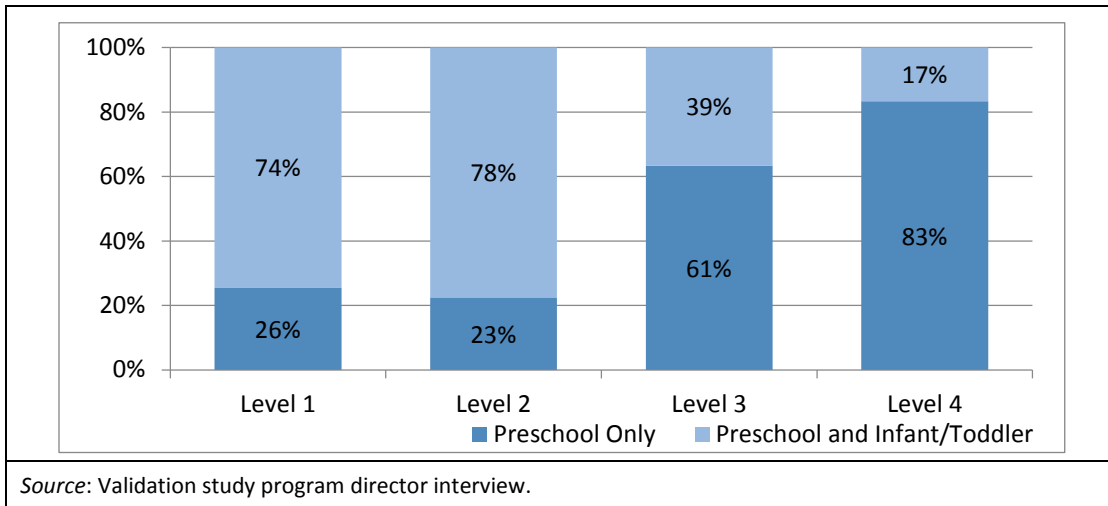
	N	Mean	Range
Level 1	39	53	14–122
Level 2	40	82	12–289
Level 3	41	79	14–344
Level 4	6	42	16–76

Source: Validation study program information questionnaire

Ages of children served and MA QRIS Level

Overall, there were higher proportions of programs serving infants and toddlers in the system’s lower tiers (Levels 1 and 2) than in the upper tiers. At Level 3, approximately 39 percent of programs served infants and/or toddlers, and at Level 4, only 17 percent of programs offered programming for infants and/or toddlers. By contrast, at Levels 1 and 2, 74 percent and 78 percent of programs, respectively, served infant and toddlers. This difference persisted even when Head Start programs are removed. Among non-Head Start programs in the sample, 74 percent of Level 1 and 91 percent of Level 2 programs served infants and toddlers compared with 56 percent at Level 3 and 25 percent at Level 4 (Figure 8).

Figure 8. Age groups served by programs, by MA QRIS Level

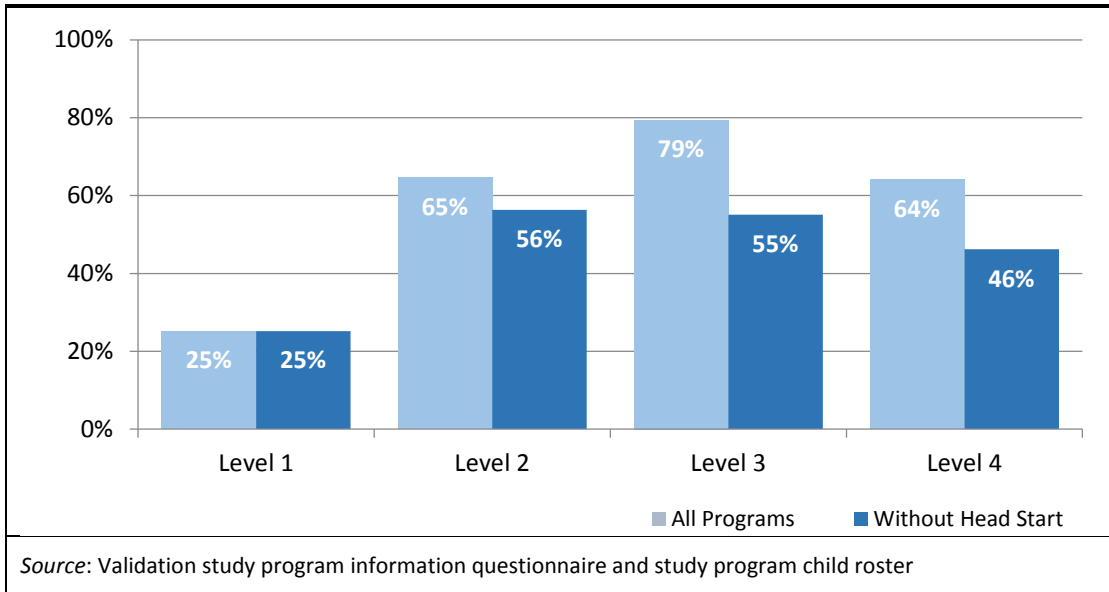


It should be noted that in general, about 60 percent of programs in the MA QRIS serve both infant/toddlers and preschool aged children. That being said, these analyses do suggest that programs serving infants and toddlers have not been as quick to progress to the system’s higher levels and, thus, may represent a group that may benefit from targeted outreach and/or quality support in the future.

Participation in subsidy initiatives and MA QRIS Levels

Finally, there were notable differences in the proportion of programs serving children receiving a subsidy and MA QRIS Level. Approximately one-quarter of Level 1 programs had enrolled children receiving some form of subsidy, such as vouchers, contract slots, or Head Start funding, whereas at Levels 2, 3, and 4, the proportion of programs with subsidized enrollment ranged from 64 percent (Level 4) to 79 percent (Level 3). This difference between Level 1 and the other levels persists even with Head Start programs removed, and to some extent, likely reflects EEC requirements that programs participating in certain funding initiatives join and, in some cases, progress in QRIS in order to remain eligible for that funding. In addition, EEC has worked intensively with programs participating in its Universal Pre-K (UPK) initiative to support their progress to Level 3(Figure 9).

Figure 9. Children receiving subsidized tuition by MA QRIS Level



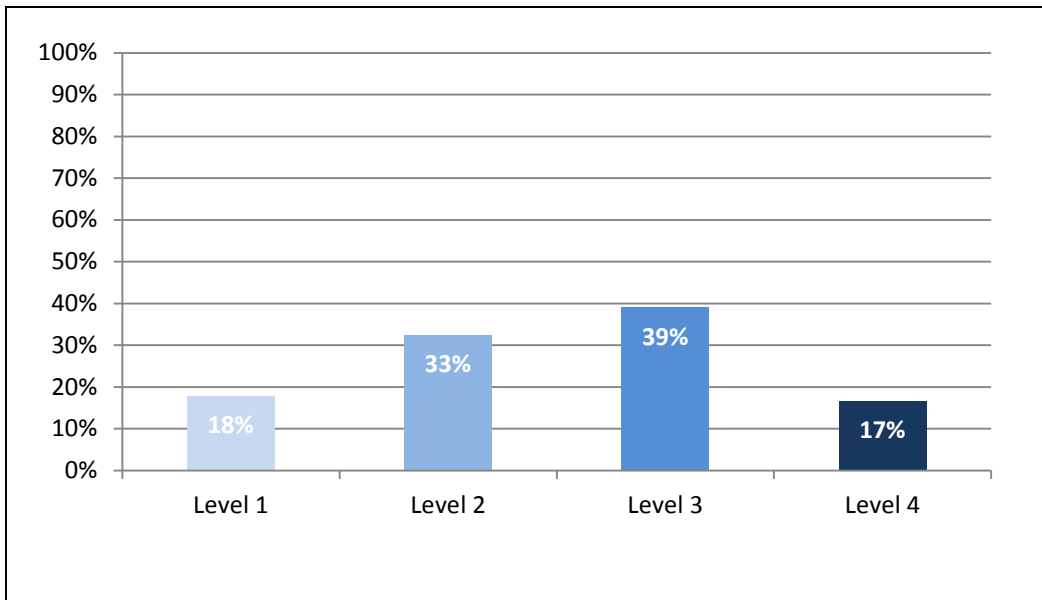
High-needs communities and the MA QRIS

The validation study also explored whether the MA QRIS was reaching high-needs communities. Overall, this data suggests that center-based early education programs are located in diverse communities throughout the state, with programs in the higher MA QRIS Levels more clustered in high-need and linguistically diverse communities (Figure 10).

Communities were classified using results of a community needs assessment completed by the MA Department of Public Health in 2010. This assessment identified priority communities in need of enhanced supports for maternal health and child development; these communities are participating in EEC’s home visiting initiative. Providers within these high-risk communities serve economically disadvantaged families from within 17 communities in Massachusetts.²⁰ Analysis of these data indicates that a larger percentage of programs in higher levels are located in these identified high-need communities. For example, 39 percent of Level 3 and 33 percent of Level 2 programs are located in these communities, in contrast to 18 percent of Level 1 programs.

²⁰ These communities are Holyoke, Springfield, Chelsea, Lawrence, Lowell, New Bedford, Fall River, Lynn, Southbridge, Worcester, Brockton, Boston, Pittsfield, Revere, Everett, North Adams, and Fitchburg. Once a classification system was set up, the QPM ‘city’ field had to be cleaned to match the community names in the MAPC classification. Since addresses are entered by individual providers, 212 communities in the QPM (or 19 percent) were neighborhoods or other community names not identified in the MAPC. To guide changes to the city field in the QPM, entries were searched using Google as well as conducting zip code matching with the Zip Code Database (<http://www.unitedstateszipcodes.org/zip-code-database>).

Figure 10. Programs located in MA high-needs communities by MA QRIS Level



Summary and implications

As mentioned previously, because the validation study sample was designed, through random selection of programs from each level, to be representative of programs in that level more broadly, the over- or under-representation of programs by some key characteristics may be useful in identifying programs that may benefit from additional outreach and support as MA QRIS continues to be implemented. Specifically, independent community providers and those serving infants and toddlers may need additional outreach or support from MA EEC to fully engage in the system and/or progress to its upper tiers. In addition, programs that are not as dependent on subsidies and for-profit organizations reflect other populations of programs for which new implementation strategies may be beneficial to promote movement to higher levels.

The findings of this study are divided into eight sections as follows:

Section I: Descriptive analyses of MA QRIS Standards, Criteria, and Verification

Section II: Statistical analyses of MA QRIS Standards and Criteria

Section III: Program Quality and the Massachusetts QRIS

Section IV: Classroom Quality and Key Program Attributes

Section V: Education, experience and the Massachusetts QRIS

Section VI: Child Outcomes and the Massachusetts QRIS

Section VII: MA QRIS Program Re-Leveling

Section VIII: Perceptions of the MA QRIS among Programs and Providers

Collectively, data and the subsequent findings come from several sources including program-level (director interview, program information questionnaire, staff roster, and supervisor supplement), classroom-level (ECERS-R, ITERS-R, Arnett Caregiver Interaction Scale, and ERS supplemental teacher interview), and child-level (direct child assessments, teacher rating scales).

Programs were grouped using their MA QRIS Level at the end of data collection (June 2015) in order to ensure that analyses properly represented quality at the time of the study. Throughout the report, Level 4 is presented in descriptive analyses when appropriate, but not included in statistical testing due to its small sample size. Further, all data regarding infant/toddler classroom rely exclusively on Levels 1, 2, and 3, as just one Level 4 study program had an infant/toddler classroom.

When comparing data between groups, it is important to determine whether observed differences are statistically significant or are possibly the result of random variation and chance. Statistical tests provide a value for that purpose, called a p-value. For all validation study data analysis, the term “statistically significant” is used to connote findings with an associated p-value of .05 or less. This is a level chosen in many social science studies and provides 95 percent confidence that observed differences are not due to random change. In other words, if a statistical test demonstrates that the differences in responses between subgroups are significant — in this case by QRIS level — then, in the simplest terms, such differences are very unlikely to have occurred simply due to chance. The report also notes when analyses “approached significance” which is associated with a p-value of between .10-.06. Given the small sample sizes in the study and for some analyses, findings approaching significance should be given some consideration as they represent a trend of practical importance in the data.

Section I: Descriptive analyses of MA QRIS Standards, Criteria, and Verification

The Massachusetts QRIS for center-based programs comprises eight quality standards defined by 79 criteria by which those standards are verified. Programs enter the system at Level 1, which is equivalent to licensing, and progress to higher levels through the MA EEC verification process that determines whether the criteria for a specific level are met. Verification requirements differ by criteria and include review of program-provided documentation and direct observation of program and classroom practices. Programs are allowed to apply for exemptions for up to two criteria per level, which is intended to make the system more flexible and take into account certain programming constraints.

To gain insight into how the MA QRIS standards and individual criteria are operating in practice, the validation study gathered extensive data through interviews, surveys, and observations to independently assess the extent to which programs at each MA QRIS Level exhibited practices and characteristics consistent with the criteria.²¹ Data collected from these sources were analyzed in two ways: first, by holistically examining not only the individual criteria but also the verification procedures to identify ambiguities, repetitions and inconsistencies on both the criterion and verification sides of the MA QRIS process (this section); and second, by statistically exploring how standards are related to each other (see Section II: Statistical analyses of MA QRIS Standards and Criteria).²² Verification standards are outlined in EEC's *Program Quality Specialist Verification Guidance Document for Center-Based/School Based Programs*, which identifies how each individual criterion is assessed. This guidance is intended to help standardize the rating process, promote inter-rater reliability among program quality specialists, and ensure consistency of ratings.

An important outcome of this work is to provide MA EEC with information and analyses that can support the refinement of the MA QRIS system and the individual criteria that comprise it in order to reduce redundancy and increase the clarity and efficiency of the system. To this end, the analyses presented in this section offer information regarding each individual MA QRIS criterion and identify those that may warrant further review and modification, either with respect to the criterion itself or its verification procedures. Where appropriate, analyses also identify criteria that do not vary across levels and thus, do not appear to distinguish quality among levels, as well as those that may be serving as barriers to progression.

It should be noted that since its implementation, the Massachusetts QRIS has undergone significant changes to its verification procedures and requirements. As such, in some cases, the percentage of programs, at a specific MA QRIS Level, not meeting requirements might reflect differences in verification at the time the program was granted a MA QRIS Level as opposed to issues with the verification processes. In some cases, use of exemptions and changes in exemptions procedures as the system has been implemented may also explain differences in the proportion of programs at each MA QRIS Level meeting a criterion. Also, researchers collected data primarily through interviews and surveys and did not require documentation from programs regarding individual criteria. This was a result of not wanting to burden participating programs with administrative work and data collection constraints. As such, with the exception of criteria referencing classroom quality and/or requirements on the ECERS-R and ITERS-R which was collected by trained and reliable study observers, data presented below are primarily based on self-report, most commonly by the director of the program.

Overarching considerations

Table 10 below presents the individual criteria for each MA QRIS Level and recommendations for criteria-specific revisions. Additionally through review of the data, criteria and verification processes, nine overarching recommendations for revision were identified: (1) better alignment of MA QRIS criteria and verification

²¹ ECERS-R and ITERS-R observations were scored such that data were collected for all quality indicators contained within the instrument, not just those up to the point that a score was able to be determined, allowing for the assessment of individual practices. In addition, some additional observational indicators were recorded during observations to assess specific practices related to some QRIS standards.

²² QRIS study instruments were developed prior to the most recent verification guidance, and as a result, there are instances in which the validation study was able to measure some but not all aspects of a particular criterion.

standards contained within a single MA QRIS document outlining both the criteria and verification process; (2) clarification that CEUs or credits are needed for MA QRIS professional development (3) timeframes for trainings to reinforce best practices; (4) reduction of memorandums of understanding requirements which appear to serve as barriers to progress for programs; (5) limitation of the use of overall scale scores for verification and increased focus on specific subscale/items that are more relevant; (6) reduction of compound criteria; (7) incorporation of Continuous Quality Improvement Plans as a self-assessment criteria; (8) consolidation of PD (professional development) requirements into the Professional Development Standard; and (9) greater consistency between MA QRIS requirements and other credentialing systems and initiatives. The recommendations are described below in detail.

(1) Better alignment of QRIS criteria and verification and distribution of a single document outlining both criteria and verification

The *Program Quality Specialist Verification Guidance Document for Center-Based/School Based Programs* conveys important information about how MA QRIS criteria are interpreted and assessed. At present, this exists as a separate document from the MA QRIS criteria and is used primarily by the Program Quality Specialists for verification purposes. While it is important to have detailed reference materials for verification, requiring programs to cross-reference two documents — the QRIS standards and the verification guidance — in order to fully understand what is a requirement may contribute to perceptions of complexity in the system, particularly in cases where the guidance modifies or substantively interprets an otherwise ambiguous criterion. As such, an important overarching recommendation involves better aligning language in some criteria with the specific verification requirements or potentially creating one document in which the criteria and verification procedures are presented in conjunction with one another. A reference document could be made available with greater detail to serve as an additional resource, if needed.

(2) Clarification that CEUs or credits are needed for MA QRIS trainings

As referenced throughout the verification guidance, MA QRIS requires that teachers receive continuing education units (CEUs) or college credit for professional development and training. Reliable data regarding CEUs could not be collected through the study as most teachers and administrators were unable to identify whether CEUs for professional development were received. Specifically, when asked about whether educators had received CEUs for MA QRIS-required trainings, depending on the training, between 54 percent and 66 percent of interviewed teachers indicated that they did not know whether they had received CEUs for the professional development. Similarly, administrators, even at higher levels, often indicated that they were not aware if their teachers had received CEUs for MA QRIS trainings.

These findings suggest a lack of knowledge about CEU and credit requirements in the field, and, in turn, point to a potential need for MA QRIS to build awareness for the importance of formal CEU- or credit-bearing coursework if that remains an important aspect of the system. At a minimum, EEC may want to consider making explicit reference to CEU or credit requirements in the MA QRIS criteria to reduce ambiguity and emphasize the need for CEU-approved training. Another potential strategy would be to assign individual educators levels outside of their program based on their receipt of MA QRIS-aligned training and/or other qualifications. For example, educators could be certified as a “MA QRIS Level 2” or a “Level 3 teacher.” This approach parallels the MA Licensing approach such that teachers are certified as either a “teacher” or “lead teacher.” This could potentially promote further educator engagement in MA QRIS as well as institute a career ladder whereby educators can build their credentials through participation in professional development and attainment of degrees or other qualifications. Additionally, MA EEC may want to consider expanding the MA QRIS professional development requirements to include other types of formal training beyond CEU professional development.

(3) Timeframes for the trainings to reinforce best practices

Another potential overarching consideration relates to the timeframes in which trainings for the MA QRIS need to be completed. Currently, the MA QRIS does not place time parameters on PD, so PD taken at any time were considered as meeting MA QRIS requirements for the study. Data indicated that for each of the trainings, more

than 60 percent of educators in observed classrooms indicated that the training had been taken within the last one to two years. For a sizeable number, however, the training was taken more than three years and sometimes more than five years prior to the study. For example, about 22 percent of educators said they had received training in the Massachusetts Guidelines for Preschool Learning Experiences within the last three to five years, and an additional 16 percent reported that the training was done more than five years ago (Table 8). We recommend that MA EEC establish timeframes for some key QRIS PD to ensure that teachers are current with changes to a curriculum, assessment standards, and the Massachusetts Guidelines.

Table 8. Timeframe in which most recent training was completed

	Number trained	% within last 1-2 yrs	% within last 3-5 yrs	% more than 5 yrs
The curriculum the program uses	150	77%	16%	7%
How to do classroom observations/self-assessments	170	69%	23%	8%
Assessment for all children, including those with and without disabilities and special needs	161	74%	20%	6%
Documenting student progress	137	67%	25%	8%
Working with children from diverse languages and cultures	132	64%	27%	10%
The MA Guidelines for Preschool Learning Experiences	112	62%	22%	16%
The MA Early Learning Guidelines for Infants & Toddlers	86	69%	20%	12%
Special diets, allergies, and specialized feeding issues	164	89%	9%	2%

(4) Reduction of memorandums of understanding as barriers to progresses for programs

For several MA QRIS criteria, programs are expected to provide services or connect families to services. When these services are not provided directly by program staff, QRIS verification typically requires that programs have a memorandum of understanding (MOU), or a formal collaboration agreement with the service providers. Many programs reported having ongoing relationships with multiple agencies and service providers, though relatively few programs indicated that they had formal MOUs in place (Table 9). In interviews, administrators often noted challenges in getting partner agencies to commit to signing formal agreements, even though they were working collaboratively with agencies in the provision of specified services. This was particularly true with regard to programs working with early intervention, the public schools, or large social service agencies. This represented a substantial frustration for many programs, particularly given the emphasis on these agreements within the MA QRIS.

Table 9. Programs delivering services through collaboration and MOU status

	Works with an agency	Has an MOU with that agency	
		#	%
Family literacy for families	41	15	37%
Specialists and consultants to work with educators	61	17	28%
Parenting education for families	31	7	23%
Child development for families	24	5	21%
Adult education for families	47	9	19%
English as a Second Language for families	41	6	15%
Job training for families	39	5	13%

Notably, programs at higher levels — which data indicated were more likely to be part of larger agencies — were often able to offer these services through in-house resources or through umbrella organizations, such that they were able to meet QRIS requirements without needing an MOU. Providers at lower levels, on the other hand, were more likely to be dependent on collaboration and subsequently more affected by the need for these formalized agreements. To better support the field, EEC may want to consider outreach to promote the importance of MOUs among agencies that typically partner with programs in service delivery or consider relaxing requirements regarding proof of collaboration as opposed to necessitating a formal MOU.

(5) Limitation of the use of overall scales scores for verification and increased focus on specific subscale/items that are more relevant

Throughout the MA QRIS, a number of criteria reference specific practices or policies that are verified by using programs’ overall scores on measurement tools, including the Environment Rating Scales (ERS), the Program Administration Scale (PAS), and the Classroom Assessment Scoring System (CLASS) or Arnett Caregiver Interaction Scale (Arnett CIS). As such, multiple criteria have the same verification process, indicating that criteria themselves do not uniquely assess diverse aspects of quality. Additionally because overall scale scores take into account multiple aspects of quality to produce an aggregate score, it is possible for programs to meet verification standards but not have in place the specific practices or policies referenced in the criteria being verified, even if the tool addresses the practice at some level. Alternatively, a program could have the specific policy and practice in place but fail to meet the criterion due to it not meeting dimensions of quality unrelated to that criterion which result in a reduced overall score. As a result, direct measure of these criteria — for example, through the use of specific items/subscales in the measurement tools or by requiring documentation of practice — may be desirable. This could also increase the focus of programs on key practices of quality promoted through the MA QRIS system.

(6) Reduction of compound criteria

There are also several instances in which MA QRIS criteria contain multiple requirements. Typically, programs were meeting some but not all of these requirements, and not surprisingly, there appears to be a negative relationship between the number of requirements incorporated into a criterion and the percentage of programs meeting that criterion, typically with requirements in the middle and end of compound criteria having the lowest percentages of compliance. Inclusion of too many requirements in an individual criterion may be confusing to programs, as it may not be clear whether all aspects of the criterion are required or, if only some are required. There also appears to be some redundancy of requirements across compound criteria, further underscoring the need to streamline criteria. Where possible, it may be beneficial to refine criteria to address the individual components in order to clarify a requirement, thereby ensuring that all aspects of the criterion are equally addressed and reduce redundancy.

(7) Incorporation of Continuous Quality Improvement Plans as a self-assessment criteria

Verification guidance for a number of criteria incorporate programs' Continuous Quality Improvement Plans (CQIPs) CQIPs are created by programs and verified by the Program Quality Specialist as part of programs progressing to a higher Level of the MA QRIS. As such, CQIPs function as *criteria* (e.g., something a program creates, submits and is verified) versus functioning as a verification process. To more clearly emphasize the need for programs to have a CQIP and the use of the tool as an customized quality planning and building guide, MA EEC may want to consider incorporating the CQIP as a MA QRIS self-assessment criteria, specifically requiring that programs have a plan in place to address areas of needed focus and strength for each of the MA QRIS Standards of Quality. Resources related to using the CQIPs as a self-reflective tool for planning and program development may be needed to facilitate the use of the CQIP such as: an outline of QRIS requirements, training and technical assistance on the CQIP and self-reflective practice, and templates for an Individual Professional Development Plans (IPDPs) that parallel MA QRIS requirements at each level.

(8) Consolidation of PD requirement into the Professional Development Quality Standard

Although the MA QRIS has a quality standard defined as *Program Staff Qualifications and Professional Development*, PD requirements for educators are incorporated within multiple MA QRIS Quality Standards, outside of the *Program Staff Qualifications and Professional Development* standard. This makes the PD requirements for each Level difficult to track and creates redundancy within the system. As such, researchers recommend that the MA QRIS place all PD requirements in *the Program Staff Qualifications and Professional Development Standard* for quality to facilitate the tracking of PD and reduce redundancy.

(9) Greater consistency between MA QRIS requirements and other credentialing systems and initiatives

Review of the criteria indicated overlap between MA QRIS requirements and other MA EEC credentialing systems, EEC licensing, and EEC initiatives, but not complete alignment. As a result, programs and educators must work with two sets of requirements. This was particularly true for administrator requirements in which the MA QRIS system parallels MA Licensing credentials for Director 1 and Director 2 but does not offer complete consistency. Subsequently, administrator must work with 2 sets of PD requirements to both promote the career ladder and advancement in the MA QRIS. Another example is the requirement for parent input for child assessments. As part of the requirement for parental input, the MA QRIS outlines timeframes for child assessments. These time frames are consistent with MA licensing for preschoolers and infants but not for toddlers. Not surprisingly, a large number of programs did not meet the requirement for toddler but did meet the requirements for other ages. As such, researchers suggest that the MA QRIS work to create greater consistency between the MA QRIS, MA licensing, MA credentialing, and another initiatives to reduce burden, clarify requirements, and increase efficiency.

Considerations for individual standards and criteria

In the sections that follow, the MA QRIS quality standards are discussed in detail, with data from the Validation Study presented regarding the percentage of programs meeting each criterion by level (Table 10). This provides a high-level overview of each criterion including the level at which the criterion resides, the criterion itself, EEC verification guidance for assessment of the criterion; the proportion of programs at each level that met the criterion as measured by validation study data; and considerations derived from these analyses. For convenience, the table has been color-coded. Green represents criteria that appear to be working well, yellow represents criteria in which changes to the verification process might be warranted, orange represents criteria for which modification of the criteria themselves may be warranted, red represents criteria that could be removed from the QRIS or may require significant revisions, and blue is for recommendations for criteria to be added to the MA QRIS system. Grey represents criteria in which reliable data was not obtained by the validation study. Again, it is important to emphasize that the MA QRIS criteria, verification standards, and exemption process for the MA QRIS has changed considerably since the system's inception. Additionally, programs are permitted two exemptions per level, so programs might not consistently be meeting standards for

their assigned level. Regardless, the data provide important insight into the functioning of individual criteria and areas for potential revision.

Table 10: MA QRIS quality standards, criteria and verification process; percent meeting criteria and considerations for refinement

MA Quality Standard 1. Curriculum and Learning							
Sub-Scale 1A. Curriculum, Assessment, and Diversity							
	QRIS Criterion	PQS Verification Guidance	Estimated percent meeting				Comments and considerations
			L1	L2	L3	L4	
Level 2	1A.2.1. Educators demonstrate completion of formal professional development in curriculum, screening tools, and formative assessment.	<p>Check PQR to ensure that at least one educator in each classroom has received formal PD in:</p> <ul style="list-style-type: none"> • curriculum • screening tools • formative assessment <p>If a program has four classrooms, the PQR should indicate that at least four educators have been trained in all three of these content areas.</p>	56%	73%	83%	67%	<ul style="list-style-type: none"> • Consider revising criterion to incorporate language regarding CEUs and credits. Data indicate that many administrators and teachers are not readily aware of training that teachers have received. • Consider breaking apart criterion into two: (1) training in curriculum and (2) training in screening tools and assessments, to clarify that all are needed • Clarify criterion to indicate that one teacher per classroom needs to meet training requirements. • Consider time frames for training
Level 2	1A.2.2. Materials reflect the language and culture of the children in the classroom, their communities, and represent the diversity of society.	<ul style="list-style-type: none"> • Verify that ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 3, and subscale score requirements of 2 or 3 depending on the subscale • Verify that the Program has submitted a Program Improvement Plan 	36%	68%	68%	83%	<ul style="list-style-type: none"> • Revise verification procedure to require a self-assessed score of 3 on the ERS Promoting Acceptance of Diversity item (not overall score) OR • Consider moving to Level 3 in which a direct visit by PQS can verify presence of materials
Level 2	1A.2.3 Program has a Continuous Quality Improvement Plan addressing goals of the program related to QRIS	<ul style="list-style-type: none"> • Verify aspects of the Program Improvement Plan 					<ul style="list-style-type: none"> • Consider adding requirement to MA QRIS, as opposed to utilizing as verification tool to stress its importance and utilization
Level 2	1A.2.4 Administrator has taken a training in conducting the ERS, for each scale relevant to their program	<ul style="list-style-type: none"> • Verify ERS training by an ERSI certified trainer 					<ul style="list-style-type: none"> • Consider adding requirement to MA QRIS, as a means of promoting a more accurate self-assessment process at Level 2 and facilitating the movement of programs (see ERS quality section)

	QRIS Criterion	PQS Verification Guidance	Estimated percent meeting				Comments and considerations
			L1	L2	L3	L 4	
Level 3	1A.3.1. Staff include parental input in the progress reports.	<ul style="list-style-type: none"> • Example of a completed progress report. • Must demonstrate the program shares progress reports with: <ul style="list-style-type: none"> ○ parents of preschool children at least <u>3 times in a 12-month period</u>, AND ○ parents of infants/toddlers or children with disabilities at least <u>4 times in a 12 month period</u>. • Description of the process (i.e. policy or procedure) for sharing the report with parents • Report must have an area for parents to comment (share observations, goal setting, etc.) <p>* Note: PQS Guidance also allows for verification of 1A.3.1 for NAEYC and Head Start programs.</p>	5%	25%	39%	33%	<ul style="list-style-type: none"> • Consider revision of criterion to include verification requirements for frequency. Data suggest that programs were not consistently aware of the frequency requirements for progress reports, particularly for toddlers which seem to be primarily on a preschool schedule, which is the frequency scheduled employed by <i>MA Licensing</i> (see data below in criterion description). • Clarify criterion regarding parental input so that the requirement for "parental input" is clear. (i.e. <i>Progress reports also include an area for parent comments to share ideas, goals and observations</i>).
Level 3	1A.3.2. Staff has received formal professional development in the curriculum; using the MA Guidelines for Preschool Learning Standards or Infant /Toddler Learning; documenting children's progress; and working with children from diverse languages and cultures and second language acquisition.	<ul style="list-style-type: none"> • Verify FORMAL professional development is listed in the PQR (college credit, min. .5 CEU, or .5 CEC) for all • Formal training documented in PQR must indicate all of the following have been completed by lead teachers: <ul style="list-style-type: none"> ○ Curriculum Development ○ Using the MA Guidelines for Preschool Learning Experiences and Infant/Toddler Learning Guidelines (I/T only applicable if this age group is served) ○ Documenting Children's Progress ○ Working with children from diverse cultures ○ Working with early English language learners 	21%	35%	59%	67%	<ul style="list-style-type: none"> • Consider revision of criterion to indicate that CEUs are required. • Consider separating into multiple criteria so that it is clear what training is needed (see data below in criterion description). • Given that the Guidelines reflect the foundations of quality and ERS scores suggest that a significant number of Level 1 and 2 programs are in the Inadequate/Minimal range (see ERS quality section), consider moving requirement to Level 2. • Requirement for curriculum training for this criterion may be redundant with the curriculum training for criterion 1A.2.1 above. • Consider adding time frames for training.

	QRIS Criterion	PQS Verification Guidance	Estimated percent meeting				Comments and considerations
			L1	L2	L3	L 4	
Level 3	1A.3.3. Program uses screening tools, progress reports, formative assessments, and information gathered through observation to set goals for individual children across all developmental domains.	<ul style="list-style-type: none"> Verify that the program completes developmental screening on all children Verify that the program completes formative assessment on all children Description of how these tools are used to share information with staff and families, and create activities to support all children’s individual needs. Verify that the program uses an evidence-based screening tool (see appendix) Verify that the program uses an evidence-based formative assessment tool (see appendix) <p>Note: If the tool is not approved, verify that the tool being used screens and assesses all developmental domains; social/emotional, cognitive, physical health and development, language and literacy.</p>	23%	65%	95%	100%	<ul style="list-style-type: none"> Consider incorporating language into criterion to stress that the use of evidence-based tools, as opposed to program-developed assessments, are needed for this criterion. At lower levels of the QRIS, programs were less likely to conduct screenings or use an EEC approved tool for formative assessments. High usage of program-developed tools and the cost of published assessments may serve as a barrier to programs; targeted support may be necessary to help programs in lower levels access the tools and receive training in their use.
Level 3	1A.3.4. Staff demonstrate language and literacy skills either in English or the child’s language that provide a model for children.	<p>Verify that:</p> <ul style="list-style-type: none"> ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 4.5, and subscale score requirements of 3 or 4 depending on the subscale Program has submitted a Program Improvement Plan Provide T.A. on ERS Subscale 3 after observations in 60% of preschool classrooms and 100% of infant and toddler classrooms. 	51%	68%	88%	100%*	<ul style="list-style-type: none"> Consider verification through the Language and Reasoning or Language and Talking Subscales only as opposed to overall score to more directly assess language environments for children. <p>* Data for Level 4 reflect preschool classrooms only, as only one Level 4 program had an infant/toddler classroom.</p>

	QRIS Criterion	PQS Verification Guidance	Estimated percent meeting				Comments and considerations
			L1	L2	L3	L 4	
Level 4	1A.4.1. Program uses a curriculum that is aligned with MA guidelines for Preschool Learning Standards and the Infant/ Toddler Learning Guidelines.	<p>Verify that the program uses a curriculum aligned with core competencies outlined in MA Guidelines; approved curriculums include:</p> <ul style="list-style-type: none"> • High Scope • Creative Curriculum • OWL • Resources for Early Learning <p>Note: If the curriculum is not approved, program must provide an example of the curriculum. Cross-reference the curriculum with the Standards and Guidelines.</p>	72% (28%)	90% (73%)	98% (80%)	100% (83%)	<ul style="list-style-type: none"> • A portion of programs were using their own curriculum which they believe meet guidelines. Due to the constraints of the study, we were unable to verify if the curriculum actually meets the guidelines standards. The numbers in parenthesis show the percent of programs that report using a field curriculum that we can verify as meeting EEC guidelines.
Level 4	1A.4.2. Program uses progress reports, appropriate screening tools, formative assessments, and information gathered through observation to inform curriculum planning, use results to monitor each child's progress across developmental domains, and inform program decision-making (e.g. curriculum content, strategies for improved staff implementation, and professional development).	Documentation that curriculum is adapted to monitor each child's progress across developmental domains, and inform program decision-making.	41%	78%	90%	100%	<ul style="list-style-type: none"> • This criterion and its verification procedures appear to be working as intended. No edits are recommended.

MA Quality Standard 1. Curriculum and Learning							
Sub-Scale 1B. Teacher-Child Relationship and Interactions							
	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	1B.2.1. All staff receive orientation and ongoing professional development and supervision in how to support positive relationships and interactions through positive, warm, and nurturing interactions.	Formal professional development documented in the Registry. <ul style="list-style-type: none"> • <i>Educators</i>: must have Formal professional development. • <i>Staff who do not work in the classroom</i> (ex. drivers, cooks, maintenance workers): may complete orientation or professional development hours only. 	56%	95%	100%	100%	<ul style="list-style-type: none"> • This criterion and its verification procedures appear to be operating as expected. • There may be opportunities to refine language with respect to references to interactions, such that the standard would read “All staff receive an orientation and ongoing professional development and supervision in how to support positive relationships and warm and nurturing interactions.”
Level 2	1B.2.2 Program has a Continuous Quality Improvement Plan addressing goals of the program related to QRIS	<ul style="list-style-type: none"> • Verify aspects of the Program Improvement Plan 					<ul style="list-style-type: none"> • Consider adding as a requirement to the MA QRIS, as opposed to utilizing as verification tool to stress its importance and utilization.
Level 3	1B.3.1. Staff engage children in meaningful conversations, use open-ended questions and provide opportunities throughout the day to scaffold their language to support the development of more complex receptive and expressive language, support children's use of language to share ideas, problem solve, and have positive peer interactions.	<ul style="list-style-type: none"> • Verify that ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 4.5, and subscale score requirements of 3 or 4 depending on the subscale • Provide T.A. on ERS Subscale 3 after observations in 60% of preschool classrooms and 100% of infant and toddler classrooms. • Verify that the Program has submitted a Program Improvement Plan AND <ul style="list-style-type: none"> • Verify that CLASS self-assessed score of 3 or higher on the dimensions of Positive Climate and Teacher Sensitivity, and a score of 3 or lower on the dimension of Negative climate • OR Arnett Caregiver Interaction Scale self-assessed score of 3.0 or higher for all educators 	51%	68%	88%	100%	<ul style="list-style-type: none"> • This criterion addresses a program's language environments, which is also addressed at Level 3 in the curriculum, assessment, and diversity standard (1A.3.4) using a similar verification procedure (ERS scores). • Data reflect the proportion of programs receiving a score of 4 or above on the Language and Reasoning (ECERS-R) and Listening and Talking (ITERS-R), subscales. • Consider combining this criterion with criterion 1A.3.4 and/or using Language and Reasoning (ECERS-R) and Listening and Talking (ITERS-R) scores as opposed to overall scores and all subscales for more direct verification of the criterion itself.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	1B.3.2. Educators are provided with opportunities to use outside consultants or staff with expertise in the age of the children served to assist them in implementing strategies that support positive relationships/interactions and prevention/intervention techniques.	<ul style="list-style-type: none"> • Verify that program offers educators support from internal staff or consultant with certification in at least one of the following: <ul style="list-style-type: none"> ○ behavioral health ○ mental health ○ prevention/intervention techniques (i.e. CSEFEL) • Review policy and verify that educators have access to this support <i>as needed</i>. <p>If this is offered via an outside contractor, verify the MOU or provide proof that collaboration is in place.</p> <p>* Note: PQS Guidance also allows for verification of this criterion for NAEYC and Head Start programs.</p>	74% (21%)	83% (50%)	95% (80%)	83% (50%)	<ul style="list-style-type: none"> • Most programs had specialists on staff, used consultants to assist educators, or worked with an agency that provided these services. • When formal collaboration agreements (MOUs) are factored in, the proportion of programs meeting this criterion decreased substantially (see the numbers in parentheses). • Administrators often face challenges in getting partner organizations to sign formal agreements, even though they were actively collaborating. If formal agreements remain an important part of QRIS, EEC may need to provide outreach and support to promote MOUs or allow program to demonstrate active collaboration as an alternative to formal agreements. • It is also unclear what is meant by “certification” as part of the verification process: it may be beneficial to provide a list of acceptable certifications for clarification.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 4	1B.4.1. Staff utilize teaching strategies that ensure a positive classroom environment, engage children in learning, and promote critical thinking skills.	<p>Verify:</p> <ul style="list-style-type: none"> CLASS reliable rater score of 6 or 7 on the dimensions of Positive Climate and Teacher Sensitivity, and a score of 2 or 1 on the dimension of Negative climate OR Arnett Caregiver Interaction Scale average score of 3.2 or higher <p>AND</p> <ul style="list-style-type: none"> ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 5.5, and subscale score requirements of 4 or 5 depending on the subscale ERS Reliable Rater scores for each classroom meet or exceed the overall score requirement of 5.5, and subscale score requirements of 4 or 5 depending on the subscale Program has submitted a Program Improvement Plan. 	23%	23%	46%	60%*	<ul style="list-style-type: none"> Consider using Interaction subscale scores as opposed to overall scores for more direct verification of the criterion itself. Data reflects the proportion of programs receiving a score of 5 or above on the Interaction subscale of the ECERS-R and/or ITERS-R, depending on ages served. The relatively low proportion of study classrooms receiving scores at this level, even at higher levels, suggests this might be an area programs may need support to improve or maintain quality (see ERS and Quality section). If CLASS or Arnett are to be included in verification, a process will need to be developed for reliable rating to occur, as this does not currently exist. <p>* Data for Level 4 reflect preschool classrooms only, as only one Level 4 program had an infant/toddler classroom</p>

MA Quality Standard 2. Safe, Healthy Indoor and Outdoor Environments							
	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	2A.2.1. Annual consultation by a Health Consultant to monitor records, update health care policies and practices, identify program issues, assist programs in complying with health and safety requirements, and provide a written report to the program, unless needs of a child require additional consultation.	<p>Program must submit two forms:</p> <ul style="list-style-type: none"> A <i>signed</i> form from the Health Consultant or copy of the collaboration agreement through which a health consultant provides services; and A <i>signed</i> statement from the program director verifying that the visits occur on at least an annual basis. <p>Copy must have been signed within 2 years, or program must submit an updated copy.</p> <p>Health Consultant must meet EEC Licensing definition of Health Care Consultant.</p> <p>* Note: PQS Guidance also allows for verification of 1A.3.1 for NAEYC and Head Start programs.</p>	64%	85%	95%	100%	<ul style="list-style-type: none"> Overall, this criterion and its verification procedures appear to be working largely as intended. Data reflect the proportion of programs working with a health consultant, not a MA QRIS health consultant specifically.
Level 2	2A.2.2. Demonstrates healthy, safe, and clean indoor and outdoor environments.	<ul style="list-style-type: none"> Verify that ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 3, and subscale score requirements of 2 or 3 depending on the subscale. Verify that the program has submitted a Program Improvement Plan. 	87%	73%	85%	100%*	<ul style="list-style-type: none"> Data reflect the proportion of programs receiving a score of 2 or above on the Personal Care Routines subscale of the ECERS-R and/or ITERS-R, depending on ages served — a score that falls between the Inadequate and Minimal benchmarks for the ERS (see ERS Quality Section for more detail). <p>* Data for Level 4 are for preschool classrooms only, as only one Level 4 program had an infant/toddler classroom</p>
Level 2	2A.2.3 Program has a Continuous Quality Improvement Plan addressing goals of the program related to QRIS.	<ul style="list-style-type: none"> Verify aspects of the Program Improvement Plan 					<ul style="list-style-type: none"> Consider adding as a requirement to the MA QRIS, as opposed to utilizing as verification tool to stress its importance and utilization

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	2A.3.1. Program, with parental consent, provides (directly or through collaboration) vision, hearing, and dental screenings, and shares results with families.	<ul style="list-style-type: none"> Program provides an example document that tracks that all children have screenings for: <ul style="list-style-type: none"> o vision o hearing o dental Verify that the program shares the results of the screenings with the families. <p>If this service(s) is offered via an outside contractor, verify the MOU or provide proof that a collaboration is in place.</p>	23%	38%	90%	83%	<ul style="list-style-type: none"> Data represent the proportion of programs tracking children’s receipt of these services. Fewer programs provided these services, either directly or through collaboration, often because families typically accessed services on their own. Refinement of the criterion is recommended to clarify that tracking is expected even when families access their own services and clarification is needed for expectations of programs when families do not receive these services directly (i.e. are programs expected to provide these services, refer families, etc...).
Level 3	2A.3.2. Staff are trained in how to work with children with special diets, allergies, and specialized feeding issues.	<ul style="list-style-type: none"> Verify FORMAL professional development is listed in the PQR (college credit, min. .5 CEU, or .5 CEC) for all educators (lead teachers, teachers, assistant teachers and directors). Formal training documented in PQR must indicate the following has been completed: <ul style="list-style-type: none"> o special diets, allergies, and specialized feeding issues. 	41%	65%	76%	100%	<ul style="list-style-type: none"> Consider clarifying that all staff are required to have this training in criterion. Other criteria related to staff training typically require training for one educator per classroom, and in most cases where Level 3 programs were not meeting the requirement; it was because some but not all staff had been trained. Consider revising criterion to incorporate language regarding CEUs and credits. Data indicate that many administrators and teacher are not readily aware of training that teachers have received.
Level 3	2A.3.3. Demonstrates healthy, safe, and clean indoor and outdoor environments.	<ul style="list-style-type: none"> Verify that ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 4.5, and subscale score requirements of 3 or 4 depending on the subscale. Provide T.A. on ERS Subscale 2 after observations in 60% of preschool classrooms and 100% of infant and toddler classrooms. Verify that the program has submitted a Program Improvement Plan. 	21%	18%	46%	60%*	<ul style="list-style-type: none"> Less than half of Level 3 programs met the Level 3 score requirement for the Personal Care Routines subscale, which assessed health practices. Similar to above appears to be an area in which programs appear to need support to maintain and improve practice (see ERS quality sections). <p>* Data for Level 4 are for preschool classrooms only, as only one Level 4 program had an infant/toddler classroom</p>

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 4	2A.4.1. Demonstrates stimulating indoor and outdoor environments that provide access to sinks in the classroom.	Verify that: <ul style="list-style-type: none"> ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 5.5, and subscale score requirements of 4 or 5 depending on the subscale ERS Reliable Rater scores for each classroom meet or exceed the overall score requirement of 5.5, and subscale score requirements of 4 or 5 depending on the subscale Program has submitted a Program Improvement Plan Sinks in all classrooms. 	3%	3%	15%	0%*	<ul style="list-style-type: none"> Few programs met the Level 4 score requirement for the Personal Care Routines subscale (see ERS quality section). * Data for Level 4 are for preschool classrooms only, as only one Level 4 program had an infant/toddler classroom

MA Quality Standard 3. Workforce Qualifications and Professional Development							
Standard 3A. Designated Program Administrator Qualifications and Professional Development							
	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	3A.2.1. Program administrator has a Child Development Associate (CDA) Credential for the age of the children served (or higher, i.e. Associate's or Bachelor's degree), OR is enrolled in a program leading to an Associate or Bachelor's Degree in early childhood education or related field.	<p>Verify program administrator has a CDA, Associate, Bachelor or Master Degree listed in the PQR.</p> <p>A program administrator who has not competed a CDA, Associate, Bachelor or Master Degree must:</p> <ul style="list-style-type: none"> • have some college course work entered in the PQR that has been completed within the last six months. <p>Note: If the program administrator does not have any current coursework entered in the PRQ, the program must submit evidence that the PA is enrolled in an Associate, Bachelor, or Master Degree program. This can be evidence from the college or university, or a signed document from the program administrator.</p>	90%	98%	95%	100%	<ul style="list-style-type: none"> • There is relatively little variation across levels in the proportion of programs meeting this requirement. As such, EEC may want to consider removing or modifying the requirement. • One potential approach would be to modify this criterion to require an associate's degree (see discussion of criterion 3A.2.2).
Level 2	3A.2.2. Program administrator with primary responsibility for the supervision of educators in the program has a BA degree.	Verify BA Degree is listed in the PQR.	54%	83%	88%	100%	<ul style="list-style-type: none"> • This criterion duplicates a Level 3 criterion (3A.3.1) and, if left at Level 2, would obviate the need for criterion 3A.2.1, which requires that administrators have a CDA or above or are working towards a degree. • The relatively low proportion of Level 1 programs meeting this requirement suggests that this is likely to be a significant barrier to progression as a Level 2 requirement. • One approach to revision would be to require a bachelor's degree at Level 3 (retain existing 3A.2.1) and modify criterion 3A.2.1 to require an associate's degree at Level 2.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	3A.2.3. Program administrator has a minimum of 3 college credits in administration and management and 12 college credits in early childhood education/child development/special education and 2 years' experience as an administrator.	<p>College Credits:</p> <ul style="list-style-type: none"> If the program administrator is Director I and/or Director II Qualified they meet the college credit requirement College credits must be from an EEC Teacher Qualification Unit approved college or university <p>Experience:</p> <p>Program administrator must have two years of experience working as the primary supervisor to educators (ex. director, education coordinator and/or site coordinator)</p>	87%	90%	90%	83%	<ul style="list-style-type: none"> This criterion and its related components do not appear to be substantively distinguishing quality across levels. The three credit requirement may be unnecessary in light of the requirement that administrators have a CDA or above (or associate's degree if the criterion is modified). This is also required for director certification. Similarly, three credits in management and administration are required for MA director certification and therefore this is redundant with basic MA Licensing requirements. <p>The vast majority of administrators at all programs had at least two years of experience.</p>
Level 2	3A.2.4. Program administrator is trained in the MA Guidelines for Preschool Learning Experiences, the Infant / Toddler Learning Guidelines, child development, the Strengthening Families protective factors, and is knowledgeable about the core competencies in order to be able to develop their staff's professional development plans.	<ul style="list-style-type: none"> Verify FORMAL professional development is listed in the PQR (college credit, min. .5 CEU, or .5 CEC) Formal training documented in PQR must indicate all of the following have been completed: <ul style="list-style-type: none"> MA Guidelines for Preschool Learning Experiences Infant/Toddler Learning Guidelines (only applicable if this age group is served) Child Development Strengthening Families Protective Factors 	54%	60%	85%	83%	<ul style="list-style-type: none"> All or nearly all administrators had received child development training, which is required for director certification. As such, this is redundant with MA Licensing criteria and should potentially be removed as a means of simplifying this compound criterion. Like many compound criteria, many administrators had training in some but not all of the others areas. To emphasize the importance of each area, consider separating the criterion into multiple criteria — one for training in the relevant Guidelines and one for training in Strengthening Families. Consider revising criterion to incorporate language regarding CEUs and credits.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	3A.2.5. Program administrator has received professional development in supervision of adults and strategies for working with adults.	<ul style="list-style-type: none"> Verify FORMAL professional development is listed in the PQR (college credit, min. .5 CEU, or min. .5 CEC) Formal training documented in PQR must indicate all of the following have been completed: <ul style="list-style-type: none"> Supervision of adults/ strategies for working with adults. 	79%	85%	98%	100%	<ul style="list-style-type: none"> Criterion 3A.2.5 may be duplicative with MA Licensing Director Certification, which requires completion of coursework in management and administration. Consider dropping or modifying this criterion.
Level 2	3A.2.6. Program administrator has an Individual Professional Development Plan (IPDP) that addresses their process and timelines to achieve the Program Administrator Qualifications for the next level of the QRIS and increases their competency along the advanced continuum of the core competencies.	<p>Program must submit a signed document that IPDP for the program administrator is:</p> <ul style="list-style-type: none"> created annually updated annually reviewed Annually <p>* Note: PQS Guidance also allows for verification of 3A.2.6 for NAEYC and Head Start programs.</p>	51%	85%	100%	100%	<ul style="list-style-type: none"> Consider revision of this criterion or the verification process to increase alignment. Note, the verification process seems to only require an annual IPDP. Data suggest that IPDPs may be of variable quality, and programs, particularly those at lower levels, may benefit from sample IPDPs or IPDP templates to facilitate the IPDP process and its usefulness.
Level 2	3A.2.7 Program has a Continuous Quality Improvement Plan addressing goals of the program related to QRIS	<ul style="list-style-type: none"> Verify aspects of the Program Improvement Plan. 					<ul style="list-style-type: none"> Consider adding as a requirement to the MA QRIS, as opposed to utilizing as verification tool to stress its importance and utilization.
Level 3	3A.3.1. Program administrator has at least a Bachelor's degree.	Verify program administrator listed in the PQR has a BA Degree	54%	83%	88%	100%	<ul style="list-style-type: none"> This criterion duplicates a Level 2 criterion (3A.2.2). Consider retaining this criterion and dropping criterion 3A.3.1.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	3A.3.2. Program administrator has at least 9 credit-bearing hours of specialized college-level course work in administration, leadership, and management.	Verify program administrator listed in the PQR has 6 credit-bearing hours of specialized college-level course work in administration, leadership, and management.	67%	73%	71%	83%	<ul style="list-style-type: none"> • Inconsistencies between criterion and verification need addressing. • Review of Higher Ed programming indicated that relatively few early education programs offer more than one course related to leadership and management. Data suggest that this criterion may have a high exemption rate as a result as well as too high of a standard in relation to what is provided by Higher Education. • Consider changing this criterion to require Director II certification, which requires additional coursework or, alternatively, working with institutions to build additional offerings that would meet the requirement. This would incorporate language and standards already familiar to the field in order to clarify expectations and promote a career ladder for MA QRIS that is consistent with MA Licensing.
Level 3	<p>3A.3.3. Program administrator has at least 24 credit-bearing hours of specialized college-level course work in early childhood education, child development, elementary education</p> <p style="text-align: center;">OR</p> <p>Documents that a plan is in place to meet the above qualifications within five years.</p>	<p>Verify program administrator listed in the PQR has 24 credit-bearing hours of specialized college-level course work in early childhood education, child development, elementary education, or early childhood special education</p> <p style="text-align: center;">OR</p> <p>Provide copy of IPDP outlining plan.</p>	90%	100%	95%	100%	<ul style="list-style-type: none"> • Nearly all administrators, regardless of Level, reported having the requisite credits to meet criterion 3A.3.3, likely due to the fact that most had an associate's degree in an early education and care-related field. (Currently, a CDA is required at Level 2 and a bachelor's degree at Level 2 and 3.) • Consider dropping these criteria given the limited variation and the fact that it appears to be redundant with other degree requirements.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	3A.3.4. Program administrator's IPDP addresses their process and timelines to achieve the Program Administrator Qualifications for the next level of the QRIS and increases their competency along the advanced continuum of the core competencies.	Verify that program administrator's IPDP is completed at least annually. * Note: PQS Guidance also allows for verification of 3A.3.4 for NAEYC programs.	51%	85%	100%	100%	<ul style="list-style-type: none"> Annual IPDPs are required at Level 2 (criterion 3A.2.6), and as such, criterion 3A.3.4 is already addressed at a lower level. Consider dropping this criterion or modifying it to reflect a higher level of requirement relative to 3A.2.6.
Level 4	3A.4.1. Program administrator has at least 9 credit-bearing hours of specialized college-level course work in administration, leadership, and management.	Verify program administrator listed in the PQR has 6 credit-bearing hours of specialized college-level course work in administration, leadership, and management.	67%	73%	71%	83%	<ul style="list-style-type: none"> This criterion is identical to a Level 3 criterion (3A.4.1). Consider dropping this criterion or moving Level 3 criterion to Level 4. As discussed above, we recommend changing this criterion to Director II certification.
Level 4	3A.4.2. Program administrator has at least 24 credit-bearing hours of specialized college-level course work in early childhood education, child development, elementary education, or early childhood special education.	Verify Program administrator listed in the PQR has 18 credit-bearing hours of specialized college-level course work in early childhood education, child development, elementary education, or early childhood special education.	90%	100%	93%	100%	<ul style="list-style-type: none"> Nearly all administrators, regardless of Level, reported having the requisite credits to meet criterion 3A.3.3, likely due to the fact that most had an associate's degree in an early education and care-related field. (Currently, a CDA is required at Level 2 and a bachelor's degree at Level 2 and 3) Consider dropping this criterion given the limited variation and the fact that it appears to be redundant with degree requirements.
Level 4	3A.4.3. Program administrator has a minimum of 5 years' experience as an administrator.	Program administrator must have two years of experience working as the primary supervisor to educators (ex. director, education coordinator and/or site coordinator)	85%	83%	85%	83%	<ul style="list-style-type: none"> Data for this standard did not consistently differentiate among Levels, with a few programs at each level not meeting standards. Experience requirements for directors could potentially make higher level programs hesitant to promote experienced teachers into administrative roles, potentially having an adverse impact on internal career ladders.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 4	3A.4.4. Program administrator has a Professional Development Plan that addresses areas of identified needs and increases their competency along the advanced continuum of the core competencies.	<p>Verify that Program administrator's IPDP is completed at least annually.</p> <p>* Note: PQS Guidance also allows for verification of 3A.4.4 for NAEYC programs.</p>	51%	85%	100%	100%	<ul style="list-style-type: none"> Annual IPDPs are already required at Level 2 (criterion 3A.2.6), and as such, criterion 3A.3.4 is already addressed at a lower level. Consider dropping this criterion or modifying it to reflect a higher level of requirement relative to 3A.2.6.

MA Quality Standard 3. Workforce Qualifications and Professional Development							
Standard 3B. Program Staff Qualifications and Professional Development							
	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	3B.2.1. All staff working in program have a high school diploma or GED.	All EDUCATORS who are considered in the licensing ratio, referred to as “regular employees,” must have one or more of the following: <ul style="list-style-type: none"> • HS Diploma • GED • CDA Note: This also applies to any educator who is considered a “regular employee” of a Public School classroom, or a classroom in an approved License-Exempt Program	77%	93%	98%	100%	<ul style="list-style-type: none"> • This criterion and its verification procedures appear to be working as expected. • One minor modification would be to revise the criterion read “all educators” instead of all staff, better aligning the criterion with the verification method.
Level 2	3B.2.2. 50 percent of classrooms have educator(s) with a Bachelor’s degree or higher who work for the full program day.	Check PQR to ensure that at 50% of classrooms have an educator with a Bachelor Degree. (If a program has four classrooms, the PQR should indicate that at least two educators have a Bachelor Degree.)	49%	65%	90%	83%	<ul style="list-style-type: none"> • This criterion appears to have a high exemption rate at Level 2. Programs may need support to meet this requirement and progress in the system. One possibility would be to provide more gradation in requirements between Licensing and Level 2. EEC may want to consider moving the BA requirement to Level 2 and instilling an AA requirement at Level 2.
Level 2	3B.2.3 Program has a Continuous Quality Improvement Plan addressing goals of the program related to QRIS	<ul style="list-style-type: none"> • Verify aspects of the Program Improvement Plan. 					<ul style="list-style-type: none"> • Consider adding as a requirement to the MA QRIS, as opposed to utilizing as verification tool to stress its importance and utilization.

	QRIS Criterion	• PQS Verification Guidance	L1	L2	L3	L4	• Comments and considerations
Level 2	3B.2.3. All educators (lead teachers, teachers, teaching aides, etc.) have a minimum of 3 college credits in early childhood education, or related field.	<ul style="list-style-type: none"> • Licensed Programs: All programs that are licensed and have a license in good standing meet this requirement • Public School Programs: All programs that are regulated by a Public School and are in good standing with DESE meet this requirement • License-Exempt Programs: Check PQR to ensure that at least one educator in each classroom has received 3 college credits in early childhood education (if a program has four classrooms, the PQR should indicate that at least two educators have received 3 college credits in ECE). 	n/a	n/a	n/a	n/a	<ul style="list-style-type: none"> • Reliable data on educator credits could not be obtained through the study. As such, this criterion was not assessed. • Note that MA Licensing requires a minimum of 3 college credits in early childhood education or related filed. As such, this requirement is redundant for teachers and lead teachers. If the goal is to increase the education of teacher aids and assistant teachers, we recommend that the criterion be altered to state the goal specifically for the purpose of clarification and not include teachers and lead teachers in the criterion.
Level 2	3B.2.4. All educators have an IPDP that is developed in conjunction with the supervisor that addresses the identified professional development needs of that educator and the development of their competency along the initial-level of the continuum of the core competencies. The IPDP must also address the actions and timelines that need to be met in order to move to the next level of the QRIS.	<p>Program must submit a signed document that states:</p> <ul style="list-style-type: none"> • An IPDP has been created for ALL educators • All IPDPs are reviewed annually • All IPDPs are updated annually. <p>* Note: PQS Guidance also allows for verification of 3B.2.4 for NAEYC and Head Start programs.</p>	56%	83%	100%	100%	<ul style="list-style-type: none"> • At lower levels, programs were more likely to be creating IPDPs for lead teachers and teachers than they were for assistant teachers. • Consider revision of this criterion or the verification to increase alignment. • Data suggest that programs at lower levels may benefit from sample IPDPs or templates related to what EEC expects be contained in an IPDP.
Level 3	3B.3.1. 75 percent of classrooms have educator(s) with a Bachelor’s degree or higher who work for the full program day.	<p>Check PQR to ensure that at 75% of classrooms have an educator with a Bachelor Degree.</p> <p>(If a program has five classrooms, the PQR should indicate that at least four educators have a Bachelor Degree.)</p>	41%	40%	83%	83%	<ul style="list-style-type: none"> • This criterion and its verification procedures appear to be working as expected. • The low proportion of programs at Levels 1 and 2 meeting the Level 3 requirement suggests this criterion is functioning as a substantial barrier to progression.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	3B.3.2. IPDP ensures that the educator receives formal professional development in the components of the assessment process including screening, observation, use of assessment tools, and IDEA processes.	Verify that all educators' IPDPS are completed at least annually. * Note: PQS Guidance also allows for verification of 3B.3.2 for NAEYC programs.	56%	83%	100%	100%	<ul style="list-style-type: none"> Annual IPDPs are required at Level 2 (criterion 3B.2.4), and as such, this criterion is already addressed at a lower level. Consider dropping this criterion or otherwise modifying it.
Level 4	3B.4.1. All (100 percent) of the classrooms have educator(s) with a Bachelor's degree or higher who work for the full program day.	Check PQR to ensure that at 100% of classrooms have an educator with a Bachelor Degree. (If a program has five classrooms, the PQR should indicate that at least five educators have a Bachelor Degree.)	28%	23%	73%	83%	<ul style="list-style-type: none"> This criterion and its verification procedures appear to be working as expected. The low proportion of programs at Levels 1 and 2, and to a lesser extent, Level 3 meeting the Level 4 requirement suggests this criterion is functioning as a substantial barrier to progression.
Level 4	3B.4.2. Educators have a minimum of 30 college credits in early childhood education/ child development/special education	Check PQR to ensure that at least one educator in each classroom has received 30 college credits in early childhood education. (If a program has four classrooms, the PQR should indicate that at least four educators have received 30 college credits in ECE.)	n/a	n/a	n/a	n/a	<ul style="list-style-type: none"> Unable to collect reliable data regarding educator college credits through the validation study. As such, this criterion was not assessed.
Level 4	3B.4.3. IPDP reflects identified opportunities for mentoring/coaching/ supervision available to educators, etc.	Document signed by program administrator that IPDP is completed regularly.	56%	83%	100%	100%	<ul style="list-style-type: none"> The verification of an Annual IPDPs are required at Level 2 (criterion 3B.2.4), and as such, the verification of this criterion does to address the criterion or provide address an aspect of quality not already verified above. Consider dropping this criterion or otherwise modifying verification process to require mentoring or coaching as part of the IPDP process.

MA Quality Standard 4. Family and Community Engagement							
	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	4A.2.1. Program completes Strengthening Families Self-Assessment and uses data to engage in continuous improvement.	Description of program improvement plan based on Strengthening-Families self-assessment including: <ul style="list-style-type: none"> • current goals • activities for strengthening family and community engagement 	23%	80%	76%	67%	<ul style="list-style-type: none"> • Interviews suggest Head Start programs may not be using Strengthening Families, but are using guidelines regarding family and community engagement which were seen to be aligned. • Consider amending to allow program to meet the requirement if they use a comparable tool at Level 2, and requiring the Strengthening Families protocol at Level 3. • Some programs did not see Strengthening Families as aligned with the needs of their clientele. Programs serving a larger proportion of children on subsidies were more likely to be using information gathered from this tool to inform programming (see criterion description below).
Level 2	4A.2.2. Programs offer opportunities for parents to meet with classroom staff, at least monthly.	Signed document demonstrating opportunities are: <ul style="list-style-type: none"> • planned • formal • intentional • initiated by the program Note: These opportunities can be face to face, electronically, or via phone.	79%	85%	93%	100%	<ul style="list-style-type: none"> • May be helpful to further clarify the purpose and types of engagement expected at Level 2 in verification guidance (i.e. formal conferences, drop-in or telephone hours, etc...) • The level of formality required may influence frequency expectations. Monthly might be too frequent for formal conferences but could be appropriate to less formal engagement.
Level 2	4A.2.3. Program has developed informational materials on the program that are in the language of the community, are available for staff to use in the community, and are given to prospective families.	<ul style="list-style-type: none"> • Verify that ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 3, and subscale score requirements of 2 or 3 depending on the subscale • Verify that the Program has submitted a Program Improvement Plan. 	67%	68%	93%	83%	<ul style="list-style-type: none"> • Verification procedures do not appear to match criterion. • When assessed directly, many Level 2 programs were not meeting the requirement, largely attributable to their not having materials translated into other languages when other languages are spoken in the community (see detailed criterion description below).

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	4A.2.4. Program maintains ongoing communication with the school/early intervention program, CFCE grantee, mental health providers to facilitate collaboration and coordination of services that support children and families.	<ul style="list-style-type: none"> PAS self-assessed score of 3 or higher List of agencies that program collaborates with Description of procedure for referring children and families Description of procedure for following up with services/supports for children/families. 	82%	98%	100%	100%	<ul style="list-style-type: none"> Overall PAS score for verification score does not address this criterion directly, and as such, consider dropping from the verification method.
Level 2	4A.2.5. Program participates in community events.	<ul style="list-style-type: none"> A list of events that the Program Staff has participated in over the past 12 months Events have to be in collaboration with other community organizations. <p>Note: Class field trips are not verification of this standard unless it is a community event</p>	77%	83%	95%	100%	<ul style="list-style-type: none"> The proportion of Level 2 programs who met the requirement was somewhat lower than might be anticipated. Consider clarifying the purpose of participation and provide examples of the types of events that would qualify.
Level 2	4A.2.6 Program has a Continuous Quality Improvement Plan addressing goals of the program related to QRIS	<ul style="list-style-type: none"> Verify aspects of the Program Improvement Plan. 					<ul style="list-style-type: none"> Consider adding as a requirement to the MA QRIS, as opposed to utilizing as verification tool to stress its importance and utilization.
Level 3	4A.3.1. A daily two-way communication system is available between the educators and families through a variety of means	<p>Program must have two or more of the following means of communication:</p> <ul style="list-style-type: none"> phone email notebook/log book checklists other <p>* Note: PQS Guidance also allows for verification of 4A.3.1 for NAEYC and Head Start programs.</p>	92%	98%	95%	100%	<ul style="list-style-type: none"> Across all Levels, all or nearly all programs indicated that they used multiple modes of communication with families on a frequent basis. Consider dropping this requirement or otherwise modifying it.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	4A.3.2. Families are encouraged to volunteer in the program, to assist in the classroom, and share cultural and language traditions or other interests such as their jobs, hobbies, and other relevant information.	<ul style="list-style-type: none"> Verify that ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 4.5, and subscale score requirements of 3 or 4 depending on the subscale Provide T.A. on ERS Subscale 7 after observations in 60% of preschool classrooms and 100% of infant and toddler classrooms. Verify that the Program has submitted a Program Improvement Plan. 	56%	65%	95%	100%	<ul style="list-style-type: none"> Consider using ERS indicators addressing family involvement (ECERS indicator 38.5.4 and ITERS indicator 33.5.4) for verification. Current verification method (overall ERS and subscale scores) does not assess the criterion directly. * Data for Level 4 reflects preschool classrooms only, as only one Level 4 program had an infant/toddler classroom.
Level 3	4A.3.3. Program ensures that there are translators available, as needed, at meetings, workshops and conferences to ensure strong communication between the program and families.	<p>Program must submit a description of process.</p> <p>If the program uses staff, identify the number of staff used as translators.</p> <p>If this service(s) is offered via an outside contractor, verify the MOU or provide proof that a collaboration is in place.</p> <p>* Note: PQS Guidance also allows for verification of 4A.3.3 for NAEYC and Head Start programs.</p>	82%	83%	93%	100%	<ul style="list-style-type: none"> Data suggest that MOUs have generally been difficult for programs to obtain from partners, even though collaborations may be on-going. Programs may need support in establishing these agreements or EEC may want to consider dropping the MOU requirement.
Level 3	4A.3.4. Program representative(s) participate in local community group work that is related to early childhood, and the cultural groups served by the program and/or family support.	Program Administration Scale (PAS) score of 5 or higher.	69%	78%	88%	100%	<ul style="list-style-type: none"> Overall PAS score does not address this criterion directly; consider assessing directly or dropping criterion.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	4A.3.5. Program ensures young children and their families have access to developmental, mental health, health, and nutrition services either through private pay arrangements OR are offered such services through other programs.	<ul style="list-style-type: none"> Verify that the program ensures that children and families have access to all of the following: <ul style="list-style-type: none"> developmental screenings mental health screenings health screenings nutrition screenings Verify that if children and families are offered this service via an outside contractor, an MOU or proof of collaboration is in place. <p>* Note: PQS Guidance also allows for verification of 4A.3.5 for NAEYC and Head Start programs.</p>	18%	43%	83%	83%	<ul style="list-style-type: none"> Some aspects of this criterion may be redundant with other requirements: <ul style="list-style-type: none"> Programs are required to conduct developmental screenings at Level 3 as part of criterion 1A.3.3. All or nearly all programs were tracking physical exams; programs are required to maintain documentation of annual physical examinations as part of licensing. Further clarification of how programs are expected to ensure access might be beneficial (e.g. track and referral).
Level 4	4A.4.1. Parents participate on the Advisory Board for the program and are actively involved in the policy and decision making for the program.	<p>Verify that:</p> <ul style="list-style-type: none"> ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 5.5, and subscale score requirements of 4 or 5 depending on the subscale ERS Reliable Rater scores for each classroom meet or exceed the overall score requirement of 5.5, and subscale score requirements of 4 or 5 depending on the subscale Program has submitted a Program Improvement Plan Technical Assistance Visit has been completed. <p>* Note: PQS Guidance also allows for verification of 4A.4.1 for NAEYC and Head Start programs.</p>	23%	43%	83%	100%	<ul style="list-style-type: none"> Consider using ERS indicators that address family involvement in program decision-making (ECERS indicator 38.7.3 and ITERS indicator 33.7.3) for verification. Current verification method (overall ERS and subscale scores) does not assess the criterion directly. <p>* Data for Level 4 reflects preschool classrooms only, as only one Level 4 program had an infant/toddler classroom</p>

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 4	4A.4.2. Program provides or connects families to education, training and support programs (such as family literacy, adult education, job training, child development, parenting, English as a second language, etc.).	<ul style="list-style-type: none"> Documentation that the Program connects families to education, training and support. MOU or proof that a collaboration is in place with the agencies who can provide these services. <p>* Note: PQS Guidance also allows for verification of 4A.4.2 for Head Start programs.</p>	13%	38%	78%	67%	<ul style="list-style-type: none"> Although examples are provided, guidance does not specify how many or what services are required for a program to meet this criterion. Some programs noted that services of this nature were not typically required by the families they served; clarification regarding expectations in these cases may be beneficial. Data suggest MOUs have been difficult for programs to obtain, even as they regularly collaborated with partners. Programs may need support or EEC may want to consider dropping this requirement.
Level 4	4A.4.3. Program ensures all children and families have access to comprehensive screenings, referrals and services including developmental screening, mental health screening, speech screening, speech therapy, physical therapy, occupational therapy, dental health care, and nutrition services.	<p>Document signed by program administrator demonstrating that children have access to the following screenings, referrals, and services:</p> <ul style="list-style-type: none"> developmental screening mental health screening speech screening speech therapy physical therapy occupational therapy dental health care nutrition services <p>These may be provided either through the program or another service delivery such as EI, public school, CFCE Program, or by family arrangement.</p> <ul style="list-style-type: none"> If this is offered via an outside contractor, verify the MOU or provide proof that a collaboration is in place. <p>* Note: PQS Guidance also allows for verification of 4A.4.3 for Head Start programs.</p>	15%	30%	76%	67%	<ul style="list-style-type: none"> Some aspects of this criterion may be redundant with other requirements: <ul style="list-style-type: none"> Programs are required to conduct developmental screenings at Level 3 as part of criterion 1A.3.3 Speech services, physical therapy, and occupation therapy would typically only be received by children referred for special services through Early Intervention or special education. Programs are required to work with Early Intervention and special education providers at Level 2 (criterion 4A.2.4). One aspect of criterion 4A.4.3 that is unique is the requirement that children have access to dental health care. Dental screenings are required at Level 3 as part of criterion 2A.3.1. Further clarification of how programs are expected to ensure access might be beneficial (e.g. track and referral).

MA QRIS Quality Standard 5. Leadership, Management, and Administration							
5A. Leadership, Management, and Administration							
	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	5A.2.1. Communication and updates on the program are provided to educators and families.	<p>Program must provide a statement that program updates are provided:</p> <ul style="list-style-type: none"> at least quarterly to staff and to families in their primary, or preferred, language to the extent appropriate and possible explanation of how updates are provided (e.g., newsletter, email, bulletin board posting, updated copy of handbook) <p>* Note: PQS Guidance also allows for verification of 5A.2.1 for Head Start programs.</p>	100%	100%	100%	100%	<ul style="list-style-type: none"> All programs were providing program updates to families and staff at least quarterly. As this criterion is not distinguishing quality, consider dropping this requirement or otherwise modifying it.
Level 2	5A.2.2. Program has a written admissions policy that promotes an awareness of and respect for differences among children and families, a respect for the child and their family's culture and language, and is responsive to the inclusion of a variety of learning needs.	<p>Verify that:</p> <ul style="list-style-type: none"> ERS self-assessed scores for each classroom meets or exceeds the overall score requirement of 3, and subscale score requirements of 2 or 3 depending on the subscale Program has submitted a Program Improvement Plan Program has an admissions policy promoting diversity. <p>* Note: PQS Guidance also allows for verification of 5A.2.2 for NAEYC and Head Start programs.</p>	77%	90%	88%	83%	<ul style="list-style-type: none"> Overall ERS score does not address this criterion directly; consider dropping from verification. Examples of admissions policies that promote diversity might also be beneficial.
Level 2	5A.2.3. Program has a written business plan.	Copy of business plan including annual operating budget	69%	95%	100%	100%	<ul style="list-style-type: none"> Overall, this criterion and its verification appear to be working as intended. One minor modification might be to include a reference to an annual operating budget in the criterion itself.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	5A.2.4. Staff are paid for planning time.	PAS overall score of 3 or higher	67%	78%	90%	83%	<ul style="list-style-type: none"> Overall PAS score does not address this criterion directly; consider more direct verification by assessing program policies. May also want to consider clarification of whether planning time needs to be when children are not present (i.e. can nap time be considered planning time)
Level 2	5A.2.5. Program has policies that support teacher retention.	Teacher retention policies may include: <ul style="list-style-type: none"> vacation and/or sick time education/tuition reimbursement pay increases/career ladder free or reduced child care retirement plan or contribution 	100%	100%	100%	100%	<ul style="list-style-type: none"> All programs reported having at least one of the policies listed in verification guidance. Because the criterion, as currently construed, is redundant and not distinguishing quality, consider dropping or modifying it.
Level 2	5A.2.6 Program has a Continuous Quality Improvement Plan addressing goals of the program related to QRIS.	<ul style="list-style-type: none"> Verify aspects of the Program Improvement Plan. 					<ul style="list-style-type: none"> Consider adding as a requirement to the MA QRIS, as opposed to utilizing as verification tool to stress its importance and utilization.
Level 3	5A.3.1. Program tracks and monitors absences of individual children and contacts families when children are absent more than 20 percent in a month.	Program must provide evidence that they: <ul style="list-style-type: none"> Use a paper-based or computer based system for tracking absences. Contact families when children are absent more than 20% in a month Provide support to families to help reduce the child's absences. 	97%	95%	95%	100%	<ul style="list-style-type: none"> All or nearly all programs tracked children's absences and contacted families. Interview data suggest that programs' policies were mixed, often dependent on whether they participated in subsidy programs. Consider incorporating requirement regarding support into the criterion itself if this is to remain a part of verification and clarify expectations for attendance policies for private pay families.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	• Comments and considerations
Level 3	5A.3.2. Program director, staff and family input is solicited on an annual basis through a survey to evaluate the program.	PAS score of 5 or higher.	36%	83%	93%	100%	<ul style="list-style-type: none"> • Consider combining with criterion 5A.3.3, which requires use of survey results for program planning and improvement. • Overall PAS score does not address this criterion directly; consider more direct verification by assessing program policies.
Level 3	5A.3.3. Results of the annual survey are used to develop a comprehensive written program improvement plan.	Verify that the results of family and staff survey are incorporated into Program Improvement Plan.	38%	75%	90%	100%	<ul style="list-style-type: none"> • Consider combining with criterion 5A.3.2, which requires administration of annual surveys for families and staff.
Level 3	5A.3.4. Program has an annual review conducted of the accounting records by an independent party who has accounting or bookkeeping expertise.	Proof of annual review by independent party.	62%	85%	85%	100%	<ul style="list-style-type: none"> • This criterion and its verification appear to be working largely as expected. • Most programs at Level 2 were also engaged in this practice.
Level 3	5A.3.5. Program tracks and monitors teacher turnover and has plan for addressing turnover.	Provide a description of the system for tracking teacher turnover and the plan for addressing teacher turnover.	36%	50%	80%	100%	<ul style="list-style-type: none"> • Most programs indicated that they tracked teacher turnover. Fewer indicated that they had a plan to address turnover. This may point to a need to focus more explicitly on programs' plans for addressing turnover when verifying this criterion.
Level 4	5A.4.1. Program staff and advisory board are involved in the development of the business plan and it is reviewed periodically for updating.	Documentation that program staff and advisory board are involved in the development of the business plan and it is reviewed periodically for updating.	8%	40%	63%	50%	<ul style="list-style-type: none"> • At higher levels, programs were more likely to involve staff and advisory boards in the development or review of business plans, but the overall proportion was low across all levels. Data suggest that more clarification of the requirements need to be specified including distinguishing the difference between being "involved" in its development versus simply "approving" the plan and if approving a plan meets the requirement for "involved in the development".

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 4	5A.4.2. Program has a system of technology that allows for data collection and tracking program information.	Program has a system of technology that maintains and tracks: <ul style="list-style-type: none"> • children’s health • services • absenteeism • children’s educational information • staff qualifications • professional development • financial records/information • teacher turn-over rates • other (if other, please explain) 	3%	10%	39%	33%	<ul style="list-style-type: none"> • Data systems were common among higher level programs, but few of those systems contained all of the listed items. • It may be helpful to clarify whether all data elements are required. • Systems were most commonly used for maintaining data on children. Consider separating this criterion into two criteria — one for child data and one for staff data — to ensure both are addressed.
Level 4	5A.4.3. An outside audit is conducted annually by a certified public accountant.	Verification that an outside audit is conducted annually by a certified public accountant.	59%	75%	78%	100%	<ul style="list-style-type: none"> • This criterion and its verification appear to be working as intended.
Level 4	5A.4.4. Program shares the results of the program quality rating with the families, staff, governing board and funders.	Verify that: <ul style="list-style-type: none"> • ERS self-assessed scores for each classroom meet or exceed the overall score requirement of 5.5, and subscale score requirements of 4 or 5 depending on the subscale • ERS Reliable Rater scores for each classroom meet or exceed the overall score requirement of 5.5, and subscale score requirements of 4 or 5 depending on the subscale • Program has submitted a Program Improvement Plan • Technical Assistance Visit has been completed. 	n/a	n/a	n/a	n/a	<ul style="list-style-type: none"> • Current verification (overall ERS and subscale scores) does not address this criterion directly. • More direct verification of the criterion would be difficult, as it would imply sharing a rating before that rating was granted. • Consider dropping this requirement or otherwise modifying it.

MA QRIS Quality Standard 5. Leadership, Management and Administration							
5B. Supervision							
	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 2	5B.2.1. Program provides recognition for staff in annual evaluation as well as in public forum, as appropriate (i.e. verbal recognition in group setting or written recognition in newsletter).	Verify that Program provides recognition for staff in: <ul style="list-style-type: none"> Annual Evaluation Public Forum (ex. staff meeting, newsletter, bulletin board) 	92%	98%	100%	100%	<ul style="list-style-type: none"> All or nearly all programs conducted annual evaluations of staff. Consider dropping or otherwise modifying this criterion.
Level 2	5B.2.2. Staff are given feedback on instructional practice on a monthly basis.	Verify that the all educators are: <ul style="list-style-type: none"> observed by a supervisor while working with children at least monthly are given feedback on instructional practice. 	67%	73%	100%	100%	<ul style="list-style-type: none"> More than a quarter of Level 2 programs were not meeting the ERS indicator related to feedback (every other month vs. monthly) based on teacher reports. This suggests a need for more robust verification protocols that go beyond program policies and/or administrator reports.
Level 2	5B.2.3 Program has a Continuous Quality Improvement Plan addressing goals of the program related to QRIS	<ul style="list-style-type: none"> Verify aspects of the Program Improvement Plan. 					<ul style="list-style-type: none"> Consider adding as a requirement to the MA QRIS, as opposed to utilizing as verification tool to stress its importance and utilization.
Level 3	5B.3.1. Program uses at least 3 types of internal communication on a monthly basis to inform staff of program activities, policies, etc.	Provide description of at least three types of communications. Ex: <ul style="list-style-type: none"> email postings meetings phone calls newsletters memos other 	97%	98%	98%	100%	<ul style="list-style-type: none"> All programs were using at least three forms of communication on a monthly basis. As this criterion is not distinguishing quality, consider dropping this requirement or otherwise modifying it.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	5B.3.2. Staff receive at least one benefit (paid vacation time, sick time, health insurance, tuition/PD reimbursement or retirement plan option).	<ul style="list-style-type: none"> Program Administration Scale (PAS) score of 5 or higher. Must provide an example of at least one benefit, ex. paid vacation time, sick time, health insurance, tuition/PD reimbursement, or retirement plan option 	97%	100%	100%	100%	<ul style="list-style-type: none"> All programs were offering at least one benefit listed in the PQS guidance (discussion below). As this criterion is not distinguishing quality, consider dropping this requirement or otherwise modifying it. Overall PAS score does not address this criterion directly. If the criterion is maintained, consider dropping PAS score from verification method.
Level 3	5B.3.3. Staff are given feedback that give examples of best practice at least twice a month.	Program Administration Scale (PAS) score of 5 or higher	15%	23%	20%	50%	<ul style="list-style-type: none"> Overall PAS score does not address this criterion directly; consider more direct verification by assessing program policies. Reliable measurement of this criterion may be difficult and EEC may want to consider dropping this criterion as unmeasurable or otherwise revising it to reflect a more directly identifiable and directly measurable practice.
Level 3	5B.3.4. The program has a system to support the career development of staff through a career ladder (e.g., regularly scheduled time to meet with a supervisor or mentor to monitor progress towards career goals).	Program Administration Scale (PAS) score of 5 or higher.	79%	95%	100%	100%	<ul style="list-style-type: none"> Overall PAS score does not address this criterion directly; consider more direct verification by assessing program policies. All or nearly all programs at Levels 3 and above had one of the career ladder supports listed in the criterion in place. Consider moving this criterion to Level 2.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 3	5B.3.5. Staff salary scales reflect the educational levels, experience, and performance levels, as determined by the annual evaluation of the staff members, and is comparable with the current wage level of others in the community with the same levels of education.	Program Administration Scale (PAS) score of 5 or higher.	54%	45%	22%	50%	<ul style="list-style-type: none"> Overall PAS score does not address this criterion directly; consider more direct verification by assessing program policies. When assessed directly, many programs were not meeting this requirement, especially as it relates to considering experience (for new staff) and performance (for existing staff). This may be a more challenging criterion for programs with more infrastructures or established pay scales to meet, particularly Head Start programs.
Level 3	5B.3.6. Educators are provided with morning, midday, and afternoon breaks.	Verified by Personal Needs of Staff Item on ECERS-R and ITERS-R.					<ul style="list-style-type: none"> Based on ERS data, educators report minimal break time. Given the high demands of the job, break time support may be an area in which MA EEC wants to provide added support for educators.
Level 4	5B.4.1. Program offers a benefit package that includes vacation, sick time, and health insurance.	Verify that the program has a written policy that shows that benefit packages are offered that includes: <ul style="list-style-type: none"> vacation sick time health insurance. 	46%	88%	90%	100%	<ul style="list-style-type: none"> This criterion and its verification appear to be working as expected. Few programs at Level 1 meet this requirement, largely because they were not providing health insurance. It may be beneficial to clarify who must receive these benefits for a program to be considered meeting the standard (full time, part time, etc...)
Level 4	5B.4.2. Staff are provided ongoing mentoring that includes demonstration of best practices on a weekly basis.	Verify that there are regular opportunities for teaching staff to engage in reflective teaching practices, peer group coaching and mentoring.	3%	23%	20%	67%	<ul style="list-style-type: none"> This criterion would benefit from further clarification of EEC requirements for meeting the criterion.

	QRIS Criterion	PQS Verification Guidance	L1	L2	L3	L4	Comments and considerations
Level 4	5B.4.3. Program demonstrates systematic opportunities for educators to engage in reflective teaching practices through the use of peer groups, coaches, and/or mentors.	Verify that there are regular opportunities for teaching staff to engage in reflective teaching practices, peer group coaching and mentoring.	3%	23%	20%	67%	<ul style="list-style-type: none"> This criterion, as verified, is identical to criterion 5B.4.2 above and therefore is considered repetitive. Consider dropping this criterion or modifying verification to make it distinct.
Level 4	5B.4.4. Program has an incentive program that rewards each educator who achieves the next step on the career ladder.	Program Administration Scale (PAS) score of 7.	82%	93%	93%	100%	<ul style="list-style-type: none"> Overall PAS score does not address this criterion directly; consider more direct verification by assessing program policies. All or nearly all programs at Levels 4 had one of the following career ladder incentives in places — pay scale based on professional qualifications, greater responsibility with increasing education and credentials, merit-based pay increases, or more senior job titles for educators with higher credentials — as did nearly all at Levels 2 and 3.

MA QRIS Quality Standard1A: Curriculum, Assessment and Diversity

As noted in the program quality specialist guidance, utilization of a developmentally appropriate, comprehensive curriculum aligned to the state’s standards; use of assessments that reflects the curriculum and state standards; and adaptations to meet individual children’s needs reflect critical elements of a high-quality early education and care program. These aspects of quality are addressed in the Curriculum, Assessment, and Diversity quality standard for the MA QRIS, which currently includes eight criteria: two at Level 2, four at Level 3, and two at Level 4.

1A.2.1. (Level 2) Educators demonstrate completion of formal professional development in curriculum, screening tools, and formative assessment.

For a program to meet this criterion, PQS guidance requires that at least one educator in each classroom has received formal professional development in curriculum, screening tools, and formative assessment. Data gathered through administrator questionnaires indicate that, at Levels 2 and above, most programs are meeting this criterion. As Table 11 shows, at Levels 2 and 3, 88 percent of programs have one educator in each classroom who has received curriculum training, and at Level 4, 83 percent of programs have an educator trained in curriculum in each classroom. For formative assessment training, 78 percent of programs at Level 2 and 88 percent of programs at Level 3 reported at least one educator in each classroom had received this training. (Data regarding training in the use of screening tools were not gathered on study instruments.) However, when considering if one person from each classroom had both of the trainings, the percentage decreases. It may be beneficial to consider separating these requirements into two criteria to allow programs to track PD more efficiently and emphasize the need for at least an educator in each classroom to have training on curriculum AND one educator from each classroom to have training on assessment.

Table 11. Programs meeting measured aspects of criterion 1A.2.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 1A.2.1	56%	73%	83%	67%
At least one educator in each classroom has received training in curriculum	79%	88%	88%	83%
At least one educator in each classroom has received training in assessment	69%	78%	88%	83%

Data regarding professional development were also collected directly from lead teachers in observed study classrooms. Among these teachers, 80 percent in Level 2 programs, 91 percent in Level 3 programs, and 100 percent in Level 4 programs indicated they had received curriculum training. Similarly, 80 percent of teachers in Level 2 programs, 96 percent of teachers in Level 3 programs, and 100 percent of teachers in Level 4 programs indicated they had received training in formative assessment. Although these data reflect only a subset of classrooms at a program, they do suggest that it is possible that educators may have more training than is being reported to administrators. These data appear to suggest that continued promotion of formal systems for reporting and tracking trainings, for example IPDPs, and/or educator use of the Professional Qualifications Registry (PQR) might be beneficial to ensure training is documented and tracked effectively by directors and administrators. It should be noted that researchers were unable to determine if training received by teachers were CEU-level trainings, as most educators and administrators were unable to verify whether CEUs were given and the PQR is not consistently up to date. Due to a lack of understanding of CEUs evident during data collection, it is anticipated that the above percentages would go down if only CEU trainings in curriculum and assessment were considered.

1A.2.2. (Level 2) Materials reflect the language and culture of the children in the classroom, their communities, and represent the diversity of society.

Criterion 1A.2.2 is assessed using a program's ERS scores and whether those scores meet Level 2 requirements overall and for individual subscales. However, because overall and subscale scores take into account multiple items of quality, it is possible for programs to meet verification standards but not have representative materials in their classroom. As such, for the QRIS validation, the study incorporated an item into its classroom observation protocols in which observers were asked to identify if the materials in the classroom reflected the diversity of children present.

As Table 12 shows, approximately two-thirds of programs at Levels 2 and 3 met this criterion when it was assessed directly through observation, suggesting that verification of this criterion using overall ERS scores may be insufficient to ensure that the requirement is met. One possible alternative would be to use of the Promoting Acceptance of Diversity item of the ERS to capture a clearer picture of materials related to diversity. As can be seen from classroom observations using the ERS, almost all preschool classrooms met the Level 2 cut-off on the ERS (3=Minimal quality) for the Promoting Acceptance of Diversity item. More variability, however, was noted among infant and toddler classrooms. Given that Level 2 is an ERS self-assessed-only level, which makes verification more subjective, another alternative would be to move this criterion to Level 3. This would require a visit from a program quality specialist that includes an observation, during which time the specialist could verify the presence of diverse materials and offer technical assistance as required.

Table 12. Programs meeting measured aspects of criterion 1A.2.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets observation requirement for criterion 1A.3.1 for all study classrooms (up to two classrooms per program, one infant/toddler and one preschool).	36%	68%	68%	83%
Preschool classroom materials reflect the diversity of children in the classroom (programs with observed preschool classroom only)	44%	71%	70%	80%
Infant or toddler classroom materials reflect the diversity of children in the classroom (programs with observed infant or toddler classroom only)	56%	90%	86%	100%
Percent of programs meeting cut-off a 3 ECERS-R Promoting Acceptance of Diversity Item	95%	100%	100%	100%
Percent of programs meeting cut-off a 3 ITERS Promoting Acceptance of Diversity Item	70%	93%	80%	100%

1A 3.1. (Level 3) Staff include parental input in progress reports.

To meet this criterion, programs must develop progress reports for all children and share those progress reports with families; meet frequency guidelines outlined in the PQS guidance — three times in a 12-month period for preschoolers and four times in a 12-month period for infants, toddlers, and children with disabilities; and have an area on the report for parents to offer comments, such as goals they may have for their child or other observations. Data collected through the validation study show that, regardless of QRIS level, nearly all programs develop and share progress reports with families, ranging from 95 percent of programs at Level 2 to 100 percent of programs at Level 4 (Table 13). Similarly, most programs at Level 3 and 4 gather family input as part of the progress report process, which is consistent with that aspect of this Level 3 criterion.

However, even at Levels 3 and 4 relatively few programs met the requirement regarding the frequency of progress reporting for all ages served (39 percent and 33 percent, respectively). This was most often the result of programs indicating that they prepared progress reports for toddlers on a similar schedule as they did for preschoolers, rather than at the same frequency as infants and children with special needs as the PQS guidance

requires. We suggest that the criteria incorporate timelines to clarify parameters as ambiguity seems to exist regarding expectations, particularly for the toddlers. Notably, MA Licensing groups toddlers with preschoolers in terms of the frequency requirement, which may be a contributor to this confusion. MA Licensing regulations require that progress reports are developed for preschools and toddlers every six months (twice annually) and for infants and children with special needs every three months (four times each year). In addition, further clarification of what is required in terms of “parent input” would be beneficial. For the study, a program was considered to be including parent input in progress reports if they actively sought to collect and record parent input, for example if they asked parents to offer observations or participate in the development of the child’s learning goals.

Table 13. Programs meeting measured aspects of criterion 1A.3.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured requirements for criterion 1A.3.1	5%	25%	39%	33%
Program develops written progress reports for all children	97%	95%	98%	100%
Program shares written progress reports with families	95%	95%	98%	100%
Program develops and share progress reports with sufficient frequency for all age groups served	5%	25%	39%	33%
...for <u>preschoolers</u> at least three times each year	10%	59%	78%	80%
...for <u>toddlers</u> at least four times each year	17%	23%	25%	n/r
...for <u>infants</u> at least four times each year	84%	72%	58%	n/r
...for <u>children with special needs</u> at least four times each year	56%	58%	51%	67%
Program includes parental input in progress reports	51%	80%	93%	100%

n/r – not reported due to sample size of 1

1A.3.2. (Level 3) Staff have received formal professional development in the curriculum; using the MA Guidelines for Preschool Learning Standards or Infant /Toddler Learning; documenting children's progress; and working with children from diverse languages and cultures and second language acquisition.

To meet this criterion, verification guidance requires that all of a program’s lead teachers have received formal professional development in multiple areas, including curriculum; are using the MA Guidelines for Preschool Learning Experiences and, if applicable, Infant/Toddler Learning Guidelines; are documenting student progress, working with children from diverse cultures; and working with early English language learners. Information collected from program directors indicates that at Level 3 and above, most programs (ranging from 73 percent to 88 percent) meet requirements for each individual training area. However, only slightly more than half (59 percent) had staff that met all required trainings (Table 14). Notably, training in the Massachusetts Guidelines, an important foundational aspect of curriculum and learning, was the least commonly received training both at Levels 3 and 4 where it is required for the MA QRIS.

Because many Level 3 and 4 programs appear to be meeting some, but not all of the requirements, it may be beneficial to separate this criterion into its component parts, for example, a criterion related to Guidelines training; a criterion related to working with children from diverse cultures and English language learners; and a criterion related to documenting student progress. (The requirement regarding curriculum training may be redundant with a requirement in criterion 1A.2.1, and as such, might reflect an aspect of this criterion that could be removed). This will stress the importance of each of the trainings, and the need for each. Further, because curriculum training is also required for standard 1A.2.1 (Level 2), with no indication of the differential requirements, this part of the criterion appears to be redundant and so should be removed. In terms of field practice, training in the guidelines may be more beneficial at a lower level given that the guidelines are designed to serve as general standards for practice.

Table 14. Programs meeting measured aspects of criterion 1A.3.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured requirements for criterion 1A.3.2	21%	35%	59%	67%
At least one educator in each classroom has received training in curriculum	79%	88%	88%	83%
At least one educator in each classroom has received training in the Massachusetts Guidelines for all relevant ages served	46%	65%	71%	83%
At least one educator in each classroom has received training in documenting student progress	56%	80%	83%	83%
At least one educator in each classroom has received training in working with children from diverse languages and cultures	33%	58%	80%	83%

1A.3.3. (Level 3) Program uses screening tools, progress reports, formative assessments, and information gathered through observation to set goals for individual children across all developmental domains.

Programs are verified for criterion 1A.3.3 if they complete developmental screenings on all children using an evidence-based tool, conduct formative assessments on all children using an evidence-based tool, share the results of those screenings and assessments with families, and use the information to create activities to support all children's individual needs. At Levels 3 and 4, at which this criterion is required, data collected through the validation study show that all programs conduct formative assessments and all or nearly all programs conduct screenings with all children (Table 15).

Table 15. Programs meeting measured aspects of criterion 1A.3.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 1A.3.3	23%	65%	95%	100%
Program conducts screenings with all children	26%	73%	95%	100%
Program conducts formative assessments with all children	77%	88%	100%	100%
Program conducts formative assessment with all children using an EEC-approved tool (descriptive only, not included in the proportion of programs meeting measured aspects of criterion 1A.3.3)	33%	77%	98%	100%

Among programs at Levels 3 and 4, all those who were conducting formative assessments did so using an EEC-approved formative assessment tool, which include Teaching Strategies Gold, Work Sampling System, and the High Scope Child Observation Record (COR). At Level 1, programs were most commonly using their own self-developed tool for ongoing assessment, and at Level 2, some programs were using self-created tools. As such, only 33 percent of programs at Level 1 and 77 percent of programs at Level 2 were using a tool that had been pre-approved by EEC. The use of an EEC-approved assessment may create a barrier for some programs as they seek to move up within the system, particularly those currently at Level 1. It may be necessary for EEC to develop strategies to support programs in obtaining these tools and being trained in their use — for example, by offering subsidies or assessment grants to account for the high costs materials and training often associated with these types of assessments.

1A.3.4. (Level 3) Staff demonstrate language and literacy skills either in English or the child's language that provide a model for children.

Attainment of this criterion is assessed through a program's ERS scores and whether those scores meet Level 3 requirements both overall and for individual subscales. However, the aspect of quality most evident in this criterion, namely a program's language and literacy environment and interaction might be assessed more directly using the Language and Literacy subscale of the ECERS-R and the Listening and Talking subscale of the ITERS-R. As Table 16 indicates, at Levels 3 and 4, all or nearly all programs achieved a score of 4 or above (the QRIS requirement for Level 3 for these subscales). At lower levels, fewer programs (51 percent at Level 1 and 68 percent at Level 2) attained scores at this level or above.

Table 16. Programs meeting measured aspects of criterion 1A.3.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets score of 4 or above for all study classrooms (up to two classrooms per program, one infant/toddler and one preschool).	51%	68%	88%	nr
Preschool classroom received a score of 4 or above on the ECERS-R Language and Reasoning subscale (programs with an observed preschool classroom only)	62%	85%	90%	100%
Infant or toddler classroom received a score of 4 or above on the ITERS-R Listening and Talking subscale (programs with an observed infant or toddler classroom only)	59%	74%	93%	nr

1A.4.1. (Level 4) Program uses a curriculum that is aligned with MA guidelines for Preschool Learning Standards and the Infant/ Toddler Learning Guidelines.

Programs are verified for criterion 1A.4.1 if they demonstrate that they use a curriculum aligned with the core competencies outlined in the Massachusetts Guidelines for Preschool Learning Standards and/or the Infant/Toddler Learning Guidelines. Programs can demonstrate this by using one of the EEC-approved curricula — which include High Scope, Creative Curriculum, Opening the World of Learning (OWL), or Resources for Early Learning — or they can provide an example of the curriculum they use for EEC review against the Guidelines. The study was unable to assess programs that could meet requirements using the latter approach, and thus data regarding curriculum use are somewhat limited.

Across all levels, large proportions of programs reported having an identified curriculum for all ages served (Table 17). However, programs in Levels 2, 3, and 4 were much more likely to be using a comprehensive, published curriculum, of which Teaching Strategies Creative Curriculum was the most common (used by 86 percent of programs that used a published curriculum). Among programs in the three higher levels of the QRIS, at least three-quarters were using a comprehensive published curriculum for children at all ages, compared with approximately one-quarter of programs at Level 1. The remainder was either using a self-developed curriculum or did not have an identified curriculum. Although some self-developed curricula may be aligned with the Massachusetts Guidelines, and thus meet the requirements, a more thorough review of the curriculum would be needed to determine whether the self-made curriculum meets this QRIS criterion. Additionally, analyses for the study indicated a link between curriculum, and quality (discussed in Program Characteristics and Quality section) which further underscores the need to support the implementation of a vetted curriculum. The relatively low proportion of programs using a comprehensive, vetted curriculum at Level 1 and the sizeable proportion of programs Level 2 not doing so highlights this as a potential area of focus, particularly given that curriculum implementation often requires significant planning, training, and financial support. Ultimately, the development of strategies to ensure that smaller programs with less access to resources and adequate supports — such as mentor programs or forming learning communities among programs, as well as EEC-sponsored curriculum support grants — may be critical to supporting movement within the system.

Table 17. Programs meeting measured aspects of criterion 1A.4.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program uses an identified curriculum with all ages served	72%	90%	98%	100%
Administrator interview: Program uses a comprehensive curriculum in preschool classrooms (only programs that serve preschoolers)	79%	92%	98%	100%
Administrator interview: Program uses a comprehensive curriculum in infant and/or toddler classrooms (only programs that serve infants and/or toddlers)	72%	90%	100%	100%
Program uses a <u>prepackaged comprehensive</u> curriculum with all ages served	28%	73%	80%	83%
Program uses a prepackaged comprehensive curriculum in preschool classrooms (only programs that serve preschoolers)	31%	74%	80%	80%
Program uses a prepackaged comprehensive curriculum in infant/toddler classrooms (only programs that serve infants and/or toddlers)	24%	68%	88%	100%

1A.4.2. (Level 4) Program uses progress reports, appropriate screening tools, formative assessments, and information gathered through observation to inform curriculum planning, use results to monitor each child's progress across developmental domains, and inform program decision-making (e.g. curriculum content, strategies for improved staff implementation, and professional development).

Programs receiving a Level 4 designation are required to provide documentation that data from children’s assessments and screenings are used to monitor progress across developmental domains and inform program decision-making (criterion 1A.4.2). Data gathered in administrator interviews indicate that at Level 4, all programs were engaged in both these practices, as the criterion requires (Table 18). These practices also appear to be common among Level 3 programs and, to a lesser extent, those at Level 2.

Table 18. Programs meeting measured aspects of criterion 1A.4.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets measured aspects of criterion 1A.4.2	41%	78%	90%	100%
Program tracks children's progress across all developmental domains.	79%	93%	98%	100%
Program uses child assessment and/or screening data when making decisions about the program (e.g. changes to the curriculum or resources uses, in planning professional development, or when purchasing materials)	51%	83%	93%	100%

MA QRIS Quality Standard1B: Teacher-Child Relationships and Interactions

Another important aspect of quality involves positive teacher-child interactions that can foster children’s self-regulation and emotional well-being. This standard of the QRIS comprises a total of four criteria, one at Level 2, two at Level 3, and one at Level 4.

1B.2.1. (Level 2) All staff receive orientation and ongoing professional development and supervision in how to support positive relationships and interactions through positive, warm, and nurturing interactions.

Criterion 1B.2.1 is based on information contained within the Professional Qualifications Registry, or PQR. For the validation study, this criterion was assessed using indicators on the ERS observations for preschool and infant/toddler classrooms that addressed both orientation for new staff and in-service training. As shown in Table 19, all or nearly all programs at Level 2 and above offered orientation for new staff that addressed interaction, discipline methods, and appropriate activities, and regularly provided in-service training, which is consistent with this Level 2 requirement. Far fewer programs at Level 1 (56 percent) provided both orientation and training. As such, this criterion appears to be operating as expected. However, there may be opportunities to refine the language of the criterion by removing some redundancy with respect to references to interactions. An alternative might be to drop the portion that reads *interactions through positive*, such that the criterion would read “*All staff receive orientation and ongoing professional development and supervision in how to support positive relationships, and warm, and nurturing interactions.*”

Table 19. Programs meeting measured aspects of criterion 1B.2.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 1B.2.1	56%	95%	100%	100%
Thorough orientation for new staff including interaction with children and parents, discipline methods, appropriate activities.	72%	98%	100%	100%
In-service training provided regularly by program (Ex. Staff participate in workshops, guest speakers and videos used for on-site training)	74%	98%	100%	100%

1B.3.1. (Level 3) Staff engage children in meaningful conversations, use open-ended questions and provide opportunities throughout the day to scaffold their language to support the development of more complex receptive and expressive language, support children's use of language to share ideas, problem solve, and have positive peer interactions.

For the purposes of the MA QRIS, this criterion is verified using a program's overall ERS score and ERS subscale scores and whether those scores exceed predetermined minimum scores both overall and for individual subscales, as well as overall Arnett CIS scores. Because overall and subscale scores take into account multiple items of quality, it is possible for a program to not meet verification requirements for this particular criterion due to things that are unrelated to the criterion itself, which focuses on the program's overall language and literacy environment. As such, the validation study used scores on ECERS-R Language and Literacy subscale and the ITERS-R Listening and Talking subscale to more directly assess this criterion.

At Levels 3 and 4, all or nearly all programs achieved a score of 4 or above, the MA QRIS requirement for Level 3 for these subscales (Table 20; for Level 4, only preschool classrooms were considered, as only one study program had an infant or toddler classroom). At lower levels, fewer programs attained these scores, or 51 percent of Level 1 programs and 68 percent of those at Level 2. Notably, the language environment in the classroom is also addressed within the curriculum, assessment, and diversity standard at Level 3 (1A.3.4. Staff demonstrate language and literacy skills either in English or the child's language that provide a model for children). As such, this criterion may be redundant. The criterion could be changed to emphasize other aspects of the classroom environment or interactions more directly given that the standard in which this criterion resides relates directly to teacher-child relationships and interactions.

Table 20. Programs meeting measured aspects of criterion 1B.3.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets score of 4 or above for all study classrooms (up to two classrooms per program, one infant/toddler and one preschool)	51%	68%	88%	nr
Preschool classroom received a score of 4 or above on the ECERS-R Language and Reasoning subscale (programs with an observed preschool classroom only)	62%	85%	90%	100%
Infant or toddler classroom received a score of 4 or above on the ITERS-R Listening and Talking subscale (programs with an observed infant or toddler classroom only)	59%	74%	93%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

1B.3.2. (Level 3) Educators are provided with opportunities to use outside consultants or staff with expertise in the age of the children served to assist them in implementing strategies that support positive relationships/interactions and prevention/intervention techniques.

Programs are verified as meeting criterion 1B.3.2 if the program quality specialist can determine that the program offers educators support from internal staff or consultants with certification in behavioral health, mental health, and prevention/intervention techniques — if the support is offered by an outside contractor, an MOU or proof that collaboration is in place must also be provided.²³

As shown in Table 21, 95 percent of Level 3 programs and 83 percent of Level 4 programs indicated that specialists worked directly with educators to assist in planning and implementing intervention or prevention strategies for children, as did 74 percent of Level 1 programs and 83 percent of Level 2 programs. Programs were much less likely to have formal arrangements in place for these services, for example using their own staff or having an MOU in place with the provider. When this is taken into account, the percentage of programs at each level meeting the requirement decreases substantially. In interviews, several programs said that they worked with Early Intervention or the local public school where children were receiving services, but that they were unable to enter into a formal MOU. As mentioned previously, if MOUs continue to be required, programs may need support in being able to develop these, particularly with the public schools. It should also be noted that PQS guidance does not specify what constitutes appropriate certification under the criteria; clarification of MA EEC expectations regarding expertise and certification may be beneficial for programs.

Table 21. Programs meeting measured aspects of criterion 1B.3.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 1B.3.2	21%	50%	80%	50%
Specialists or consultants work directly with educators to assist planning and implementing intervention or prevention strategies for children	74%	83%	95%	83%
Program has a formal agreement or MOU with the person who supports educators and planning or implementing intervention or prevention strategies or that person is a member of their staff	21%	50%	80%	50%

²³ Head Start programs and those accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

1B.4.1. (Level 4) Staff utilize teaching strategies that ensure a positive classroom environment, engage children in learning, and promote critical thinking skills.

Within the MA QRIS, this criterion is verified using a program's overall ERS score and ERS subscale scores. Because overall and subscale scores take into account multiple items of quality, it is possible for a program to not meet verification requirements for this particular criterion due to things that are unrelated to the criterion itself. As such, the validation study used scores on Interaction subscale of the ECERS-R and ITERS-R, which assess the positive classroom environment more directly.

Approximately 60 percent of Level 4 preschool classrooms had Interaction subscale scores of 5 or above, which corresponds to the Good benchmark on the ERS (Table 22; ITERS-R scores are not reported, as only one Level 1 program in the study had an infant/toddler classroom). Interaction subscale scores tended to be higher in infant/toddler classrooms than in preschool classrooms, and as such, at all levels, a larger proportion of infant and toddler classrooms had Interaction scores of 5 or above. The relatively low proportion of study classrooms at Levels 3 and 4 meeting this benchmark suggest that programs may need additional support and training if they are to have and maintain high-quality interaction consistent with expectations for the system's upper tiers.

Table 22. Programs meeting measured aspects of criterion 1B.4.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets score of 5 or above on the Interaction subscale for all study classrooms (up to two classrooms per program, one infant/toddler and one preschool)	23%	23%	46%	nr
Preschool classroom received a score of 5 or above on the ECERS-R Interaction subscale (programs with an observed preschool classroom only)	26%	36%	54%	60%
Infant or toddler classroom received a score of 5 or above on the ITERS-R Interaction subscale (programs with an observed infant or toddler classroom only)	33%	48%	67%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

MA QRIS Quality Standard2: Safe, Healthy Indoor and Outdoor Environments

The QRIS also comprises a series of criteria related to program environments, to ensure that these environments support the implementation of the curriculum and also provide support for the health, safety, and nutrition of young children to ensure their optimum development and well-being. The safe, healthy indoor and outdoor environments standard comprises six criteria, two at Level 2, three at Level 3, and one at Level 4.

2A.2.1. (Level 2) Annual consultation by a Health Consultant to monitor records, update health care policies and practices, identify program issues, assist programs in complying with health and safety requirements, and provide a written report to the program, unless needs of a child require additional consultation.

To be verified as meeting criterion 2A.2.1, programs must provide a signed document from the health consultant or a copy of the collaboration agreement through which a health consultant is provided, as well as a signed statement from the administrator that visits occur on at least an annual basis.²⁴ In interviews conducted as part of the validation study, 85 percent administrators at Level 2 programs and all or nearly all of those at Levels 3 and 4 indicated that their program had a health consultant and met with their health consultant annually (Table 23). For the purposes of the study, health consultants referred to a Massachusetts-licensed physician, registered nurse, nurse practitioner, or physician's assistant with pediatric or family health training who works with the

²⁴ Programs accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

program around health care policy and health related issues. It did not refer to a QRIS health consultant specifically. Health consultants were less common at Level 1: 77 percent of programs at Level 1 had a health consultant, and 64 percent met annually with their health consultant.

Table 23. Programs meeting measured aspects of criterion 2A.2.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Meets all measured aspects of 2A.2.1	64%	85%	95%	100%
Program has a health consultant	77%	95%	100%	100%
Meet with the health consultant at least annually	64%	85%	95%	100%

2A.2.2., 2A.3.3, and 2A.4.1. (Levels 2, 3, 4) Demonstrates healthy, safe, and clean indoor and outdoor environments.

For the purposes of MA QRIS, criterion 2A.2.2 is verified using a program's ERS scores and whether those scores met EEC thresholds both overall and for individual subscales. The validation study used scores on the Personal Care Routines Subscale of the ECERS-R and ITERS-R, as these assessed a program's health practices more directly. Programs were considered to be meeting the standard if their score for the Personal Care Routines subscale was 2 or higher, which is the threshold set by MA EEC for Personal Care Routines at Level 2. This criterion is part of series within the safe, healthy indoor and outdoor standard that requires progressively higher ERS scores. It includes criterion 2A.3.3 and parts of 2A.4.1, which are measured similarly for the MA QRIS, with increasing score requirements as the level increases.

Overall most classrooms had scores that met the Level 2 score threshold for Personal Care Routines, which falls between the Inadequate and Minimal benchmarks on the ERS (Table 24, see ERS quality section for a more detailed discussion). The proportion of programs meeting higher benchmarks such as a 3 (currently required at Level 3) or a 4 (currently required at Level 4) decreases substantially (Table 25). As such, criteria related to health practices are likely to reflect a substantial barrier to progression in the MA QRIS without significant support and assistance for programs.

At Level 3, 63 percent of observed preschool classrooms were meeting this benchmark with a score of 3 or above, as were 20 percent of observed infant and toddler classrooms. Overall, slightly less than half of Level 3 programs met benchmarks in all observed classrooms — up to two classrooms, one preschool and one infant or toddler classroom. As discussed elsewhere in this report, personal care and health practices represent a significant potential growth area, particularly for infants and toddler, even for highly rated programs. Among programs at lower levels, significantly fewer were meeting these benchmarks, 21 percent at Level 1 and 18 percent at Level 2. These data suggest that programs are likely to need support if they are to improve their practice so they can progress to higher levels, or maintain their practice within the higher levels.

Using validation study measures, very few programs met the criteria for 2A.4.1. Across all levels, the proportion observed preschool classrooms receiving a score of 4 or above ranged from 0 percent to 20 percent; none of the observed Level 4 classrooms had a score of 4 or above. Sinks tended to be common, particularly in infant and toddler rooms and preschool classrooms at Level 2 or above. In fact, only one program, a preschool classroom, met the ERS requirement for the Personal Care Routines but did not have a sink in the classroom; all the infant and toddler classrooms that met the requirement had sinks in the classroom. As such, the requirement regarding sinks, in and of itself, does not appear to be contributing substantively to whether or not programs are meeting this criterion. As mentioned previously, the data suggest that programs are likely to need support in Personal Care Routines and related health practices if they are to progress to higher levels and, for higher level programs, to maintain their practice over time.

Table 24. Programs meeting measured aspects of criterion 2A.2.2 (Level 2), by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets score of 2 or above on the Personal Care Routines subscale for all study classrooms (up to two classrooms per program, one infant/toddler and one preschool)	87%	73%	85%	nr
Preschool classroom received a score of 2 or above on the ECERS-R Personal Care Routines subscale (programs with an observed preschool classroom only)	92%	87%	98%	100%
Infant or toddler classroom received a score of 2 or above on the ITERS-R Personal Care Routines subscale (programs with an observed infant or toddler classroom only)	89%	81%	67%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

Table 25. Programs meeting higher level requirements (2A.3.3, Level 3 and 2A.4.1, Level 4) for Personal Care Routines, by MA QRIS level

	Level 1	Level 2	Level 3	Level 4
Program meets score of 3 or above on the Personal Care Routines subscale for all study classrooms (criterion; Level 3)	21%	18%	46%	nr
Preschool classroom received a score of 3 or above on the ECERS-R Personal Care Routines subscale (programs with an observed preschool classroom only)	28%	31%	63%	60%
Infant or toddler classroom received a score of 3 or above on the ITERS-R Personal Care Routines subscale (programs with an observed infant or toddler classroom only)	22%	26%	20%	nr
Program meets score of 4 or above on the Personal Care Routines subscale for all study classrooms (criterion; Level 4)	3%	3%	15%	nr
Preschool classroom received a score of 4 or above on the ECERS-R Personal Care Routines subscale (programs with an observed preschool classroom only)	5%	5%	20%	0%
Infant or toddler classroom received a score of 4 or above on the ITERS-R Personal Care Routines subscale (programs with an observed infant or toddler classroom only)	7%	3%	13%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

2A.3.1. (Level 3) Program, with parental consent, provides (directly or through collaboration) vision, hearing, and dental screenings, and shares results with families.

To be verified as meeting criterion 2A.3.1, programs must provide an example document that tracks when all children have vision, hearing, and dental screenings, and verify that results are shared with families. If the service is offered by an outside contractor, the program must also provide an MOU or other proof that collaboration is in place. Overall, on program questionnaires, all or nearly all programs at Level 3 and 4 indicated that they were tracking whether children had received hearing and vision screenings, compared with less than half of programs at Levels 1 and 2 (Table 26). Similarly, most programs in the higher two levels — 90 percent at Level 3 and 83 percent at Level 4 — were tracking dental screenings, compared with 43 percent at Level 2 and 26 percent at Level 1.

Table 26. Programs meeting measured aspects of criterion 2A.3.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 2A.3.1	23%	38%	90%	83%
Tracks whether children have received hearing screenings	41%	48%	93%	100%
Tracks whether children have received vision screenings	41%	48%	93%	100%
Tracks whether children have received dental screenings	26%	43%	90%	83%

In Levels 3 and 4, about half of programs indicated that they provided or had an agreement with a provider for all three of these screenings, compared with one in five programs at Level 2 and very few at Level 1 (Table 27). Programs are only expected to provide or have an agreement for these services if they are not accessed directly by parents, in which cases the expectation within the MA QRIS may vary depending on the clientele served.

As can be seen from the data, although programs were tracking screenings and children's need for medical screenings, significantly less had a formal agreement in place for screenings and services.

Table 27. Programs that provide or have an agreement in place for screenings, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Provide or have an agreement in place for all screenings referenced in criterion 2A.3.1	3%	20%	46%	50%
Tracks whether children have received hearing screenings	8%	25%	56%	50%
Tracks whether children have received vision screenings	3%	28%	56%	50%
Tracks whether children have received dental screenings	5%	30%	59%	83%

2A.3.2. (Level 3) Staff are trained in how to work with children with special diets, allergies, and specialized feeding issues.

Criterion 2A.3.2 requires that all educators receive formal professional development (i.e. CEUs or credits) in how to work with children with special diets, allergies, and specialized feeding issues. Data from the program questionnaire used for the validation study show that all programs at Level 4 met this requirement, as did 76 percent of those at Level 3 (Table 28). Among those at Level 3 that did not meet this requirement, this was largely attributable to the training having been completed by some but not by all staff. Among Level 1 and 2 programs, the proportions meeting this Level 3 requirement were 41 percent and 65 percent, respectively. As with other criteria related to training, the validation study was not able to gather reliable data on whether training received CEUs or credits, so this aspect of the criteria is not included in the study measure. As mentioned previously, if CEUs or credits remain important to MA QRIS verification, MA EEC may find it beneficial to incorporate language regarding this requirement into the criterion itself to clarify the requirement and emphasize its importance to the field.

Table 28. Programs meeting measured aspects of criterion 2A.3.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 2A.3.2	41%	65%	76%	100%
Program questionnaire: Staff have received this training	77%	85%	93%	100%
Program questionnaire: <u>All</u> staff have received this training	41%	65%	76%	100%

MA QRIS Quality Standard 3A: Designated Program Administrator Qualifications and Professional Development

The Massachusetts QRIS also includes a series of requirements related to administrator qualifications and professional development. This standard comprises six criteria at Level 2, four criteria at Level 3, and four criteria at Level 4, or a total of 14 criteria across all levels.

3A.2.1. (Level 2) Program administrator has a Child Development Associate (CDA) Credential for the age of the children served, (or higher i.e. Associate's or Bachelor's degree) OR is enrolled in a program leading to an Associate or Bachelor's Degree in early childhood education or related field.

To meet criterion 3A.2.1, a program’s administrator must have a CDA or higher or have completed college coursework within the past six months. Data collected from administrators as part of the validation study show that most programs met this requirement, regardless of QRIS level, ranging from 90 percent at Level 1 to 100 percent at Level 4 (Table 29). At Levels 2 and above, more than 80 percent of programs met this requirement because their administrator has a Bachelor’s degree in an ECE-related field, whereas, at Level 1, half of administrators had a Bachelor’s degree, with a relatively large proportion of programs (28 percent) meeting the requirement through their administrator having an Associate’s degree. Regardless of level, relatively few programs met this requirement as a result of their administrator being enrolled in a degree program in an ECE-related field (between 0 percent and 5 percent, depending on level; Table 30). Given the relatively limited variation for criterion 3A.2.1, this criterion may be one that could be dropped from the QRIS.

Table 29. Programs meeting measured aspects of criterion 3A.2.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
3A.2.1. Program administrator has a Child Development Associate (CDA) Credential for the age of the children served (or higher i.e. Associate's or Bachelor's degree), OR is enrolled in a program leading to an Associate or Bachelor's Degree in early childhood education or related field.	90%	98%	95%	100%

Table 30. Reason program meets criterion 3A.2.1

	Level 1	Level 2	Level 3	Level 4
Administrator has a bachelor’s degree or higher in an ECE-related field	54%	83%	88%	100%
Administrator has an associate’s degree in an ECE-related field	28%	13%	5%	0%
Administrator has a CDA	3%	0%	0%	0%
Administrator is enrolled in a degree program in an ECE-related field	5%	3%	2%	0%
Does not meet criterion 3A.2.1	10%	2%	5%	0%

3A.2.2. (Level 2) Program administrators with primary responsibility for the supervision of educators in the program, has a BA degree.

Criterion 3A.2.2 requires that the administrator with primary responsibility for the supervision of staff has at least a Bachelor’s degree in an ECE-related field. Data from administrator questionnaires indicate that all programs at Level 4 met this requirement, as do most at Levels 2 and 3 (Table 31). However, only about 54 percent of Level 1 programs met this requirement, suggesting that a bachelor’s degree requirement at Level 2 could pose a barrier to program’s progression in the system. Notably this requirement is identical to that of criterion 3A.3.1, which also requires a bachelor’s degree, but at Level 2. One possible approach to eliminating this redundancy would be to amend criterion 3A.2.1 to require an associate’s degree at Level 2 — which 98 percent of current administrators at Level 2 programs have (see Table 30 above) — eliminate criterion 3A.2.2,

and retain criterion 3A.3.1. This would create a sequence within QRIS whereby an associate's degree is required at Level 2 and a bachelor's degree is required at Level 3, resulting in an interim step between Level 1 and Level 3. This might be more manageable for programs at Level 1 who have administrators without BA degrees.

Table 31. Programs meeting measured aspects of criterion 3A.2.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 3A.2.2	54%	83%	88%	100%
Administrator has a bachelor's degree	54%	83%	93%	100%
Administrator's bachelor's degree is in a related field	54%	83%	88%	100%

3A.2.3. (Level 2) Program administrator has a minimum of 3 college credits in administration and management and 12 college credits in early childhood education/child development/special education and 2 years' experience as an administrator.

Programs meet criterion 3A.2.3 if administrators demonstrate that they have the necessary credits for this criterion or are Director I or Director II certified — in which case they would have obtained the required credits — and the administrator has a least two years of experience as an administrator. Overall, regardless of level, data show that all or nearly all administrators had the required credits in early education and care or a related field, ranging from 98 percent to 100 percent (Table 32). Similarly, the vast majority of administrators had the requisite credits in administration and management (between 83 percent and 100 percent, depending on level) and at least two years of experience (between 93 percent and 100 percent, depending on level). Across all levels of the QRIS, program administrators had an average of at least 12 years of experience, ranging from 12.5 years at Level 2 to 16.2 years at Level 3 (Table 33).

As a result of the high percentage of programs meeting each individual requirement, there was little variation in the overage percentage of programs meeting all three requirements contained in this criterion, suggesting that this might be a criterion that could be dropped from the system or otherwise modified, as it is not distinguishing quality. Further, it is possible that the requirement regarding experience could be at odds with MA EEC's emphasis on the development of a career ladder for early education and care providers, as it could deter some programs from promoting otherwise qualified staff to administrative positions due to concerns that the promotion could negatively affect the program's QRIS level.

Table 32. Programs meeting measured aspects of criterion 3A.2.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 3A.2.3	87%	90%	90%	83%
Administrator has at least four courses in early childhood education, child development, early childhood special education, and/or special education.	100%	100%	98%	100%
Administrator has at least one course in administration and management.	95%	95%	100%	83%
Administrator has at least two years of experience.	92%	95%	93%	100%

Table 33. Average years of experience of administrator, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Average years of administrative experience for program administrators	14.4	12.5	15.1	16.2

3A.2.4. (Level 2) Program administrator is trained in the MA Guidelines for Preschool Learning Experiences, the Infant / Toddler Learning Guidelines, child development, the Strengthening Families protective factors, and is knowledgeable about the core competencies in order to be able to develop their staff's professional development plans.

To be verified for criterion 3A.2.4, a program's administrator must have completed formal professional development, as documented in the PQR, in the Massachusetts Guidelines for Preschool Learning Experiences, the Massachusetts Guidelines for Infants and Toddlers, child development, and the Strengthening Families Protective Factors. Although the MA QRIS requires CEUs or credits in order for training to be considered as meeting the standard, the validation study was not able to collect reliable data on CEUs, and as such, this is not included in study measures of the standard. As with criteria related to educator training, it is recommended that MA EEC consider incorporating language regarding the requirement for CEUs or credits into the language of the criteria itself to emphasize the importance of formal training for QRIS purposes. The aspect of the criterion related to administrators' knowledge about the core competencies is not easily measurable, and as such, is not currently addressed in verification guidance. MA EEC might want to consider removing this language from the criterion itself to emphasize the measurable aspects of the criterion that are currently being assessed as part of verification procedures.

As Table 34 displays, although most administrators had received training in each of these areas individually, fewer had completed training in all four areas, or 54 percent at Level 1, 60 percent at Level 2, 85 percent at Level 3, and 83 percent at Level 4. One possible approach to emphasize the importance of training in all of the areas would be to divide the criterion into multiple criteria, one for training in the relevant guidelines and one for training in the Strengthening Families Protective Factors. Notably, child development training was nearly universally received among administrators (it is a requirement for Director I certification), and as such, this aspect of the criterion does not appear to be distinguishing quality, and could be removed.

Table 34. Programs meeting measured aspects of criterion 3A.2.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 3A.2.4	54%	60%	85%	83%
Administrator has had training in relevant guidelines for all ages	74%	75%	88%	100%
Administrator has had training in the PS guidelines (programs with preschool-aged students only)	79%	87%	90%	100%
Administrator has had training in the infant and toddler guidelines (programs with infants and toddlers only)	79%	77%	81%	100%
Administrator has had training in the child development	100%	100%	98%	100%
Administrator has had training in the Strengthening Families Protective Factors	64%	78%	93%	83%

3A.2.5. (Level 2) Program administrator has received professional development in supervision of adults and strategies for working with adults.

To be verified as meeting criterion 3A.2.5, a program's administrator must have formal professional development in the supervision of adults and strategies for working with adults documented in the PQR. Again,

to qualify for the MA QRIS, training must either have received college credit or CEUs, although the validation study was unable to collect reliable data regarding CEUs, and as result, this aspect of the criterion was not able to be measured. MA EEC may want to consider making the requirement for CEUs or credits explicit in the criterion language itself to emphasize this aspect of the MA QRIS to the field.

Overall, administrators at nearly all programs at Levels 3 and 4 indicated that they had received professional development in the supervision of adults and strategies for working with adults, as did most at Level 2 (Table 35). At Level 1, 79 percent of administrators said they had received this training, or a similar proportion to those at Level 2.

Table 35. Programs meeting measured aspects of criterion 3A.2.5, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program administrator has received professional development in supervision of adults and strategies for working with adults	79%	85%	98%	100%

3A.2.6. (Level 2) Program administrator has an Individual Professional Development Plan (IPDP) that addresses their process and timelines to achieve the Program Administrator Qualifications for the next level of the QRIS and increases their competency along the advanced continuum of the core competencies.

At Level 2, program administrators are required to have an IPDP that is created, updated, and reviewed annually.²⁵ The validation study gathered data on administrators’ IPDPs, but did not collect data on the frequency with which IPDPs were created and reviewed. Notably, criterion 3A.2.6 is one of three criteria within the designated workforce qualifications and professional development criterion, all of which are verified in an identical manner (the other criteria are 3A.3.4 and 3A.4.4).

Overall, all administrators at Levels 3 and 4 indicated that they had an IPDP, and nearly all at Levels 2 (85 percent) indicated that they had an IPDP (Table 36). This practice was far less common at Level 1, where only half of administrators said they had an IPDP. Notably, 25 percent of administrators at Level 1 programs and 15 percent of administrators at Level 2 programs indicated that their IPDP did not include a timeline for meeting any of the unmet requirements related to progression in QRIS (training, degree, or college credits). This suggests that IPDPs may be of variable quality, and programs at lower levels may benefit from sample IPDP documents or further instructions related to EEC expectations regarding IPDPs.

Table 36. Programs meeting measured aspects of criterion 3A.2.6, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program administrator has an IPDP	51%	85%	100%	100%

The aspect of the criterion indicating that IPDPs be related to “increasing their (administrator) competency along the advanced continuum of the core competencies” is not easily measurable and, as such, is not currently addressed in verification guidance. MA EEC may want this to consider removing this language from the criterion and stress more measurable outcomes such as progress toward the next MA QRIS level and/or specific professional development related to core competencies.

²⁵ Head Start programs and those accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

3A.3.1. (Level 3) Program administrator has at least a Bachelor's degree.

To be verified for criterion 3A.3.1, a program’s administrator is required to have a bachelor’s degree in an ECE-related field. As mentioned previously, this requirement is the same as that contained within 3A.2.2. As depicted in Table 37, all programs at Level 4 met this requirement, as do most at Levels 2 and 3. In all or nearly all cases in which an administrator had a degree, the degree was in a related field. However, only about 54 percent of administrators at Level 1 programs have a degree, suggesting that a bachelor’s degree requirement at Level 2 could pose a substantial barrier to progression in the system, and as such, it may be beneficial to retain 3A.3.1 as is and eliminate or modify the Level 2 requirement, 3A.2.2.

Table 37. Programs meeting measured aspects of criterion 3A.3.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured components of criterion 3A.2.1.	54%	83%	88%	100%
Administrator has a BA degree	54%	83%	93%	100%
Administrator's BA is in a related field	54%	83%	88%	100%

3A.3.2. (Level 3) Program administrator has at least 9 credit-bearing hours of specialized college-level course work in administration, leadership, and management.

In consistent with the how the criterion is written, a program is verified as meeting criterion 3A.3.2 if the program administrator has completed 6 credit-bearing hours of specialized college-level course work in administration, leadership, and management. This criterion also appears at Level 4 as criterion 3A.4.1, which under current verification guidance is verified in the same manner. Overall, 67 percent of administrators at Level 1 programs, 73 percent at Level 2 programs, 71 percent at Level 3 programs, and 83 percent at Level 4 programs reported having at least two college courses in administration and management (Table 38).

Table 38. Programs meeting measured aspects of criterion 3A.3.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Administrator has at least two courses in administration and management.	67%	73%	71%	83%

To some extent, the relatively low proportion of administrators with two courses in administration and management may be explained by the fact that relatively few early education and care educator preparation programs in Massachusetts offer more than one course in these areas. If this remains an important requirement for Level 3, it may be necessary to develop strategies to promote access to existing courses or work with educator preparation programs to develop additional courses in administration and management. An alternative might be to change the criterion to require Director II certification, which includes additional coursework in any of the following areas: policy management, supervision, families and communities, child care policy, health and safety in early childhood programs, or children with disabilities, as a way to build skills and credentials and further align QRIS with certification requirements.

3A.3.3. (Level 3) Program administrator has at least 24 credit-bearing hours of specialized college-level course work in early childhood education, child development, elementary education OR Documents that a plan is in place to meet the above qualifications within five years.

To meet criterion 3A.3.3, verification indicates that program must demonstrate that the administrator has completed at least 18 (different from criterion) credit-bearing hours of specialized coursework in early education, child development, elementary education, or early childhood special education, or provides a copy of an IPDP outlining a plan for the educator to attain these credits. For the validation study, attainment of this

requirement was measured by the number of courses administrators reported in these areas or whether the administrator indicated that he or she was actively enrolled in a degree program in an ECE related field. As exhibited in Table 39, all or nearly all administrators met this requirement, regardless of QRIS level.

To some extent, the high percentage of administrators with the required coursework for criterion 3A.3.3 may reflect the fact that most administrators — particularly those at Level 2 programs and above — have at least an associate’s degree in early education and care or a related field. Since a degree in an ECE-related field is required as part of QRIS at Level 3, these specific requirement regarding credits may not be necessary, as nearly all administrators with a bachelor’s degree in an ECE-related field (97 percent) reported having the requisite number of credits to meet criterion 3A.3.3. As such, not only are there inconsistencies between the criteria and verification for this requirement, , this Level 3 criteria also appears to be redundant with the MA QRIS Level 3 educational degree requirements for program administrators (see Table 40).

Table 39. Programs meeting measured aspects of criterion 3A.3.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets criterion 3A.3.3	90%	100%	95%	100%

Table 40. Reason program meets criterion 3A.3.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Administrator has at least 6 courses in early education or related field	90%	100%	93%	100%
Administrator is enrolled in a degree program	0%	0%	2%	0%
Program does not meet 3A.3.3	10%	0%	5%	0%

3A.3.4. (Level 3) Program administrator's IPDP addresses their process and timelines to achieve the Program Administrator Qualifications for the next level of the QRIS and increases their competency along the advanced continuum of the core competencies.

This criterion is verified for programs whose administrators have an IPDP that is created, updated, and reviewed annually, which is identical to the verification requirement for criterion 3A.2.6 (Level 2) and criterion 3A.4.4 (Level 4).^{26,27} Because programs must already demonstrate that they meet this requirement to achieve Level 2, this criterion is redundant, and MAEEC may want to consider removing it to streamline QRIS or modifying it to differentiate it from the lower level requirement (Table 41).

Table 41. Programs meeting measured aspects of criterion 3A.3.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program administrator has an IPDP	51%	85%	100%	100%

²⁶ Programs accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

²⁷ The validation study gathered data on administrators' IPDPs, but did not collect data on the frequency with which IPDPs were created and reviewed.

3A.4.1. (Level 4) Program administrator has at least 9 credit-bearing hours of specialized college-level course work in administration, leadership, and management.

Inconsistent with how the criterion is recorded; a program is verified as meeting criterion 3A.4.1 if the program administrator has completed 6 credit-bearing hours of specialized college-level course work in administration, leadership, and management. This is identical to the requirements for criterion 3A.3.2. As such, it is recommended that 3A.4.1 be removed, as it reflects a redundant requirement given that programs at Level 4 must already demonstrate that they meet this requirement to achieve Level 3. Alternatively, this criterion could be modified to differentiate it from the lower level requirement (see Table 42).

Table 42. Programs meeting measured aspects of criterion 3A.4.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Administrator has completed two courses in administration and management	67%	73%	71%	83%

3A.4.2. (Level 4) Program administrator has at least 24 credit-bearing hours of specialized college-level course work in early childhood education, child development, elementary education, or early childhood special education.

To meet criterion 3A.4.2, a program must demonstrate that its administrator has completed at least 18 credit-bearing hours of specialized coursework in early education, child development, elementary education, or early childhood special education which is incongruous with the criterion. Additionally, this criterion is similar to the requirements of criterion 3A.3.3, although unlike criterion 3A.3.3, criterion 3A.4.2 does not allow a program to meet the requirement by having a plan in place for its administrator to meet this requirement within five years.

As Table 43 shows, all or nearly all administrators met this requirement, regardless of QRIS level. It is unclear from the data however, if these hours can be applied towards a degree. When considering the MA QRIS as a mechanism for supporting the career ladder for educators and the professionalization of the field, an emphasis on degrees and obtaining degree-related credits may be important focus at the higher levels of MA QRIS. It is important to note that most research on the importance of education in the field of early care and education has focused primarily on degrees. As such, credit thresholds which may not represent the same quality or consistency in coursework as degree related programs may not produce the same impact as a focused degree program of study.

Table 43. Programs meeting measured aspects of criterion 3A.4.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program administrator has completed at least 6 courses in ECE, child development, elementary education, or early childhood special education	90%	100%	93%	100%

3A.4.3. (Level 4) Program administrator has a minimum of 5 years' experience as an administrator.

At Level 4, programs are required to have an administrator with at least five years of experience working as the primary supervisor to educators. This reflects an increase relative to the Level 2 requirement of two years of experience (criterion 3A.2.3). Regardless of level, between 83 percent and 85 percent of administrators at study programs (Table 44) reported having at least five years of experience required as part of criterion 3A.4.3. Because there are no notable differences across level, this criterion does not appear to be distinguishing quality and may be worth being considered for removal or modification. As mentioned previously, it is also possible that requirements regarding experience could be at odds with MA EEC's emphasis on the development of a career ladder for early education and care providers, as it could deter some programs from promoting otherwise

qualified staff to administrative positions due to concerns that the promotion could negatively affect the program's MA QRIS level or chances of progression within the MA QRIS.

Table 44. Programs meeting measured aspects of criterion 3A.4.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program administrator has 5 or more years of experience as an administrator of an early education and care program	85%	83%	85%	83%

3A.4.4. (Level 4) Program administrator has a Professional Development Plan that addresses areas of identified needs and increases their competency along the advanced continuum of the core competencies.

Criterion 3A.4.4 is verified for programs whose administrators have an IPDP that is created, updated, and reviewed annually, which is identical to the verification requirement for 3A.2.6 (Level 2) and 3A.3.4 (Level 3).^{28,29} As such, it is recommended that 3A.4.4 be removed, as it reflects a redundant requirement given that programs at Level 3 must already demonstrate that they meet this requirement to attain Level 2. Alternatively, this criterion could be modified to differentiate it from the lower-level requirement (see Table 45 for MA QRIS Level data).

Table 45. Programs meeting measured aspects of criterion 3A.4.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program administrator has an IPDP	51%	85%	100%	100%

MA QRIS Quality Standard 3B: Program Staff Qualifications and Professional Development

As noted in the program quality specialist guidance, research indicates that the workforce engaged in early childhood education must have formal professional development in early childhood education and content knowledge in order to support program quality and impact child outcomes. To this end, the MA QRIS also contains a series of criteria related to educator qualifications and professional development. This includes a total nine criteria, four at Level 2, two at Level 3, and three at Level 4.

3B.2.1. (Level 2) All staff working in program have a high school diploma or GED.

To reach Level 2, all educators considered in licensing ratios must have at least a high school diploma, a GED, or a CDA. Validation study shows that consistent with this, all or nearly all programs at Levels 2 and above (93 percent to 100 percent) met this requirement (Table 46). Fewer programs at Level 1 (77 percent) met this requirement. In all but a couple of instances, when a program did not meet this requirement, it was typically attributable to one educator who did not have a high school diploma, GED, or CDA.

²⁸ The validation study gathered data on administrators' IPDPs, but did not collect data on the frequency with which IPDPs were created and reviewed.

²⁹ Programs accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

Table 46. Programs meeting measured aspects of criterion 3B.2.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
All educators included in licensing ratios have a high school diploma, GED, or CDA	77%	93%	98%	100%

3B.2.2. (Level 2) 50 percent of classrooms have educator(s) with a Bachelor's degree or higher who work for the full program day.

At Level 2, programs are required to have at least 50 percent of classrooms with an educator with a bachelor's degree. To be verified for this criterion, the PQR must indicate at least one educator with a degree for every two classrooms (for example, if a program has four classroom, PQS guidance indicates that the program must have at least two educators with a bachelor's degree). This criterion is part of series of requirements within the program staff qualifications and professional development standard, with requirements increasing as the levels increase. At Level 3, programs are required to have a sufficient number of degreed educators for 75 percent of classrooms (criterion 3B.3.1), and at Level 4, programs are required to have sufficient number of degreed educators for 100 percent of classrooms (criterion 3B.4.1).

Validation study data show that about two-thirds of Level 2 programs reported having a sufficient number of degreed educators to meet the Level 2 requirement, likely the result of what has been described as high exemption rate for criterion in this sequence (Table 47). These data also suggest that this criterion may function as a barrier to programs' progression within QRIS, as slightly less than half of Level 1 programs in the study would meet the requirement without an exemption. With regard to the requirements for higher levels, 83 percent of programs at Levels 3 and 4 reported sufficient degreed staff to meet the Level 3 requirement, and 83 percent of Level 4 programs reported sufficient degreed staff to meet the Level 4 requirement, suggesting lower exemption rates at the higher levels (Table 48). This may be, in part, attributable to the prevalence of Head Start programs in these two higher levels.

Table 47. Programs meeting measured aspects of criterion 3B.2.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program has sufficient educators with a degree in an ECE-related field for 50% of classrooms	49%	65%	90%	83%

Table 48. Programs meeting measured aspects of related criteria, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program has sufficient educators with a degree in an ECE-related field for 75% of classrooms (criterion 3B.3.1)	41%	40%	83%	83%
Program has sufficient educators with a degree in an ECE-related field for all classrooms (criterion 3B.4.1)	28%	23%	73%	83%

3B.2.3. (Level 2) All educators (lead teachers, teachers, teaching aides, etc.) have a minimum of 3 college credits in early childhood education, or related field.

Criterion 3B.2.3 requires that all educators have at least three college credits in early childhood education or a related field, which is a requirement for educator certification and licensing. As such, criterion 3B.2.3 is automatically verified for all center-based programs that are licensed and have a license in good standing. The validation study was unable to gather reliable data on individual educators' credit attainment. Although a staff

roster worksheet asked programs to provide information about college coursework for their educators, complete credit information was only obtainable for about 35 percent of educators listed on the staff roster. As such, data are not reported for criteria related to educators' credits (which includes criterion 3B.2.3 and criterion 3B.4.2).

3B.2.4. (Level 2) All educators have an IPDP that is developed in conjunction with the supervisor that addresses the identified professional development needs of that educator and the development of their competency along the initial-level of the continuum of the core competencies. The IPDP must also address the actions and timelines that need to be met in order to move to the next level of the QRIS.

Another series of criteria within the program staff qualifications and professional development standard pertains to educators' individual professional development plans. To be verified as meeting criterion 3A.2.4, a program must submit a document stating that an IPDP has been created for all educators, that these IPDPs are reviewed annually, and that these IPDPs are updated annually.³⁰ This criterion is part of series of criteria included in the program staff qualifications and professional development standard, all of which have the same verification guidance and, as such, appear to be functionally identical.

As shown in Table 49, most programs at Level 2 (83 percent) and all programs at Levels 3 and 4 reported having IPDPs for all teachers, including assistant teachers if the program has them.³¹ At Level 1, slightly more than half of programs (56 percent) were creating IPDPs for all teachers, including assistant teachers if the program had them. At Levels 1 and 2, programs were less likely to be creating IPDPs for assistant teachers than they were for teachers. Notably, programs at Level 2 and above were much more likely to incorporate QRIS requirements in lead teacher's and teachers IPDPs. Approximately 92 percent of Level 2 programs, 88 percent of Level 3 programs, and 100 percent of Level 4 programs incorporated QRIS requirements in staff IPDPs, compared with 54 percent of Level 1 programs. This may suggest that programs at lower levels may benefit from sample IPDP documents to better help lower level programs connect IPDPs to QRIS requirements. It also suggests the need for the criteria to clarify that IPDP are required for all educators, including assistant teacher and aids.

Table 49. Programs meeting measured aspects of criterion 3B.2.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 3B.2.4	56%	83%	100%	100%
Program develops IPDPs for lead teachers and teachers	67%	90%	100%	100%
Program develops IPDPs for assistant teachers (only reported for programs that have assistant teachers)	53%	82%	100%	100%

3B.3.1. (Level 3) 75 percent of classrooms have educator(s) with a Bachelor's degree or higher who work for the full program day.

To be verified for this criterion, a program must have a sufficient number of educators with a bachelor's degree for 75 percent of its classrooms (for example, if a program has five classroom, PQS guidance indicates that the program must have at least four educators with a bachelor's degree). At Level 2, programs are required to have a sufficient number of degreed educators for 50 percent of classrooms (criterion 3B.2.2), and at Level 4 programs are required to have sufficient number of degreed educators for 100 percent of classrooms (criterion 3B.4.1).

Validation study data show that 83 percent of programs at Level 3 and above met criterion 3B.3.1 (Table 50). As mentioned previously (see Table 37 and discussion on page 68-69), the Level 3 criterion related to educator

³⁰ Head Start programs and those accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

³¹ The validation study gathered data on administrators' IPDPs, but did not collect data on the frequency with which IPDPs were created and reviewed.

degree status appears to have a lower exemption rate than the Level 2 criterion, which may be, in part, attributable to the prevalence of Head Start programs in higher levels. At Levels 1 and 2, about 40 percent of programs could meet the Level 3 requirement, suggesting that criterion 3B.3.1 could pose a substantial barrier to progression for many programs at lower levels.

Table 50. Programs meeting measured aspects of criterion 3B.3.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program has sufficient educators with a bachelor’s degree in an EEC-related field for 75% of classrooms	41%	40%	83%	83%

3B.3.2. (Level 3) IPDP ensures that the teacher receives formal professional development in the components of the assessment process including screening, observation, use of assessment tools, and IDEA processes.

Criterion 3B.3.2 is verified for programs that complete IPDPs for all educators annually, which is equivalent to the verification requirements for 3B.2.4 (Level 2) and 3B.4.3 (Level 4).^{32, 33} As such, as currently verified, this criterion appears to reflect a redundant requirement, as programs must already demonstrate that they meet this requirement to attain Level 2. Additionally, the language of the criterion, itself, references that the IPDP should ensure that teachers receive formal professional development in screening, observation, and the use of assessment tools; these requirements are redundant with Level 2 training criterion within the Curriculum, Assessment, and Diversity Standard (criterion 1A.2.1). Therefore, it is recommended that this criterion be removed or modified to differentiate it from other lower-level requirements (see Table 51).

Table 51. Programs meeting measured aspects of criterion 3B.3.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 3B.3.2	56%	83%	100%	100%
Program develops IPDPs for lead teachers and teachers	67%	90%	100%	100%
Program develops IPDPs for assistant teachers (only reported for programs that have assistant teachers)	53%	82%	100%	100%

3B.4.1. (Level 4) All (100 percent) of the classrooms have educator(s) with a Bachelor’s degree or higher who work for the full program day.

To be verified for this criterion, a program must have a sufficient number of educators with a bachelor’s degree for 100 percent of its classrooms (for example, if a program has five classroom, PQS guidance indicates that the program must have at least five educators with a bachelor’s degree). At Level 2, programs are required to have a sufficient number of degreed educators for 50 percent of classrooms (criterion 3B.2.2), and at Level 3 programs are required to have sufficient number of degreed educators for 75 percent of classrooms (criterion 3B.3.1).

Validation study data show that 83 percent of programs at Level 4 have a sufficient number of educators with a bachelor’s degree for 100 percent of classrooms, as do 73 percent of programs at Level 3 (Table 52). Few programs at Level 1 and Level 2, 28 percent and 23 percent, respectively, could meet this requirement. As such, in the short term, it is unlikely that criterion 3B.4.1 is functioning to inhibit progression, given most programs already at Level 3 could meet the Level 4 requirement, over the longer term, the requirement is likely to be one that programs in lower levels will need support if they are to eventually progress to the system’s highest tiers.

³² Programs accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

³³ The validation study gathered data on administrators’ IPDPs, but did not collect data on the frequency with which IPDPs were created and reviewed.

Table 52. Programs meeting measured aspects of criterion 3B.4.1 by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program has sufficient educators with a degree in an EEC-related field for all classrooms	28%	23%	73%	83%

3B.4.2. (Level 4) Educators have a minimum of 30 college credits in early childhood education/ child development/special education.

Criterion 3B.4.2 requires that at least one educator in each classroom has received 30 college credits in early childhood education. For example, PQS guidance states that if a program has four classrooms, the PQR should indicate that at least four educators have received 30 college credits in early childhood education. Although a staff roster worksheet asked programs to provide information about college coursework for their educators, complete credit information was only obtainable for about 35 percent of educators listed on the staff roster. As such, data are not reported for criteria related to educators’ credits (which includes criterion 3B.2.3 and criterion 3B.4.2).

3B.4.3. (Level 4) IPDP reflects identified opportunities for mentoring/coaching/ supervision available to educators, etc.

Criterion 3B.4.3 is verified for programs whose administrators submit a document indicating that IPDPs are completed regularly, which is equivalent to the verification requirement for 3B.2.4 (Level 2) and 3B.3.2 (Level 3).³⁴ Therefore, it is recommended that this criterion be removed or its verification procedures be otherwise modified to differentiate it from the lower level requirement. Notably the criterion language references mentoring, coaching, and supervision, which could potentially be incorporated into an IPDP template for Level 4 programs (see Table 53 for data).

Table 53. Programs meeting measured aspects of criterion 3B.4.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 3B.4.3	56%	83%	100%	100%
Program develops IPDPs for lead teachers and teachers	67%	90%	100%	100%
Program develops IPDPs for assistant teachers (only reported for programs that have assistant teachers)	53%	82%	100%	100%

MA QRIS Quality Standard 4: Family and Community Engagement

Programs at higher levels of the QRIS are also expected to engage with families and their communities to share information and provide resources and support to help children thrive in both the program and at home. As the program quality specialist guidance document notes, high quality programs acknowledge the “interconnectedness between the family and a child’s approach to learning and establish a relationship with families that is built on mutual trust, respect and a willingness to involve them as full partners.” This standard comprises a total of 13 criteria, which includes five at Level 2, five at Level 3, and three at Level 4.

³⁴ The validation study gathered data on administrators’ IPDPs, but did not collect data on the frequency with which IPDPs were created and reviewed.

4A.2.1. (Level 2) Program completes Strengthening Families Self-Assessment and uses data to engage in continuous improvement.

To achieve a designation as Level 2, programs are required to conduct a Strengthening Families self-assessment, which asks programs “to compare their practices with those of exemplary programs and identify areas for improvement,” and also to provide a description of the program improvement plan based on that self-assessment that includes current goals and activities for strengthening family and community engagement.³⁵ Validation study data show at Levels 2 and above, most programs, ranging from 67 percent (Level 4) to 80 percent (Level 2) completed a Strengthening Families self-assessment (Table 54).

In interviews, several administrators noted that they did not feel as though the Strengthening Families self-assessment was well-aligned with the needs of their clientele. Data indicate that programs with relatively few children receiving subsidies were less likely to have completed a Strengthening Families self-assessment — 54 percent of programs with less than a quarter of enrolled children receiving subsidies, compared with 79 percent or programs with a higher proportion of children on subsidies. It was also noted that some Head Start programs were not using the Strengthening Families self-assessment but were following Head Start guidelines related to family and community engagement, which was seen as aligned with the goals of the Strengthening Families programs.

Table 54. Programs meeting measured aspects of criterion 4A.2.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program has completed a Strengthening Families Self-Assessment	23%	80%	76%	67%

4A.2.2. (Level 2) Programs offer opportunities for parents to meet with classroom staff, at least monthly.

At Level 2, programs are also required to offer opportunities for parents to meet with classroom staff at least monthly, a requirement that is verified through a signed document that demonstrates that opportunities are planned, formal, intentional, and initiated by the program. Per the guidance, these opportunities can be face-to-face, electronic, or via phone. For the purposes of the validation study, sharing of information between programs and families was assessed through an indicator on the ERS observation. The proportion of programs engaged in “much sharing of child-related information” ranged from 79 percent at Level 1 to 100 percent at Level 4, although differences across levels were not significant (Table 55).

Table 55. Programs meeting measured aspects of criterion 4A.2.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Much sharing of child-related information between parents and staff in all study classrooms (up to two total, one preschool and one infant/toddler)	79%	85%	93%	nr
Much sharing of child-related information between parents and staff in observed preschool classroom (ECERS Indicator 38.5.3; programs with an observed preschool classroom only)	79%	85%	93%	100%
Much sharing of child-related information between parents and staff in infant/toddler classroom (ITERS Indicator 33.5.3; programs with an observed infant/toddler classroom only)	74%	84%	93%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

³⁵ Strengthening Families Self-Assessment Tool for Center-Based Early Education and Care Programs (2014). *Center for the Study of School Policy*. Retrieved from <http://www.cssp.org/reform/strengtheningfamilies/2014/CENTER-BASED-EARLY-CARE-AND-EDUCATION-PROGRAM-SELF-ASSESSMENT.pdf>

For the purposes of QRIS, it may be helpful to clarify the specific types of interaction between parents and classroom staff expected at Level 2 — for example, whether the expectation is for formal conferences or for more informal types of communication — and/or offer specific types of meetings that would be considered to meet criterion 4A.2.2. The level of formality expected may have bearing on the appropriate level of frequency, as informal check-ins between parents and staff on a monthly basis may be appropriate, but more formal interactions might warrant a lower level of required frequency.

4A.2.3. (Level 2) Program has developed informational materials on the program that are in the language of the community, are available for staff to use in the community, and are given to prospective families.

In addition, programs are also required to develop informational materials in the language of the community. For the purposes of QRIS, verification of this criterion is assessed using a program’s ERS scores, even though this is not specifically addressed in the measure. As Table 56 indicates, nearly all programs have developed informational materials for families, regardless of QRIS level. However, when other languages are spoken in the community, programs at Levels 3 and 4 were much more likely to have those materials translated into other languages. The relatively high percentage of Level 2 programs not meeting the language requirement for materials (32 percent) suggests that current verification procedures using ERS scores may not be sufficient to assess this requirement, and MA EEC may want to consider amending verification procedures to assess this criterion more directly, should this requirement remain part of the MA QRIS. Also of note, most Level 1 programs who met the requirement regarding language did so because they self-identified as serving communities that do not commonly have other languages spoken (see Table 57).

Table 56. Programs meeting measured aspects of criterion 4A.2.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 4A.2.3	67%	68%	93%	83%
Program has written informational materials about the program	97%	100%	100%	100%
Informational materials are available in languages other than English, or no other languages commonly spoken in the community	67%	68%	93%	83%

Table 57. Reason program meets language access component of criterion 4A.2.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Has materials in other languages	26%	55%	88%	50%
No other languages in community	41%	13%	5%	33%
Does not meet requirements for 4A23 for language availability of informational materials	33%	33%	7%	17%

4A.2.4. (Level 2) Program maintains ongoing communication with the school/early intervention program, CFCE grantee, mental health providers to facilitate collaboration and coordination of services that support children and families.

Criterion 4A.2.4 is verified through a combination of a program’s self-assessed Program Administration Scale (PAS) score (required to be 3 or higher), a list of agencies with whom the program collaborates, a description of the procedure for referring children and families, and a description of the procedures for following up with services and supports. The validation study assessed two of the central aspects of this criterion, namely the program’s relationship with providers of Early Intervention Services and providers of special education services, typically the public schools in the child’s home school district.

As shown in Table 58, all or nearly all programs at Levels 2 and above indicated that they had a relationship with organizations providing both types of services, as did most programs at Level 1. For the purposes of the study, a program was considered to be working with an agency if they could identify the name of the agency and the procedures for referring a child or family, which is generally consistent with MA EEC verification guidance. Notably, the study was unable to reliably assess whether a program was working with a CFCE grantee, as many programs could identify organizations that they worked with, but not whether those organizations were CFCE grantees.

Table 58. Programs meeting measured aspects of criterion 4A.2.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 4A.2.4	82%	98%	100%	100%
Program works with an agency that provides Early Intervention services for children who need them (only programs who serve infants and toddlers)	93%	97%	100%	100%
Administrator interview: Program works with an agency that provides Special Education services for children who need them (programs that serve preschoolers)	87%	100%	100%	100%

With regard to mental health services, programs were less likely to identify an agency that they were working with for these services specifically. Overall, 46 percent of programs at Level 1, 68 percent at Level 2, 88 percent at Level 3, and 100 percent at Level 4 said they worked closely with an organization that could provide mental health services. However, the interview item asked about services for families and referenced family issues, including stress management, which may extend beyond the requirement for meeting criterion 4A.2.4, and thus was not included in the calculation of the overall proportion of programs meeting the criterion.

4A.2.5. (Level 2) Program participates in community events.

At Level 2, programs are also asked to provide a list of events that its staff participated in within the last twelve months. To be considered for criterion 4A.2.5, events must have been collaboration with other community organizations; class field trips are not normally considered community events. As reported by validation study participants (Table 59), 83 percent of programs at Level 2 participated in at least one type of community event in the last twelve months, as did all programs in Level 4 and nearly all programs in Level 3 (95 percent). Participation in community events was also common among Level 1 programs, suggesting that this requirement is unlikely to be functioning as a barrier to progression in the system.

Table 59. Programs meeting measured aspects of criterion 4A.2.5, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program participates in community events to raise awareness of the program and/or the benefits of early education	77%	83%	95%	100%

Conferences and community meetings were the most common type of community events programs reported participating in (Table 61). Relatively few Level 1 providers (8 percent) indicated that they had participated in meetings with legislators or policy makers, whereas this was much more common among programs at higher levels (43 percent at Level 2, 39 percent at Level 3, and 100 percent at Level 4). On average, programs at higher levels appeared to be engaged in different event types. Based on the data below (Table 60), we recommend that MA EEC refine the measure to specify community events that are relevant to the criteria.

Table 60. Program participation in community event types, by MA QRIS Level

	Total	Level 1	Level 2	Level 3	Level 4
Conferences for early education providers	60%	49%	60%	63%	100%
Community meetings	60%	44%	55%	73%	100%
Community health or wellness events or fairs	52%	36%	58%	61%	67%
Neighborhood celebrations	42%	31%	40%	56%	33%
Meetings with legislators or policy makers	33%	8%	43%	39%	100%
Other types of events	9%	13%	10%	5%	0%
Average number of different types of meetings selected	2.5	1.8	2.6	3.0	4.0

4A.3.1. (Level 3) A daily two-way communication system is available between the educators and families through a variety of means.

To be verified as meeting criterion 4A.3.1, programs must demonstrate that they use two or more means of communication with families.³⁶ Teachers in observed classrooms were asked to indicate the frequency with which they used a variety of communication methods with families, including check-ins during drop-off or pick-up, notes that go home with children, notebooks or logbooks, checklists, phone calls, emails, and text messages. Overall, nearly all educators indicated that they routinely used at least two different modes of communication on a regular basis, ranging from 92 percent at Level 1 to 100 percent at Level 4 (Table 61). As such, this criterion, as currently construed, does not appear to be substantively distinguishing quality across levels and would be a potential candidate for removal from the MA QRIS or for substantive revision. Notably, there were no significant differences in the average number of communication methods used in infant and toddler classrooms, but there were differences in the average number used in preschool classrooms, with educators in Level 3 classrooms using more communication methods, on average, than those in Level 2 classrooms (Table 62).

Table 61. Programs meeting measured aspects of criterion 4A.3.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
S4A31. A daily two-way communication system is available between the educators and families through a variety of means	92%	98%	95%	100%
Teacher interview (preschool): Program uses two modes of communication to share information with families	92%	97%	95%	100%
Teacher interview (infant/toddler): Program uses two modes of communication to share information with families	100%	100%	93%	100%

Table 62. Average number of modes of communication used by educators, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Average number of communication methods used, as reported by <u>preschool</u> teachers in observed classrooms	3.1	2.9	3.6	3.4
Average number of communication methods used, as reported by <u>infant and toddler</u> teachers in observed classrooms	3.0	2.9	2.7	3.0

³⁶ Head Start programs and those accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

4A.3.2. (Level 3) Families are encouraged to volunteer in the program, to assist in the classroom, and share cultural and language traditions or other interests such as their jobs, hobbies, and other relevant information.

For the MA QRIS, attainment of criterion 4A.3.2 is assessed using a program's ERS scores and whether those scores meet Level 3 requirements both overall and for individual subscales. Programs are also provided with technical assistance on the ERS Subscale 7 (Parents and Staff) as part of their visit from a Program Quality Specialist. As mentioned previously, because overall and subscale scores take into account multiple items of quality, it is possible for programs to meet verification standards but not have representative materials in their classroom. As such, to assess the criterion for the QRIS validation, the study used an indicator contained with Subscale 7 related to family involvement in the program.

As Table 63 displays, nearly all programs at Level 3 and all programs at Level 4 encourage families to be directly engaged in the program by offering a variety of alternatives for such involvement, a significantly higher proportion than those in Levels 1 and 2. To be considered to be meeting this indicator, programs must offer at least three different ways for families to be involved, such as assisting staff in field trips, sharing cultural customs with the class, helping with fundraising events, collecting materials for the program, and other such activities. MA EEC may want to consider changing its verification requirement to use the ERS observations indicators (ECERS-R Indicator 38.5.4 and ITERS-R Indicator 33.5.4) to align more directly with the criterion itself.

Table 63. Programs meeting measured aspects of criterion 4A.3.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Variety of alternatives used to encourage family involvement in children's program in all observed classrooms (up to two total, one preschool and one infant/toddler)	56%	65%	95%	nr
Variety of alternatives used to encourage family involvement in children's program in preschool classroom (ECERS Indicator 38.5.4; programs with an observed preschool classroom only)	59%	64%	95%	100%
Variety of alternatives used to encourage family involvement in children's program (ITERS Indicator 33.5.4; programs with an observed infant/toddler classroom only)	41%	55%	93%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

4A.3.3. (Level 3) Program ensures that there are translators available, as needed, at meetings, workshops and conferences to ensure strong communication between the program and families.

To be verified as meeting criterion 4A.3.3 programs must submit a description of the process they use for translation. If the program uses its own staff, programs must also identify the number of staff used as translators, or if the service is offered by an outside contractor, an MOU or other proof of collaboration must be provided prior to verification.

For the validation study, program administrators were asked to identify the frequency with which they provided translators for events and meetings. Overall, most program met this requirements across the MA QRIS Levels (Table 64). It is important to note that the high level of program meeting this criterion is partially drive by the percentages of programs that reported no other languages commonly spoken in their community. This was particularly true for Level 1 and Level 4 programs in which 41% and 33% of programs, respectively, reported having no other languages commonly spoken in their community (Table 65).

Table 64. Programs meeting measured aspects of criterion 4A.3.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
S4A33. Program ensures that there are translators available, as needed, at meetings, workshops and conferences to ensure strong communication between the program and families.	82%	83%	93%	100%

Table 65. Reason program meets criterion 4A.3.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program provides translators at events or meetings	41%	68%	83%	67%
Program indicated that no other languages are commonly spoken in the community	41%	15%	10%	33%
Program does not meet criterion 4A.3.3	18%	18%	7%	0%

4A.3.4. (Level 3) Program representative(s) participate in local community group work that is related to early childhood, and the cultural groups served by the program and/or family support.

Criterion 4A.3.4 is verified using a program’s overall Program Administration Scale (PAS) score. Although community engagement is assessed through the PAS, as is the case with regard to the use of overall scores for the ERS, a program’s overall score does not provide specific information about its practice in a particular area. As such, the validation study did not use the PAS to assess criterion 4A.3.4, and instead gathered data regarding programs’ participation in community group work through a program information questionnaire. As Table 66 shows, at Level 3, 88 percent of programs had administrators or central staff participated in early education and care organizations, and at Level 4, 100 percent of programs reported that this was the case. Among programs at Levels 1 and 2, the proportions were 69 percent and 78 percent respectively, suggesting the practice is also common at lower levels, although somewhat less so. As with other criteria that use overall scores to assess specific practice, MA EEC may want to consider verifying this criterion using a more direct assessment. Further clarity regarding the level or participation expected — for example, membership, regular attendance at meetings, serving on a committee, board, or as an officeholder — might also be beneficial.

Table 66. Programs meeting measured aspects of criterion 4A.3.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program administrator or central staff participate in early education and care organizations	69%	78%	88%	100%

Across the MA QRIS, the types of participation reported by programs varied. Regular attendance at meetings and membership in organizations tended to be the most common ways of participation across all levels (Table 67). The proportion of programs whose administrators played a more active role, for example chaired a committee, served on a board, or held an office, increased across the levels, from 8 percent at Level 1 to 46 percent at Level 3 and 83 percent at Level 4. Other types of involvement indicated on the questionnaire included participation in directors’ groups and serving as collaborators for CFCE grants.

Table 67. Reported ways of participating in early education and care organizations, by MA QRIS Level

	Total	Level 1	Level 2	Level 3	Level 4
Regularly attend meetings	64%	44%	65%	78%	100%
Members of the organization	63%	46%	63%	73%	100%
Chair a committee, serves on a board, or holds office	30%	8%	28%	46%	83%
Other involvement	8%	8%	10%	7%	0%

4A.3.5. (Level 3) Program ensures young children and their families have access to developmental, mental health, health, and nutrition services either through private pay arrangements OR are offered such services through other programs.

To be verified as meeting criterion 4A.3.5, MA EEC verifies that the program ensures children and families have access to developmental screenings, mental health screenings, health screenings, and nutrition screenings, and if these services are offered via an outside contractor, an MOU must be provided. This criterion may benefit from further clarity regarding MA EEC expectations for ensuring access. For example, programs may be expected to track whether children have received these screenings and services and refer those who have not received the services to a provider. However, clarifications regarding what programs are expected to do in the event that families are unable to pay for services or do not access the services would be beneficial.

The validation study gathered data about whether programs tracked if children had received each of these services. As shown in Table 68, most programs, regardless of level, were tracking whether children received physical examinations, and at Level 3 and above, all or nearly all programs tracked developmental screenings. Notably, programs are required to conduct developmental screenings as part of a Level 3 standard in the Curriculum, Assessment, and Diversity Standard (criterion 1A.3.3). As such, this aspect of criterion 4A.3.5 may be redundant with that other requirement and could potentially be dropped. Most programs at Levels 3 and 4 also tracked nutrition screening/services and mental health screenings, whereas these tended to be less commonly tracked among Level 1 and Level 2 programs. Although developmental screenings may often assess aspects of social/emotional development, this may suggest that few programs are using a dedicated social/emotional screening tool.

Table 68. Programs meeting measured aspects of criterion 4A.3.5, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 4A.3.5	18%	43%	83%	83%
Program tracks whether children receive physical examinations (health screenings)	90%	100%	100%	100%
Program tracks whether children receive developmental screenings	64%	88%	98%	100%
Program tracks whether children receive nutrition screenings or services	31%	55%	88%	100%
Program tracks whether children receive mental health screenings	28%	55%	85%	83%

4A.4.1. (Level 4) Parents participate on the Advisory Board for the program and are actively involved in the policy and decision making for the program.

Whereas at Level 3, programs are expected to encourage family engagement at the program (criterion 4A.3.2), at Level 4, programs are asked to actively involve parents in policy and decision-making for the program (criterion 4A.4.1) For QRIS, criterion 4A.4.1 is assessed using a program's ERS scores and whether those scores meet Level

4 requirements both overall and for individual subscales.³⁷ As mentioned previously, because overall and subscale scores take into account multiple items of quality, it is possible for programs to meet verification standards but not have representative materials in their classroom. As such, to assess the criterion for the QRIS validation, the study used an indicator contained with Subscale 7 related to family involvement in the program.

ECERS-R data show that at all programs at Level 4, parents were involved in decision-making roles, as was the case at 83 percent of Level 3 programs (Table 69). Data regarding this indicator for infant/toddler classrooms are not reported, as only one Level 4 program in the study had an infant or toddler classroom in the study. Parent involvement in decision-making roles was much less common in Level 1 and Level 2 programs (23 percent and 43 percent of programs, respectively). MA EEC may want to consider more directly assessing this criterion using the related ERS indicators (ECERS-R Indicator 38.7.3 and ITERS-R Indicator 33.7.3).

Table 69. Programs meeting measured aspects of criterion 4A.4.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Parents participate on the Advisory Board for the program and are actively involved in the policy and decision making for the program (up to two total, one preschool and one infant/toddler)	23%	43%	83%	nr
Parents involved in decision-making roles in program along with staff, preschool classroom (ECERS Indicator 38.7.3; programs with an observed preschool classroom only)	23%	46%	83%	100%
Parents involved in decision-making roles in program along with staff (ITERS Indicator 33.7.3; programs with observed infant/toddler classroom only)	19%	26%	73%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

4A.4.2. (Level 4) Program provides or connects families to education, training, and support programs (such as family literacy, adult education, job training, child development, parenting, English as a second language, etc.).

To be verified as meeting criterion 4A.2.4, programs must demonstrate that they connect families to education, training, and support and provide an MOU or other proof that collaboration is in place.³⁸ Although examples are provided, the criterion does not reference specific requirements about how many or what services specifically are required to meet the criterion. The validation study considered a program to meet this requirement if they were offering or connecting families to at least three services, with descriptive information provided for other possible thresholds. Overall, most programs at Level 4 offered or connected families to at least three of the services (83 percent), as did 88 percent of programs at Level 3 (Table 70). Fewer programs at Levels 1 and 2 did so, or 21 percent and 45 percent respectively. However, when MOUs are included as part of the standard, consistent with program quality specialist guidance, fewer programs met the requirement. Data suggests that formalized MOUs could be serving as barriers to programs meeting MA QRIS requirements, even though programs may be connecting families to agencies and services.

³⁷ Head Start programs and those accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

³⁸ Head Start programs are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

Table 70. Programs meeting measured aspects of criterion 4A.4.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of 4A.4.2	13%	38%	78%	67%
Program provides or connects families with an agency that provides three of the following services: family literacy, adult education, job training, child development training, parenting education, or English as a Second Language	21%	45%	88%	83%
Program has a collaboration agreement with agency or staff provides directly three of the above services	13%	38%	78%	67%

All or nearly all programs in the two upper tiers of QRIS provided or connected families to at least one service and had an MOU for such, or provided it directly using their own staff, as did many Level 2 programs (65 percent) (Tables 71, 72). Across all levels, child development training, parenting education, and family literacy were the most common services programs were offering or connecting families with. English as a second language (ESL), adult education, and job training were generally less common. In interviews, several program administrators noted that these services were not typically required by families they served; hence, they did not routinely provide or connect families to these services.

Table 71. Number of listed services programs offered or connected families with, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
At least one of the listed services	44% (38%)	65% (58%)	98% (98%)	83% (83%)
At least two of the listed services	26% (18%)	60% (55%)	95% (93%)	83% (83%)
At least three of the listed services	21% (13%)	45% (38%)	88% (78%)	83% (67%)
At least four of the listed services	13% (5%)	43% (38%)	78% (68%)	67% (50%)
At least five of the listed service	8% (3%)	38% (28%)	63% (54%)	67% (17%)
All six of the listed services	0% (0%)	28% (15%)	49% (37%)	33% (17%)
Average number	1.1 (0.8)	2.8 (2.3)	4.7 (4.3)	4.2 (3.2)

Numbers in parentheses reflect the proportion if only those with MOUs or provided by staff are included

Table 72. Percentage of programs offering or connecting families to services, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Parenting education	31%	55%	88%	83%
Child development training	26%	53%	88%	67%
Family literacy	21%	53%	80%	67%
English as a second language	15%	33%	76%	67%
Adult education	13%	48%	78%	67%
Job training	5%	38%	61%	67%

4A.4.3. (Level 4) Program ensures all children and families have access to comprehensive screenings, referrals and services including developmental screening, mental health screening, speech screening, speech therapy, physical therapy, occupational therapy, dental health care, and nutrition services.

To be verified as meeting criterion 4A.4.3, programs must provide a document signed by the administrator demonstrating that children have access to development screening, mental health screening, speech screening, speech therapy, occupational therapy, dental health care, and nutrition services.³⁹ If the services are provided through a partnership with another organization, an MOU or proof of collaboration must also be provided.

The validation study gathered data on whether the program tracked if children had received a number of different services aligned with the criterion (Table 73). Data were not available on how these services were provided or whether a collaboration agreement was in place. For all but two of these services, all programs at Level 4 were tracking whether children had received the service, and 67 percent of programs tracked all services asked on the questionnaire (which did not include dental health care, but instead asked about dental screenings).

Table 73. Programs meeting measured aspects of criterion 4A.4.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 4A.4.3	15%	30%	76%	67%
Program tracks whether children receive developmental screenings	64%	88%	98%	100%
Program tracks whether children receive mental health screenings	28%	55%	85%	83%
Program tracks whether children receive speech screenings	51%	80%	93%	100%
Program tracks whether children receive speech therapy	72%	75%	98%	100%
Program tracks whether children receive physical therapy	56%	70%	90%	100%
Program tracks whether children receive occupational therapy	59%	70%	90%	100%
Program tracks whether children receive dental screenings	26%	43%	90%	83%
Program tracks whether children have nutritional screening/services	31%	55%	88%	100%

Notably, most aspects of criterion 4A.4.3 are duplicative of requirements elsewhere within the MA QRIS. For example, programs are required to conduct developmental screenings as part of a Level 3 standard in the curriculum, assessment, and diversity standard (criterion 1A.3.3), and developmental screenings are also referenced in a similar Level 3 criterion in the family and community engagement standard (criterion 4A.3.5). Similarly, several of these services — speech services, physical therapy, and occupation therapy — would typically only be received by children referred for special services through Early Intervention or special education. Programs are required to work with Early Intervention and special education providers at Level 2 (criterion 4A.2.4). One unique aspect of criterion 4A.4.3 is the requirement that programs ensure access to dental health care, although the validation study did not collect data about dental health care specifically. Dental screenings are required at Level 3 as part of the safe, healthy indoor and outdoor environments standard as part of criterion 2A.3.1, and validation study data show that 90 percent of programs at Level 3 and 83 percent of those at Level 4 provided those screenings, compared with 26 percent of programs at Level 1 and 43 percent at Level 2 (Table 73).

³⁹ Head Start programs are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

MA QRIS Quality Standard 5A: Leadership, Management, and Administration

According to program quality specialist guidance documents, “high quality programs require effective leadership with management and administrative practices that ensure a stable environment, fiscal accountability, evaluation of the program’s practices and policies, and the development of relationships within the community in order to support the staff and the children and families they serve.” These practices are assessed as part of the leadership, management, and administration standard of QRIS, which comprises a total of 14 criteria. This includes five criteria at Level 2, five criteria at Level 3, and four criteria at Level 4.

5A.2.1. (Level 2) Communication and updates on the program are provided to educators and families.

To be verified as meeting criterion 5A.2.1, programs must provide a statement that program updates are provided to both staff and families at least quarterly in their primary or preferred language to the extent possible, as well as explanation of how updates are provided.⁴⁰ On program questionnaires collected through the validation study, all programs, regardless of MA QRIS level, indicated that they provided program updates at least quarterly, often using multiple communication methods (Table 74). Because this criterion shows no variation across programs or levels, MA EEC may want to consider dropping it from the MA QRIS, as it does not meaningfully contribute to level designations.

Table 74. Programs meeting measured aspects of criterion 5A.2.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5A.2.1	100%	100%	100%	100%
Program provides updates to staff at least quarterly	100%	100%	100%	100%
Program provides updates to families at least quarterly	100%	100%	100%	100%

As an alternative, MA EEC may want to consider refining the criterion to focus on specific communication strategies that they believe to be most effective. Descriptives are provided below of the different communication strategies being employed by programs in the study (see Tables 75, 76).

Table 75. Programs using communication strategies with staff at least quarterly, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Center-wide staff meetings	97%	100%	80%	67%
Team or classroom meetings	90%	93%	100%	100%
Staff bulletin board	77%	85%	76%	100%
Internal memos	74%	80%	73%	100%
Email voicemail	67%	63%	80%	83%
Meetings with educators individually	62%	83%	95%	100%
In-house newsletters	62%	68%	83%	67%
Text messaging	46%	38%	37%	83%
Message book	26%	33%	24%	67%
Routing slips	3%	5%	12%	33%

⁴⁰ Head Start programs are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

Table 76. Programs using communication strategies with families at least quarterly, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Newsletters	92%	90%	100%	100%
Bulletin board	92%	98%	98%	100%
Notes that go home with children	90%	98%	88%	100%
Email	74%	45%	50%	67%
Website	65%	56%	39%	67%
Phone calls	56%	83%	95%	83%
Family meetings	28%	43%	78%	83%
Mailed letters	13%	38%	25%	100%

5A.2.2. (Level 2) Program has a written admissions policy that promotes an awareness of and respect for differences among children and families, a respect for the child and their family's culture and language, and is responsive to the inclusion of a variety of learning needs.

Programs also need to have an admissions policy that promotes diversity, and verification of this criterion takes into account whether a program has this policy, as well as a program's ERS scores, both overall and for individual subscales.⁴¹ Regardless of level, all or nearly all programs reported that they had a written admissions policy (Table 77). Most programs at Level 2 and above indicated that their admissions policy had language to promote diversity, which is generally consistent with this Level 2 requirement. Many Level 1 programs (77 percent) also indicated that their policy had language promoting diversity, suggesting that this criterion is unlikely to be functioning as a substantial barrier to progression.

Table 77. Programs meeting measured aspects of criterion 5A.2.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5A.2.2	77%	90%	88%	83%
Program has a written admissions policy	97%	100%	100%	100%
Program's admission policy has language that promotes diversity	77%	90%	88%	83%

5A.2.3. (Level 2) Program has a written business plan.

At Level 2, programs are also required to provide a copy of their business plan, including an annual operating budget. Data show that, at Level 2 and above, all or nearly all programs met these requirements, compared with 69 percent of programs at Level 1 (Table 78). Only one program had a business plan that did not include an operating budget.

⁴¹ Head Start programs and those accredited through NAEYC are automatically verified as meeting this criterion, although this was not considered as part of the validation study.

Table 78. Programs meeting measured aspects of criterion 5A.2.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5A.2.3	69%	95%	100%	100%
Program has a written business plan	69%	98%	100%	100%
Program's business plan contains an operating budget.	69%	95%	100%	100%

5A.2.4. (Level 2) Staff are paid for planning time.

At Level 2, staff are also required to have paid planning time, which is verified if a program has an overall self-assessed Program Administration Scale (PAS) score of 3 or higher. As mentioned previously, because programs can meet overall score requirements on a measure without having a specific practice in place or could have a specific practice in place without meeting an overall score requirement, for the validation study, this criterion was assessed directly through study instruments, in this case, an item on the program questionnaire that asked if teachers had paid planning time.

Overall, nearly all Level 3 and Level 4 programs met this Level 2 requirement, by offering paid planning time to full-time staff, part-time staff, or both (Table 79). Approximately 22 percent of Level 2 programs indicated that staff did not have paid planning time. Most programs who met the requirement did so by providing paid planning time to full-time staff; fewer programs provided planning time to both full-time and part-time staff (Table 80). One potential recommendation might be to consider whether planning time must be when children are not present (i.e. whether nap time could constitute planning time). This was not explicitly required in the criterion and was not taken into account in the validation study measure so data on this are not available.

Table 79. Programs meeting measured aspects of criterion 5A.2.5, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Staff have paid planning time	67%	78%	90%	83%

Table 80. Reason program meets criterion 5A.2.5, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Provide paid planning time to both full- and part-time staff	21%	33%	27%	50%
Provide paid planning time to full-time staff only	44%	45%	63%	33%
Provide paid planning time to part-time staff only	3%	0%	0%	0%
Does not provide paid planning time to staff	33%	22%	10%	17%

5A.2.5. (Level 2) Program has policies that support teacher retention.

Criterion 5A.2.5, which requires that a program have policy that supports teacher retention, is verified if a program provides vacation and/or sick time, education/tuition reimbursement, free or reduced-fee child care, a retirement plan contribution, or if the program provides pay increases/a career ladder for staff. Data from the validation study questionnaire show that all programs met this requirement, regardless of level (Table 81). As such, criterion 5A.2.5 reflects a criterion that MA EEC may want to consider removing from the MA QRIS. Notably, for all but a handful of programs, both pay increases and one or more of the listed benefits were provided (Table 82).

Table 81. Programs meeting measured aspects of criterion 5A.2.5, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program has policies that support teacher retention	100%	100%	100%	100%

Table 82. Reason program meets criterion 5A.25, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Offers both a benefit listed in PQS guidance and pay increases	82%	100%	90%	100%
Offers a benefit listed in PQS guidance only	15%	0%	10%	0%
Offers pay increases only	3%	0%	0%	0%

5A.3.1. (Level 3) Program tracks and monitors absences of individual children and contacts families when children are absent more than 20 percent in a month.

Criterion 5A.3.1 is verified for programs if they use a paper-based or computer-based system for tracking absences, contact families when children are absent more than 20 percent in a month, and provide support to families to help reduce absences. Data gathered as part of administrator interviews indicated that regardless of MA QRIS level, all or nearly all programs were routinely tracking absences (Table 83). Programs were generally contacting families either after every absence or after a certain number of consecutive absences, not based on the proportion of absences in a period of time.

Table 83. Programs meeting measured aspects of criterion 5A.3.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5A.3.1	97%	95%	95%	100%
Program tracks absences using a paper-based or computer system	97%	95%	100%	100%
Program routinely contacts families when children are absent	100%	100%	100%	100%
Program contacts families after 3 consecutive absences or less	100%	100%	95%	100%

The validation study did not collect data directly about support provided to families, as this was not included in the criterion or its measurement method at the time the study instruments were developed. However, programs were asked about policies related to absences in open-ended format. These policies tended to vary widely. For example, programs that exclusively served private pay families often noted that they had no policy, as the issue of chronic absenteeism had not come up. Programs with children receiving vouchers or subsidies often acknowledged requirements regarding attendance and efforts to alert parents to these expectations. Provision of support or descriptions of working with parents to promote better attendance were commonly reported, although this was not always the case. If this reflects an important part of the criterion, MA EEC may want to consider referencing the support requirement in the criterion and providing guidance to programs.

5A.3.2. (Level 3) Program director, staff, and family input is solicited on an annual basis through a survey to evaluate the program.

At Level 3, programs are required to solicit input from staff and families to evaluate the program. This criterion is verified by an overall self-assessed Program Administration Scale (PAS) score of 5 or higher. As mentioned previously, because programs can meet overall score requirements on a measure without having a specific practice in place or could have a specific practice in place without meeting an overall score requirement, for the validation study, this criterion was assessed directly through the administrator interview.

Overall, these data show that, consistent with the requirement, at Level 3 nearly all programs (93 percent) conducted annual surveys of both families and staff, as did all programs at Level 4 (Table 84). Most Level 2 programs also surveyed both groups annually (83 percent), but the practice was much less common at Level 1 (36 percent). Overall, family surveys were more common than staff surveys. At Levels 2 and 3, programs were slightly more likely to be administering a family survey than a staff survey, and at Level 1, family surveys were much more commonly administered (72 percent of programs administered family surveys and 49 percent of programs administered staff surveys). A separate Level 3 criterion, 5A.3.3 (discussed below) requires that programs use these results as part of program improvement efforts. It may be beneficial to combine these two criteria into a single criterion, as they are directly related to one another and are part of the same cycle of inquiry — gathering of data and use of the information.

Table 84. Programs meeting measured aspects of criterion 5A.3.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5A.3.2	36%	83%	93%	100%
Program surveys families to evaluate the program	72%	95%	100%	100%
Family survey is conducted annually	56%	85%	100%	100%
Program surveys staff to evaluate the program	49%	85%	93%	100%
Staff survey is conducted annually	41%	83%	93%	100%

5A.3.3. (Level 3) Results of the annual survey are used to develop a comprehensive written program improvement plan.

Criterion 5A.3.3 requires that programs incorporate the results of family and staff surveys into programs' quality improvement plans. In interviews, nearly all administrators at Level 3 programs and all of those at Level 4 programs indicated that they used the results of their family and staff surveys to plan for or make changes at their program (Table 85). Fewer programs at Level 2 did so (75 percent) and substantially fewer programs at Level 1 met this requirement. As mentioned above, it may be beneficial to combine this criterion with another Level 3 criterion that requires administration of family and staff surveys on annual basis (criterion 5A.3.2), as they are directly related to one another and are part of the same cycle of inquiry — gathering of data and use of the information.

Table 85. Programs meeting measured aspects of criterion 5A.3.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5A.3.3	38%	75%	90%	100%
Program surveys families to evaluate the program	72%	95%	100%	100%
Program uses the results of the family survey to plan for or make changes at the program	62%	88%	98%	100%
Program surveys staff to evaluate the program	49%	85%	93%	100%
Program uses the results of the staff survey to plan for or make changes at the program	44%	75%	93%	100%

5A.3.4. (Level 3) Program has an annual review conducted of the accounting records by an independent party who has accounting or bookkeeping expertise.

To be verified as meeting criterion 5A.3.4, programs must provide proof of an annual review of accounting records by an independent party. On program questionnaires, this Level 3 requirement was met by most programs at Level 2 and above (85 percent of programs at Levels 2 and 3 and 100 percent of programs at Level 4). That equal proportions of programs were meeting this requirement at Levels 2 (where it is not required) and Level 3 (where it is required) suggests that this is generally not a requirement that is posing a significant barrier to progression in the short-term. At Level 1, fewer programs had independent reviews of their accounting records, or 62 percent (see Table 86).

Table 86. Programs meeting measured aspects of criterion 5A.3.4 by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
An outside agency reviews accounting records	62%	85%	85%	100%

5A.3.5. (Level 3) Program tracks and monitors teacher turnover and has plan for addressing turnover.

To be verified as meeting criterion 5A.3.5, programs need to provide a description of the system for tracking teacher turnover and the plan for addressing turnover. As reported on the validation study questionnaire, most programs at Level 3 and all programs at Level 4 were tracking teacher turnover (Table 87). Although all administrators from Level 4 programs said they had a plan or policy in place for addressing turnover, relatively few administrators from Level 3 programs responded similarly (37 percent). This finding may point to a need to focus more explicitly on programs' plans for addressing turnover or a need for further clarification of the "plan for addressing turnover" requirement.

Table 87. Programs meeting measured aspects of criterion 5A.3.5, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program tracks or monitors the number of percentage of teachers who leave the program	36%	50%	80%	100%
Program has a plan for addressing turnover	26%	50%	37%	100%

5A.4.1. (Level 4) Program staff and advisory board are involved in the development of the business plan and it is reviewed periodically for updating.

At Level 2, programs are required to have a business plan (criterion 5A.2.3) and at Level 4, they are asked to both update their plan annually and engage program staff and their advisory board in the development process; they also need to provide documentation that this is taking place (criterion 5A.4.1). Information from the validation study program questionnaire indicates that all programs at Levels 3 and 4 reviewed and updated their business plans, as did nearly all Level 2 programs and about half of Level 1 programs (Table 88). At higher levels, programs were more likely to involve staff and advisory boards in the development or review of their plan, but the overall proportion was low across all levels. In Levels 1 through 3, the low prevalence of this practice is not surprising, although the fact that only half of programs at Level 4 involve both staff and their board in the process is somewhat notable. This suggests that the current documentation requirement may be effective in determining that administrative requirements of having a plan are being met but that more nuanced aspects of this practice may require additional attention as part of the verification process. Additionally, clarification may be needed regarding the requirement for the advisory board to be "involved in the development or review of the business plan" versus simply having the board approve the plan.

Table 88. Programs meeting measured aspects of criterion 5A.4.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5A.4.1	8%	40%	63%	50%
Program has a written business plan	69%	98%	100%	100%
Business plan is reviewed or updated annually	49%	85%	100%	100%
Administrative staff or educators are involved in the development or review of the business plan	28%	73%	76%	67%
Advisory board or parent advisory board is involved in the development or review of the business plan	21%	58%	73%	83%

5A.4.2. (Level 4) Program has a system of technology that allows for data collection and tracking program information.

Criterion 5A.4.2 is verified if a program has a system of technology that maintains and tracks children’s health information, services, absenteeism, children’s educational information, staff qualifications, professional development, financial records/information, and teacher turnover rates. Data from the program questionnaire indicate that 83 percent of Level 4 programs used a database or software system to maintain and track program information, as did 90 percent of Level 3 programs and 75 percent of Level 2 programs (Table 89).

Systems generally did not contain all of the information referenced in the verification guidance, and it may be helpful to clarify whether all are required in order for programs to meet this criterion. Overall, data systems most commonly captured data regarding children, including child assessments/screenings, children’s health information, and children’s attendance or absenteeism. Services children have received were somewhat less commonly maintained in these databases, except at higher levels. Data on staff qualifications were the least likely to be maintained in these systems, although it is possible that these data are maintained elsewhere as part of human resources systems. To ensure that staff data are tracked, it may be beneficial to create a separate standard regarding use of a database for staff qualifications and professional development information, possibly at a lower level.

Table 89. Programs meeting measured aspects of criterion 5A.4.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5A.4.2	3%	10%	39%	33%
Program uses a database or software system to maintain and track program information	49%	75%	90%	83%
Database or software system contains each of the following components:	3%	10%	39%	33%
Information on child assessments/screenings	13%	55%	95%	83%
Children's health information	28%	68%	83%	83%
Children's attendance or absenteeism	13%	58%	80%	50%
Services children have received	5%	25%	71%	50%
Parent payments and/or voucher information	38%	65%	61%	83%
Staff qualifications	21%	45%	56%	67%

5A.4.3. (Level 4) An outside audit is conducted annually by a certified public accountant.

At Level 3, programs are required to have an independent review of accounting records (criterion 5A.3.4), and at Level 4, programs are required to have an outside audit conducted by a certified public accountant (criterion 5A.4.3). Data from the validation study program questionnaire indicated that all Level 4 programs met this requirement (Table 90). This practice was also common among programs at Level 2 and Level 3 (75 percent and 78 percent, respectively), and as such, does not appear to be serving as a substantial barrier to progression.

Table 90. Programs meeting measured aspects of criterion 5A.4.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all required aspects of criterion 5A.4.3	59%	75%	78%	100%
Program has an outside agency review accounting records	62%	85%	85%	100%
Individual who conducts the review of accounting records is a CPA	59%	75%	78%	100%

5A.4.4. (Level 4) Program shares the results of the program quality rating with the families, staff, governing board and funders.

Finally, Level 4 programs are also expected to share the results of the MA QRIS rating with families, staff, their governing board, and funders. Within the MA QRIS, this criterion is assessed using the program's ERS scores and whether those scores meet both overall and subscale requirements for Level 4. This practice is not assessed through the ERS, and as such, ERS scores are not an effective measure of this criterion.

MA QRIS Quality Standard 5B: Supervision

Finally, the MA QRIS contains a series of criteria that address supervision and supports the program provides to educators. This standard comprises a total of 11 criteria — two at Level 2, five at Level 3, and four at Level 4.

5B.2.1. (Level 2) Program provides recognition for staff in annual evaluation as well as in public forum, as appropriate (i.e. verbal recognition in group setting or written recognition in newsletter).

For a program to meet criterion 5B.2.1, a program quality specialist must verify that the program provides recognition for staff in the annual evaluation, as well as a public forum (for example, a staff meeting, newsletter, or bulletin board). Data from interviews conducted as part of ERS observations indicated that across all levels of the MA QRIS, all or nearly all programs prepared written evaluations for staff (Table 91). As such, this might reflect a criterion that could be removed from the MA QRIS or otherwise modified, as it is not currently functioning to substantively differentiate levels. Data were not collected regarding whether programs recognize staff in a public forum, and for the purposes of the MA QRIS, reliable information would be difficult to collect from program administrators.

Table 91. Programs meeting measured aspects of criterion 5B.2.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Written evaluation of performance shared with staff at least yearly in all observed classrooms (up to two total, one preschool and one infant/toddler)	92%	98%	100%	nr
Written evaluation of performance shared with staff at least yearly for <u>preschool</u> teacher (ECERS Indicator 42.5.2; programs with an observed preschool classroom only)	92%	97%	100%	100%
Written evaluation of performance shared with staff at least yearly for <u>infant or toddler</u> teacher (ITERS Indicator 38.5.2.; programs with an observed infant/toddler classroom only)	93%	97%	100%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

5B.2.2. (Level 2) Staff are given feedback on instructional practice on a monthly basis.

Programs are verified as meeting criterion 5B.2.2 if the program’s educators are observed by a supervisor while working with children at least once a month and are given feedback on instructional practice. For the validation study, data from an indicator collected as part of ERS observations and related interviews regarding whether frequent observations and feedback — defined by the ERS as taking place at least every other month, with some flexibility permitted, depending on the needs of individual staff — were taking place. These data were collected directly from classroom staff, as opposed to administrators, which are the source of data for MA QRIS verification.

Overall, teachers in all observed classrooms at Levels 3 and 4 indicated that observations were taking place in their classrooms at least every other month (Table 92). At Level 2, observations were taking place with this same level of frequency at 67 percent of programs at Level 1 and 73 percent of programs at Level 2. That nearly a quarter of Level 2 programs were not meeting the ERS indicator related to feedback based on teacher reports is notable given that this level of frequency is lower that required in the MA QRIS (every other month vs. monthly). This finding may highlight differences between programs’ stated policies regarding observations and their enacted practice in this area, which in turn suggests a need for more robust verification protocols that go beyond program policies and/or administrator reports. Examples might be to require log-books or an observation schedule as part of Level 2 documentation.

Table 92. Programs meeting measured aspects of criterion 5B.2.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Frequent observations and feedback given to staff in addition to annual evaluation in all observed classrooms (up to two total, one preschool and one infant/toddler)	67%	73%	100%	nr
Frequent observations and feedback given to staff in addition to annual evaluation in preschool classroom (ECERS Indicator 42.7.2; programs with an observed preschool classroom only)	69%	72%	100%	100%
Frequent observations and feedback given to staff in addition to annual evaluation in the infant or toddler classroom (ITERS Indicator 38.7.2; programs with an observed infant or toddler classroom only)	67%	71%	100%	nr

Data for Level 4 infant and toddler classrooms are not reported; only one Level 4 program in the study had an infant or toddler room.

5B.3.1. (Level 3) Program uses at least 3 types of internal communication on a monthly basis to inform staff of program activities, policies, etc.

At Level 3, programs are also required to provide a description of three types of internal communication used to communicate program activities and policies to staff. On program questionnaires, all or nearly all programs, regardless of level, indicated at least three types of communication they used to communicate important program information to staff (Table 93). Given that there is almost no variation across programs or levels with respect to this practice, this criterion could be reconsidered for removal or modified.

Table 93. Programs meeting measured aspects of criterion 5B.3.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program identified three communication mechanisms used on a monthly basis to inform staff of program activity, policies, and practices	97%	98%	98%	100%

In terms of communication methods used, meetings were the most common, as all programs indicated that meetings took place at least once a month (Table 94). Staff bulletin boards, email/voicemail, internal memos, and in-house newsletters were also common means of communicating with staff.

Table 94. Communication methods used at least monthly, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Meetings (Center-wide, team, or individual)	100%	100%	100%	100%
Staff bulletin board	77%	83%	69%	67%
Email/voicemail	62%	60%	79%	83%
Internal memos	59%	73%	64%	100%
In-house newsletters	56%	60%	74%	50%
Text messaging	38%	28%	31%	83%
Message book	23%	33%	23%	67%
Routing slips	3%	5%	10%	17%

5B.3.2. (Level 3) Staff receive at least one benefit (paid vacation time, sick time, health insurance, tuition/PD reimbursement or retirement plan option).

To meet criterion 5B.3.2, programs must achieve a score of 5 or above on the PAS and provide an example of at least one benefit, such as paid vacation time, sick time, health insurance, tuition/professional development reimbursement, or a retirement plan option. This criterion is part of a series of two criteria, the other being a Level 4 requirement that programs offer vacation time, sick time, and health insurance.

Program questionnaires indicate that across all levels, all or nearly all programs provided at least one benefit to full- and/or part-time staff, suggesting that this is a criterion EEC may want to consider removing or modifying (Table 95). In terms of benefits offered, vacation and sick time were commonly offered, regardless of program’s QRIS level. Health insurance was typically offered at nearly all programs at Level 2 (93 percent), Level 3 (97 percent), and Level 4, but only at slightly more than half (56 percent) of Level 1 programs (Table 96). It should be noted that because the criterion does not specify whether benefits must be offered to all staff, both full- and part-time, and whether benefits need to be offered to assistant teachers in addition to teachers, for the purposes of the validation study, a program was considered to be offering a benefit if that benefit was offered to any educators, full or part time. For the purposes of QRIS, it may be beneficial to clarify expectations regarding who must receive these benefits.

Table 95. Programs meeting measured aspects of criterion 5B.3.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program offers teachers at least one benefit to full and/or part time staff (Paid vacation time, sick time, health insurance, tuition/PD reimbursement or retirement plan option)	97%	100%	100%	100%

Table 96. Benefits offered to full and/or part time staff, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Vacation time	90%	100%	95%	100%
Sick time	90%	90%	98%	100%
Health insurance	56%	93%	95%	100%
Tuition/PD reimbursement	51%	70%	85%	100%
Retirement plan option	38%	70%	90%	83%

5B.3.3. (Level 3) Staff are given feedback that gives examples of best practice at least twice a month.

To be verified for criterion 5B.3.3, which requires that staff be given examples of best practice at least twice a month, programs must achieve a score of 5 or higher on the PAS. As mentioned previously, overall scores on measurement tools are unlikely to provide a good indicator of specific practice, and as such, the validation study sought data regarding this practice through administrator interviews. Specifically, administrators were asked if staff regularly engaged in sharing and/or demonstration of best practice, and if so, how often.

Relatively few administrators, regardless of level, indicated that formal sharing or demonstration of best practice took place on at least twice a month (Table 97). To be coded as meeting this requirement, the practice needed to be sufficiently institutionalized, with time set aside for it or a routine vehicle through which it was provided, such as meeting or mentoring session. Anecdotally, a number of interviewed educators indicated that the sharing of best practices was more likely to be occurring on a monthly basis, as part of monthly meetings, while others indicated that it was occurring, but on a more informal basis. For the purposes of the MA QRIS, reliable

measurement of this criterion is likely to be challenging, and MA EEC may want to consider dropping this criterion as unmeasurable or otherwise revising it to require a more readily identifiable and directly measurable practice. A similar criterion at Level 4 (5B.4.2) would also benefit from similar review and clarification.

Table 97. Programs meeting measured aspects of criterion 5B.3.3, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Teachers participate in sharing and/or demonstration of best practice more than twice each month	15%	23%	20%	50%

5B.3.4. (Level 3) Program has a system to support the career development of staff through a career ladder (e.g., regularly scheduled time to meet with a supervisor or mentor to monitor progress towards career goals).

Criterion 5B.3.4 is also assessed using a program’s overall PAS score, which assesses overall program administration but does not provide much information in relation to the specific practices referenced in the criterion. For the validation study, the program questionnaire asked administrators to indicate whether they had any career ladder supports in place, including regularly scheduled time for educators to meet with a supervisor or mentor to monitor progress towards career goals; requirement that educators with less than an AA are enrolled in formal education; funding or scholarships to take college courses; and release time to participate in college courses.

Overall, nearly all programs at Level 2 and all programs at Levels 3 and 4 had at least one of these supports in place, as did 79 percent of programs at Level 1 (Table 98). MA EEC may want to consider changing its verification approach to require documentation of career ladder supports, with examples of career ladder supports that MA EEC considers acceptable, and depending on the types of number of supports required, may want to consider moving this criterion to Level 2. The specific types of career ladder supports offered by programs, by level, are shown in Table 99.

Table 98. Programs meeting measured aspects of criterion 5B.3.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program offers at least one of the following - regularly scheduled time to meet with a supervisor or mentor to monitor progress towards career goals; requirement that educators with less than an AA are enrolled in formal education; funding or scholarships to take college courses; release time to participate in college courses	79%	95%	100%	100%

Table 99. Career ladder supports, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Funding or scholarships to take college courses	54%	70%	93%	100%
Scheduled time to meet with a supervisor or mentor to monitor progress towards career goals	54%	68%	90%	100%
Release time to participate in college courses	36%	63%	41%	50%
Requirement that educators with less than an AA are enrolled in college courses	5%	48%	49%	83%

5B.3.5. (Level 3) Staff salary scales reflect the educational levels, experience, and performance levels, as determined by the annual evaluation of the staff members, and is comparable with the current wage level of others in the community with the same levels of education.

Criterion 5B.3.5 is also verified using a program’s overall PAS score, which assesses overall program administration but does not provide information about specific practice. For the validation study, program questionnaires asked administrators what determines salary levels for new employees and salary increases for existing staff. The aspect of the criterion that addresses comparability of current wage levels of others in the community with the same levels of education was not measurable and was not able to be considered for the validation study. MA EEC may want to consider revising the criterion to remove this language to ensure the criterion is fully measurable.

As indicated in Table 100, across all levels, nearly all programs were taking education or EEC certification into account when setting salaries for new employees, ranging from 92 percent at Level 1 to 100 percent at Level 4. Experience was taken into account when setting salaries for new employees in 80 percent or more of programs at Levels 1, 2, and 4, but only about half of programs at Level 3. Level 3 and Level 4 programs were also much less likely to report taking performance into account when setting salaries for existing employees.

Table 100. Programs meeting measured aspects of criterion 5B.3.5, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5B.3.5	54%	45%	22%	50%
Program takes education or EEC certification into account when setting salaries for new employees	92%	95%	95%	100%
Program takes experience into account when setting salaries for new employees	82%	80%	54%	83%
Program takes education or EEC certification into account when determining increases for existing staff	62%	73%	73%	100%
Program takes performance into account when determining increases for existing staff	77%	65%	46%	50%

5B.4.1. (Level 4) Program offers a benefit package that includes vacation, sick time, and health insurance.

To meet criterion 5B.4.1, programs must provide a written policy showing that they offer a benefits package that includes vacation time, sick time, and health insurance. This criterion is part of a series of two criteria, the other being a Level 3 requirement that programs offer at least one benefit, which may include any of these three benefits or tuition/PD reimbursement or a retirement option.

This benefits package was common at Levels 2 and above (Table 101). Vacation and sick time tended to commonly be offered regardless of level, whereas health insurance was much less common at Level 1. As mentioned previously, for purposes of the MA QRIS, it may be beneficial to clarify expectations regarding who must receive these benefits (i.e. whether it must include both full- and part-time staff).

Table 101. Programs meeting measured aspects of criterion 5B.4.1, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5B.4.1	46%	88%	90%	100%
Program offers vacation time to teachers (full- and/or part-time)	90%	100%	95%	100%
Program offers sick time to teachers (full- and/or part-time)	90%	90%	98%	100%
Program offers fully or partially paid health insurance to teachers (full- and/or part-time)	56%	93%	95%	100%

5B.4.2. (Level 4) Staff are provided ongoing mentoring that includes demonstration of best practices on a weekly basis.

Programs are verified as meeting criterion 5B.4.2 if teaching staff have regular opportunities to engage in reflecting teaching practice, peer groups, and coaching and mentoring. In administrator interviews, five of the six administrators at Level 4 programs reported that their teaching staff engaged in all three of these activities, and four of the six indicated that these activities took place at least monthly (Table 102). This was much less common among programs at Levels 1, 2, and 3. Although many programs at these levels engaged in some of these practices, few engaged in all three. Similar to criterion 5B.3.3, this criterion could potentially benefit from further clarification of MA EEC's expectations regarding the level and frequency of practice required to be verified as meeting the criterion.

Table 102. Programs meeting measured aspects of criterion 5B.4.2, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program meets all measured aspects of criterion 5B.4.2	3%	23%	20%	67%
Program offers peer groups, mentoring/coaching, AND opportunities to share or demonstrate effective teaching practices	13%	30%	37%	83%
Teachers participate in sharing/demonstration of best practice	54%	80%	68%	100%
Teachers participate in peer groups	44%	50%	51%	83%
Teachers participate in mentoring or coaching	33%	60%	76%	83%
Peer groups, mentoring/coaching, and opportunities to share or demonstrate effective teaching practices ALL take place monthly	3%	23%	20%	67%

5B.4.3. (Level 4) Program demonstrates systematic opportunities for teachers to engage in reflective teaching practices through the use of peer groups, coaches and/or mentors.

The verification for criterion 5B.4.3 is identical to 5B.4.2 (i.e. MA EEC verifies that there are regular opportunities for teaching staff to engage in reflective teaching practices, peer groups, and coaching/mentoring). As such, for practical purposes criterion 5B.4.3 duplicates criterion 5B.4.2, and MA EEC may want to consider dropping this criterion or otherwise modifying its verification to differentiate it from criterion 5B.4.3.

5B.4.4. (Level 4) Program has an incentive program that rewards each educator who achieves the next step on the career ladder.

Finally, at Level 4 programs are required to have an incentive program that rewards each educator who achieves the next step on the career ladder, which for QRIS purposes is verified based on a program’s overall PAS score (7 or higher). Because an overall PAS score assesses overall program administration but not the specific practice referenced in the criterion, for the validation study, the program questionnaire asked administrators to indicate whether they had any career ladder incentives in place. These included: a pay scale based on professional qualifications, greater responsibility with increasing education and credentials, and/or more senior job titles for educators obtaining higher credentials. MA EEC may want to consider changing its verification approach to require documentation of career ladder supports, with examples of career ladder incentives that would be considered meeting the criterion.

Overall, nearly all programs at Level 2 and above indicated that they had at least one of the incentives listed on the questionnaire in place, as did 82 percent of programs at Level 1 (Table 103). The most commonly reported career ladder incentive was a pay scale based on professional qualifications (Table 104). Greater responsibility with increasing education and credentials, merit-based pay increases, and more senior job titles for educators obtaining higher credentials was generally less common. The one exception was with regard to Level 4 programs, five out of six of which (83 percent) reported that they offered educators greater responsibility with increasing education and credentials.

Table 103. Programs meeting measured aspects of criterion 5B.4.4, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Program has at least one of the following — a pay scale based on professional qualifications, merit-based pay increases, greater responsibility with increasing education and credentials, and/or more senior job titles for educators obtaining higher credentials	82%	93%	93%	100%

Table 104. Career ladder incentives in place at programs, by MA QRIS Level

	Level 1	Level 2	Level 3	Level 4
Pay scale based on professional qualifications	56%	88%	90%	100%
Greater responsibility with increasing education and credentials	59%	55%	37%	83%
Merit-based pay increases	44%	48%	29%	50%
More senior job titles for educators obtaining higher credentials	21%	30%	32%	50%

Summary and implications

The data suggest that MA QRIS needs revisions in terms of clarification of quality criteria and the verification processes. A thorough review of the MA QRIS criteria and verification processes and study data indicated several overarching themes regarding system criteria these themes include: Better alignment of QRIS criteria and verification, and distribution of a single document outlining both criteria and verification; Clarification that CEUs or credits are needed for MA QRIS trainings; Timeframes for the trainings to reinforce practices; Reduction of memorandums of understanding as barriers to progresses for programs; Limitation of the use of overall scales scores for verification and increased focus on specific subscale/items that are more relevant; Reduction of compound criteria; Incorporation of Continuous Quality Improvement Plans; Consolidation of PD requirement into the Professional Development Quality Standard; and Greater consistency between MA QRIS requirements and other credentialing systems and initiatives. These revisions alone would refine the system and add to the clarity and usability of the system.

Section II: Statistical Analyses of MA QRIS Standards and Criteria

In addition to the descriptive analyses presented in the previous section, statistical analyses were also undertaken to examine the ability of individual criteria to differentiate among MA QRIS Levels, and to examine the relations among the various MA QRIS criteria. It is important to note that the MA QRIS is not meant to function as a scale or assessment tool, and as such, it not designed to have the internal constructs expected in a psychometric measure. Since the MA QRIS is a progressive system in which quality builds, it is expected that there will be less distinction among criteria at Level 2, which all programs at Level 2, 3, and 4 should be meeting, versus criteria at Level 3, which typically only Level 3 and Level 4 programs should be meeting. It is also anticipated that the proportion of programs meeting Level 2 criteria should at least be distinct from the proportion of programs at Level 1 meeting the criteria. It is important to note that the MA QRIS is a voluntary system and subsequently there are programs that may belong at higher levels but have not actively engaged with the system. This influences associations among criteria at the various MA QRIS Levels. With this in mind, it is helpful to examine the associations among criteria as a means of identifying the criteria that do not statistically distinguish among levels of the MA QRIS as well as the placement of specific criteria within the various quality levels.

Criteria differentiation among MA QRIS levels

To examine whether significant differences existed among criteria by MA QRIS Level, a series of ANOVAs were conducted for each criterion, comparing the proportion of programs at each level that met a given criterion. If a criterion differentiates among levels, we would expect there would be significant differences in the proportion of programs meeting each criterion among the MA QRIS Levels. Results indicated proportional differences by MA QRIS Level, in the direction expected, for most of the MA QRIS criteria. There were however, some criteria in which statistical differences among MA QRIS Levels were not found. In some cases, the distinction between levels was approaching significance, and in other cases there was no clear difference found. In circumstances in which criteria were approaching significance among levels, a modification to the criteria might produce the intended result (see Descriptive analyses of MA QRIS Standards, Criteria, and Verification Section for modification recommendations). In instances in which no empirical distinctions were be found, MA EEC may decide to either drop the criteria or retain the criteria due to the criteria’s policy implications and field relevance versus its ability to distinguish among MAQRIS Levels. Table 105 presents the criteria and associated statistics by MA QRIS Level for criteria in which a significant difference was **NOT** noted by level.

Table 105. QRIS Criteria that do not empirically and significantly differentiate among MA Levels

QRIS Standard/ Level/ Number	Criteria	F Value	Significance
Standard 1: Curriculum and Learning: 1A. Curriculum, Assessment, and Diversity			
1A.2.1 (Level 2)	Educators demonstrate completion of professional development in curriculum, screening tools and formative assessment.	2.35	p=.08
Standard 1: Curriculum and Learning: 1B. Teacher-Child Relationship and Interactions			
1B.3.2 (Level 3)	Educators are provided with opportunities to use outside consultants to staff with expertise in the age of the children served to assist them in implementing strategies that support positive relationships/interactions and prevention/interventions	2.24	p=.09
Standard 2: Safe, Healthy Indoor and Outdoor Environments			
2A.2.2 (Level 2)	Demonstrates healthy, safe and clean indoor and outdoor environments	1.64	p=.18
Standard 3: Workforce Qualifications and Professional Development: 3A. Designated Program Administrator Qualifications and Professional Development			

3A.2.1 (Level 2)	Program administrator has a Child Development Associate (CDA) Credential for the age of the children served OR higher OR is enrolled in a program leading to an Associate or Bachelor's Degree in early childhood	.91	p=.44
3A.2.3 (Level 2)	Program administrator has a minimum of 3 college credits in administration and management and 12 college credits in early childhood education/child development/special education and 2 years of experience as an administrator.	.139	p=.94
3A.2.5 (Level 2)	Program administrator has received professional development in supervision of adults and strategies for working with adults.	2.55	p=.06
3A.3.2 (Level 3)	Program administrator has at least 9 credit-bearing hours of specialized college-level course work in administration, leadership, and management.	.27	p=.85
3A.3.3 (Level 3)	Program administrator has at least 24 credit-bearing hours of specialized college-level course work in early childhood education, child development, elementary education, or early childhood special education OR documents that a plan is in place to meet requirements.	1.64	p=.18
3A.4.1 (Level 4)	Program administrator has a least 9 credit-baring hours of specialized college-level coursework in administration, leadership, and management	.27	p=.85
3A.4.2 (Level 4)	Program administrator has at least 24 credit-bearing hours of specialized college-level coursework in early childhood education, child development, elementary education, or early childhood special education.	1.53	p=.21
3A.4.3 (Level 4)	Program administrator has a minimum of 5 years of experience as an administrator.	.04	p=.98
Standard 4: Family and Community Engagement			
4A.2.2 (Level 2)	Programs offer opportunities for parents to meet with classroom staff, at least monthly.	1.33	p=.27
4A.3.1 (Level 3)	A daily two-way communication system is available between educators and families through a variety of means.	.49	p=.69
4A.3.3 (Level 3)	Program ensures that there are translators available, as needed, at meetings, workshops, and conferences to ensure strong communication between the program and families.	1.16	p=.33
4A.3.4 (Level 3)	Program representative(s) participate in local community group work that is related to early childhood, and the cultural groups served by the program and/or family support.	1.98	p=.12
Standard 5: Leadership, Management, and Administration: 5A. Leadership, Management, and Administration			
5A.2.2 (Level 2)	Program has a written admissions policy that promotes an awareness of and respect for differences among children and families, a respect for the child and their family's culture and language, and is responsive to the inclusion of a variety of learning differences.	1.00	p=.39
5A.2.4 (Level 2)	Staff are paid for planning time.	2.29	p=.08
5A.3.1 (Level 3)	Program tracks and monitors absences of individual children and contacts families when children are absent more than 20% in a month.	.211	p=.89
5A.4.3 (Level 4)	An outside audit is conducted annually by a certified public accountant.	2.24	p=.08
Standard 5: Leadership, Management, and Administration: 5B. Supervision			

5B.2.1 (Level 2)	Program provides recognition for staff in annual evaluation as well in public forum, as appropriate.	1.40	p=.25
5B.3.1 (Level3)	Program uses at least 3 types of internal communication on a monthly basis to inform staff of program activities, policies etc.	.05	p=.99
5B.3.2 (Level 3)	Staff receive at least one benefit.	.739	p=.53
5B.3.3 (Level 3)	Staff are given feedback that gives examples of best practices at least twice a month.	1.31	p=.27
5B.4.4 (Level 4)	Program has an incentive program that rewards each educator that achieves the next step on the career ladder.	1.29	p=.28

Based on the chart, Standard 3: Workforce Qualifications and Professional Development: 3A. Designated Program Administrator Qualifications and Professional Development is the MA QRIS Quality Standard whose criteria showed the least distinction among MA QRIS Levels and is subsequently an area of particular need for refinement. Conversely, the criteria that comprised Standard 3: Workforce Qualifications and Professional Development: 3B. Program Staff Qualifications and Professional Development consistently showed proportional differences in the number of programs meeting the criteria by MA QRIS Level.

Associations among criteria within a given MA QRIS Quality Standard

In addition to examining criteria that significantly distinguish among MA QRIS Levels, we examined the relation among criteria within the various MA QRIS quality standards. Based on how the MA QRIS criteria are constructed, it is expected that the individual criteria that comprise the quality standards at each level will have some relationship with one another. Table 106 depicts criteria that have relationships that are inconsistent with their placement within the MA QRIS and subsequently may serve as barriers to programs' progress or fail to distinguish between lower and higher levels of the MA QRIS. As such, they represent potential candidates for movement within the MA QRIS system.

Table 106. Criteria with inconsistent relationships within each quality standard

	Criteria	Relationships
Standard 1 Curriculum and Learning: 1A. Curriculum, Assessment, and Diversity		
1A.2.2 (Level 2)	Materials reflect the language and culture of the children in the classroom, their communities, and represent the diversity of society.	Significantly related to Level 3 criteria but shows limited relations with other Level 2 criteria for this standard.
Standard 1 Curriculum and Learning: 1B. Teacher-Child Relationship and Interactions		
1B.3.2 (Level 3)	Educators are provided with opportunities to use outside consultants to staff with expertise in the age of the children served to assist them in implementing strategies that support positive relationships/interactions and prevention/interventions.	Shows limited relation to other criteria for this standard.
Standard 2: Safe, Healthy Indoor and Outdoor Environments		
2A.3.2 (Level 3)	Staff are trained in how to work with children with special diets, allergies, and specialized feeding issues.	Significantly related to Level 2 criteria but shows limited relation to other Level 3 criteria for this standard.

Standard 3: Workforce Qualifications and Professional Development: 3A. Designated Program Administrator Qualifications and Professional Development		
3A.2.3 (Level 2)	Program administrator has a minimum of 3 college credits in administration and management, 12 college credits in early childhood education/child development/special education, and 2 years of experience as an administrator.	Significantly related to Level 3 criteria but shows limited relation to other Level 2 criteria for this standard.
Standard 4: Family and Community Engagement		
4A.2.1 (Level 2)	Program completes Strengthening Families Self-Assessment and uses data to engage in continuous quality improvement.	Significantly related to Level 3 criteria but shows limited relation to other Level 2 criteria for this standard.
4A.2.4 (Level 2)	Program maintains ongoing communication with the school/early intervention program, CFCE grantee, mental health providers to facilitate collaboration and coordination of services that support children and families.	Significantly related to Level 3 criteria but shows limited relation to other Level 2 criteria for this standard.
4A.2.5 (Level 2)	Program participates in community events significantly related to criteria at Level 3 but shows little relationship to Level 2.	Significantly related to Level 3 criteria but shows limited relation to other Level 2 criteria for this standard.
4A.3.1 (Level 3)	A daily two-way communication system is available between educators and families through a variety of means.	Significantly related to Level 2 criteria but shows limited relation to other Level 3 criteria for this standard.

Summary and implications

Data indicated that for 68% of criteria there were significant differences in the proportion of programs meeting each criterion by MA QRIS Levels. For an additional 10% of criteria, the significant difference in the proportion of programs meeting the criteria by MA QRIS Level was approaching significance. This suggests that most criteria are distinguishing aspects of quality by level. For criteria in which no significant differences in the proportion of programs meeting the criteria by level were found, data indicated that for some criteria almost all programs, regardless of verified MA QRIS Level, were meeting the requirements and for other criteria few programs were meeting the requirements, suggesting the criterion is functioning as a barrier. For criteria that most programs are meeting, researchers recommend dropping them from the system, as they do not appear to distinguish quality by MA QRIS Level. For criteria that large numbers of programs are not meeting, researchers suggest modifying or relaxing standards. Additionally, analyses indicates that some criteria may be better suited at other QRIS Levels, based on statistical associations among criteria by Quality Standards and QRIS Levels.

Section III: Program Quality and the Massachusetts QRIS

The Massachusetts QRIS is designed to assess and improve the quality of early education and care in Massachusetts, and a central feature of the system involves the assignment of levels to participating programs. To determine whether the system and its levels meaningfully distinguish the quality, the validation study investigated whether programs in different MA QRIS Levels showed distinctions in observed classroom quality. Observed classroom quality was assessed by reliable raters using two measures:

- The Early Childhood Environment Rating Scale-Revised (ECERS-R) and the Infant and Toddler Environment Rating Scale-Revised (ITERS-R), which assess global quality in preschool and infant and toddler classrooms, respectively.
- The Arnett Caregiver Interaction Scale (Arnett CIS), which assesses the quality of caregiver interaction.

These tools have been widely used in QRIS validation studies, as well as in early education and care quality studies more broadly.

ERS cut-offs and MA QRIS cut-offs

The Early Childhood Environment Rating Scale-Revised (ECERS-R) and the Infant and Toddler Environment Rating Scale-Revised (ITERS-R) — referred to collectively as the Environment Rating Scales or the ERS — assess the quality of the overall classroom environment as well as the quality of that environment across seven subscales: Space and Furnishing, Personal Care Routines, Language and Reasoning (or in the case of the ITERS, Listening and Talking), Activities, Interaction, Program Structure, and Parents and Staff. The ERS scales have been used extensively in a variety of research and evaluation projects. Scores range from 1 to 7. A score of 1 represents Inadequate care, a score of 3 represents Minimal care, a score of 5 represents Good care, and a score of 7 represents Excellent care. ERS cut-off points are distinct from those used for the different levels of the MA QRIS and therefore the data are presented both in terms of ERS cut-offs (see Tables 109, 111) as well as MA QRIS cut-offs to represent programs' compliance with meeting MA QRIS level criteria (Tables 113, 114). The decision to use diverse cut-off points for the MA QRIS in contrast to the ERS quality benchmarks was based, in part, on data gathered from the MA QRIS Reliable Rater observations during the 2011–2013 time period, as well as research^{42, 43} on the aspects of quality that are most salient in predicting child development. MA QRIS Reliable Rater data suggested a need for the MA QRIS to make accommodations in the MA QRIS requirements to meet the unique needs of programs in MA and to create a progressive system that facilitates the movement of programs from self-assessed ERS quality requirements (Level 2) to a program quality observation verified by a Program Quality Specialist visit (Level 3). To provide a reference point, the relationship between ERS benchmarks and the MA QRIS ERS Criteria is presented in Table 107. Please note that the ERS is not required for MA Licensing and subsequently is not included in the requirements for Level 1.

⁴² c. Howes, M. Burchinal, R. Pianta, D. Bryant, D. Early, R. Clifford, et al. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, 23 pp. 27-50.

⁴³ A.J. Mashburn, R.C. Pianta, B.K. Hamre, J.T. Downer, O.A. Barbarin, D. Bryant, et al. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79, 732-749.

Table 107. ERS scores in relation to MA QRIS ERS criteria by MA QRIS level

	ERS Score 2-2.94 (between Inadequate & Minimal)	ERS Score 3-3.94 (Minimal)	ERS Score 4-4.94 (between Minimal and Good)	ERS Score 5-5.94 (Good)
Space and Furnishings	Level 2	Level 3	Level 4	—
Personal Care Routines	Level 2	Level 3	Level 4	—
Language Talking/Language Reasoning	---	Level 2	Level 3	Level 4
Interactions	---	Level 2	Level 3	Level 4
Activities	---	Level 2	Level 3	Level 4
Program Structure	---	Level 2	Level 3	Level 4
Parents and Staff	---	Level 2	Level 3	Level 4
Overall Average Item Score	---	Level 2 (minimum of 3.5 required)	Level 3 (minimum of 4.5 required)	Level 4 (minimum of 5.5 required)

Observed classroom quality and ERS cut-offs

In general, observed classroom quality, as measured by the ECERS-R and ITERS-R, varied across programs in the study, with the average quality (\bar{x} = 4.23 on ECERS-R and \bar{x} =3.75 for the ITERS-R) for the overall sample, which is between the Minimal to Good quality ranges, as indicated by the ERS benchmarks.

Observed quality preschool and ERS cut-offs

ECERS-R Overall Average Item Score for the sample ranged from 2.05 to 5.88, with a mean of 4.23. Scores on the seven subscales of the ECERS-R varied, with the highest mean score on the Parents and Staff Subscale (\bar{x} =5.57) and the lowest mean on the Personal Care Routines Subscale (\bar{x} =2.85). Table 108 presents the ranges, means, and standards deviations for the subscales and overall scores on the ECERS-R. It should be noted that the mean average score on the ECERS-R is higher than the average score reported by in other states.^{44,45}

⁴⁴ Lahti, M., Elickerb, J., Zellman, G., & Fiene, R., (2014.) Approaches to validating child care quality rating and improvement systems (QRIS): Results from two states with similar QRIS type designs. *Early Childhood Research Quarterly*, 30, 280-290

⁴⁵ Tout, K., Starr, R., Isner, T., Cleveland, J., Albertson-Junkans, I., Soli, M., & Quinn, K., (2011). Evaluation of Parent Aware: Minnesota's Quality Rating and Improvement System Pilot, Child Trends: Minneapolis, MN.

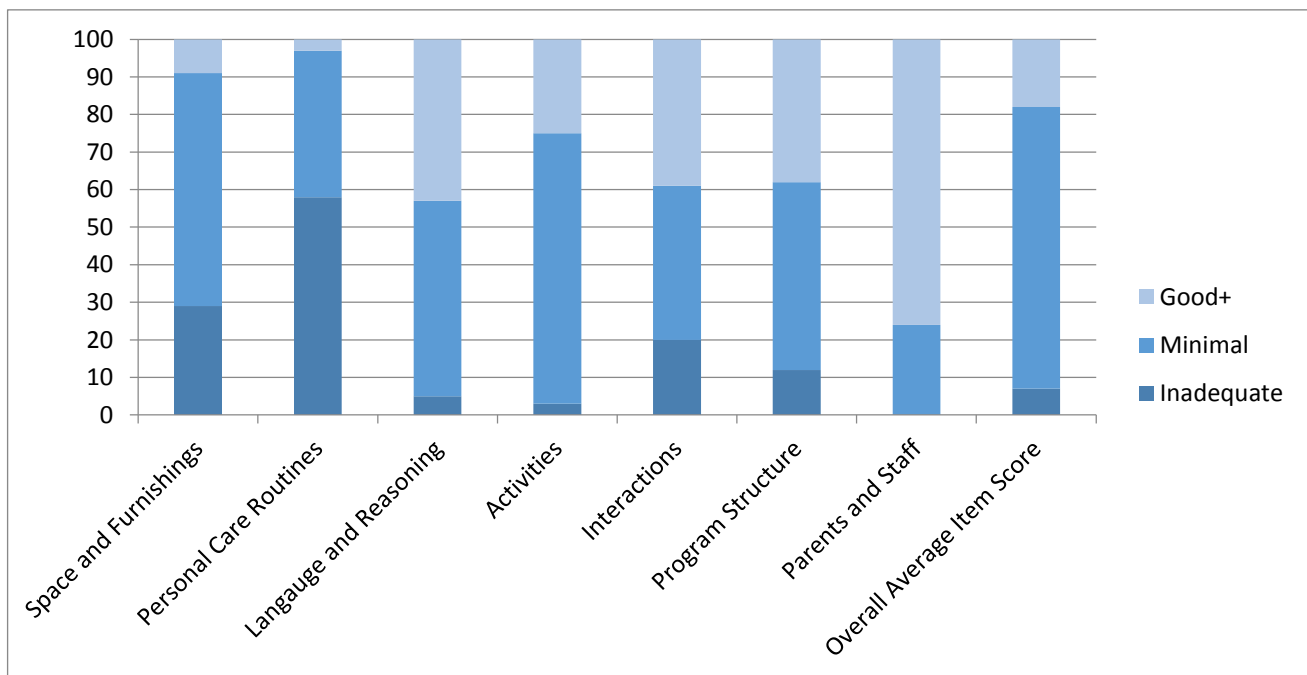
Table 108. Average ECERS-R Subscale ranges, means, and standard deviations

Subscale	Minimum	Maximum	Mean	Standard Deviation
Space and Furnishings	1.63	5.88	3.56	.90
Personal Care Routines	1.17	5.33	2.85	.79
Language and Reasoning	1.25	7.00	4.72	1.12
Activities	2.20	6.90	4.39	.95
Interactions	1.00	7.00	4.37	1.63
Program Structure	1.00	7.00	4.51	1.38
Parents and Staff	3.00	7.00	5.57	.92
Overall Average Item Score	2.05	5.88	4.23	.82

As can be seen in Table 109Table, the Language and Reasoning, Interactions, and Program Structure Subscales had scores with the greatest variability and standard deviations over 1 (based on a 7-point scale). Few classrooms (7 percent) had an Overall Average Item Score in the Inadequate range, suggesting that most of the preschool programs in the Massachusetts QRIS have at least a Minimal range of quality, regardless of level. Six classrooms however, did receive a score of a 3 or less (Inadequate range) across all subscales, except the Parents and Staff Subscale. Five of these classrooms were from Level 1 programs and 1 classroom was from a Level 2program.

Averages only offer part of the picture of quality. The distribution and frequencies of scores offer a greater understanding of the quality of programs and the percent of programs reaching the quality benchmarks of the ECERS-R (see Figure 11).

Figure 11. Percent of programs meeting ECERS-R benchmarks



For the most part, preschool program quality was in the Minimal range across all subscales. Approximately, 75 percent of programs had overall quality ratings in the Minimal range. Of note, the Parents and Staff Subscale had a majority (76 percent) of programs fall in the Good+ range and conversely the subscale for Personal Care Routines had a majority of classrooms (58 percent) fall in the Inadequate range. The vast majority of programs (72 percent) also fell in the Minimal range for Activities, with about one-quarter reaching the ERS Good Benchmark. The Interaction subscale had the greatest variability, with 20 percent falling in the Inadequate range, 41 percent falling in the Minimal range, and 39 percent falling in the Good+ range.

The distribution of quality scores as measured on the ERS is presented below (see Table 109). Again, it is essential to stress that the MA QRIS standards for each level do NOT mirror the requirements for the MA QRIS Levels. The table is presented to enhance the understanding of where programs are falling on the ERS scale as a reference point and to inform policy.

Table 109. Percent of Preschool Programs at Meeting ERS Benchmarks

Subscale	Assigned MA QRIS Level	Inadequate %	Minimal %	Good+ %
Space and Furnishings	Level 1	39	56	5
	Level 2	36	59	5
	Level 3	17	73	10
Personal Care Routines	Level 1	72	25	3
	Level 2	69	31	0
	Level 3	37	58	5
Language and Reasoning	Level 1	8	64	28
	Level 2	5	55	40
	Level 3	2	35	63
Activities	Level 1	8	87	5
	Level 2	3	72	25
	Level 3	0	60	40
Interactions	Level 1	28	46	26
	Level 2	23	41	36
	Level 3	12	34	54
Program Structure	Level 1	18	56	26
	Level 2	18	46	36
	Level 3	2	49	49
Parents and Staff	Level 1	0	41	59
	Level 2	0	36	64
	Level 3	0	0	100
Overall Average Item Score	Level 1	18	79	3
	Level 2	5	68	16
	Level 3	0	71	29

Data indicate a progressive and step-wise pattern of data in which the percent of the programs in the Inadequate range decreased as MA QRIS Levels increased; and the percent of programs in the Good range increased along with an increase in MA QRIS Level. Differences can be seen between MA QRIS Levels 1 and 2 with a greater portion of Level 2 programs scoring in higher ranges on the ERS, despite the fact that Level 1 and Level 2 are both ERS self-assessed levels. Importantly the data pattern is progressive in nature, reflecting incremental improvements in quality with each MA QRIS Level. In order for MA QRIS Leveling to be effective, programs must be able to move between levels. Within this pattern of incremental improvement with each level, gaps in observed quality between neighboring levels are smaller than the gap between non-adjacent levels. Of particular note in the data is that 18 percent of Level 1 programs were rated by the Overall Average Item Score as providing Inadequate care. It is also important to note the substantive percentage of programs at all MA QRIS Levels rated in the Inadequate range for the Space and Furnishings, Personal Care Routines, and Interactions Subscales. To understand this better, this finding is explored in greater detail below in the Item and indicator analyses: Strengths, challenges and licensing Section.

Observed quality infant and toddler and ERS cut-offs

Similar to the preschool classrooms, infant and toddler classrooms also varied in quality. The Overall Average Item Score for the ITERS-R was 3.75 with a range of 2.38 to 5.50. Descriptive data for the seven subscales and the overall score are presented below (Table 110).

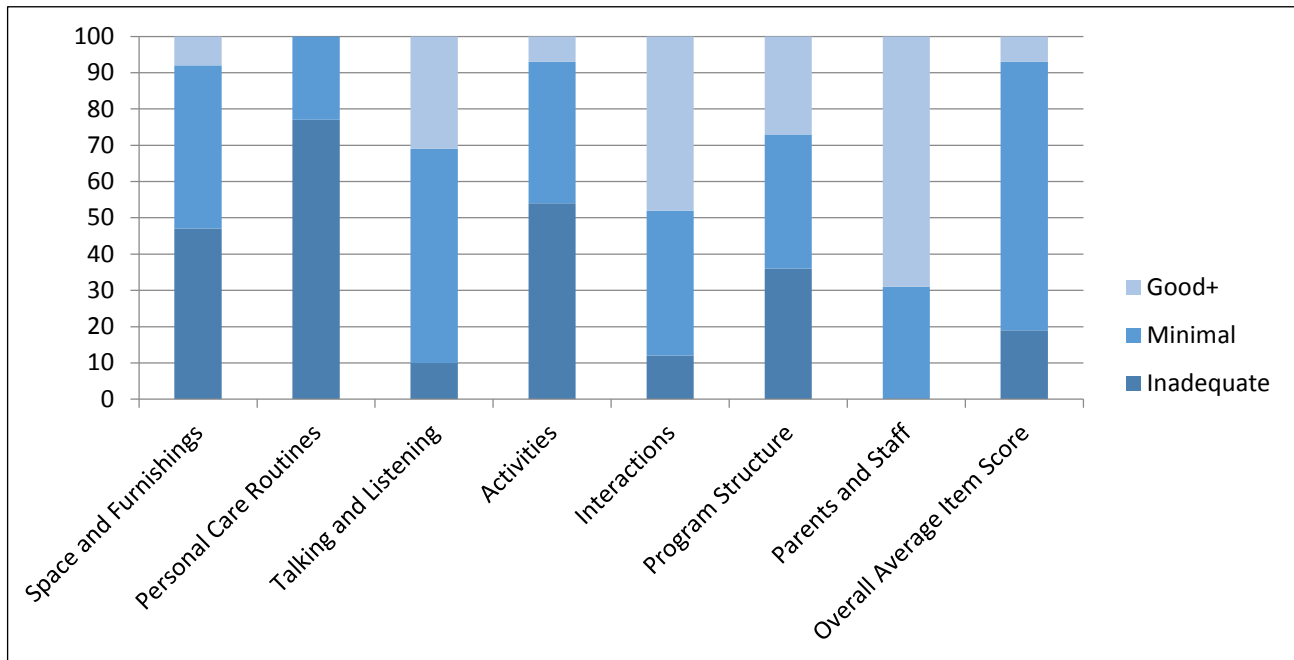
Table 110. Average ITERS-R Subscale ranges, means, and standard deviations

Subscale	Minimum	Maximum	Mean	Standard Deviation
Space and Furnishings	1.80	5.80	3.21	.89
Personal Care Routines	1.17	4.33	2.45	.73
Listening and Talking	2.00	7.00	4.31	1.08
Activities	1.38	5.78	3.07	1.06
Interactions	2.00	7.00	4.68	1.46
Program Structure	1.50	7.00	3.63	1.43
Parents and Staff	3.71	6.71	5.40	.80
Overall Average Item Score	2.38	5.50	3.75	.76

Consistent with the preschool classrooms, the Parents and Staff Subscale had the highest average score ($\bar{x}=5.40$) and the Personal Care Routines Subscale had the lowest average ($\bar{x}=2.45$). For the infant and toddler classrooms, Listening and Talking, Activities, Interactions, and Program Structure Subscales had the greatest variability in scores, with standard deviations over 1.

Figure 12 presents the frequencies of infant and toddler programs meeting ERS benchmarks.

Figure 12. Percent of programs meeting ITERS-R benchmarks



The vast majority of infant and toddler programs fell in the Minimal range (scores of 2.95-4.94) for the Overall Average Item Score. About 19 percent of programs fell in the Inadequate range for overall quality. For the Personal Care Routine Subscale (77 percent) and the Activities Subscale (54 percent), a majority of programs fell within the Inadequate range. For the Parents and Staff Subscale the vast majority of programs fell in the Good+ range. The Interaction Subscale (12 percent in Inadequate, 40 percent in Minimal range, and 48 percent in the Good+ range) and the Program Structure Subscale (36 percent in the Inadequate range, 37 percent in the Minimal range, and 27 percent in the Good+ range) had the greatest variability in scores with significant percentages of programs in each category.

Overall, about 19 percent of programs had classrooms rated as having below Minimal levels of quality. About 12 percent (n=9) of programs exhibited scores that were a 3 or below on all subscales of the ITERS-R, with the exception of Parents and Staff; six of these programs were in Level 1 and three were in Level 2. The percent of programs that met ERS quality benchmarks by MA QRIS Level is presented below (see Table 111). Consistent with preschool classrooms, the MA QRIS Levels standards for Infants and Toddlers do not consistently coincide with ERS benchmarks. Data are only presented as a means of comparison and for informing future policy.

Table 111. Percent of infant and toddler programs at meeting ERS benchmarks

Subscale	Assigned MA QRIS Level	Inadequate %	Minimal %	Good+ %
Space and Furnishings	Level 1	63	33	4
	Level 2	45	42	13
	Level 3	27	73	0
Personal Care Routines	Level 1	72	12	0
	Level 2	74	26	0

Subscale	Assigned MA QRIS Level	Inadequate %	Minimal %	Good+ %
	Level 3	80	20	0
Talking and Listening	Level 1	19	66	15
	Level 2	7	70	23
	Level 3	0	60	40
Activities	Level 1	70	26	4
	Level 2	45	49	6
	Level 3	40	53	7
Interactions	Level 1	15	52	33
	Level 2	13	39	48
	Level 3	7	26	67
Program Structure	Level 1	52	30	18
	Level 2	33	38	29
	Level 3	20	47	33
Parents and Staff	Level 1	0	52	42
	Level 2	0	29	71
	Level 3	0	0	100
Overall-Average Item Score	Level 1	33	63	4
	Level 2	13	77	10
	Level 3	0	93	7

A similar pattern of quality by level to that of the preschool classrooms is noted on some of the ITERS-R subscales for infant and toddler classrooms. For the Listening and Talking, Interactions, Program Structure, and Parents and Staff Subscales, fewer programs were in the Inadequate range as programs progressed through the MA QRIS Levels, and a greater percentage of classrooms were rated in the Good+ range as programs moved upward through the MA QRIS Levels. For the Space and Furnishings, Personal Care Routines, and Activities Subscales, the association between level and the percentage of programs in the Inadequate versus Good range was not consistently found. The data suggest very little difference among MA QRIS Levels regarding Personal Care Routines, with most programs falling in the Inadequate range. It is also noteworthy that a significant percentage of programs in the sample were rated in the Inadequate range of quality for Space and Furnishings, Personal Care Routines, Activities, and Program Structure, across all MA QRIS Levels. Additionally, about one-third (33 percent) of Level 1 programs appear to be functioning in the Inadequate range for Overall quality of care.

Observed classroom quality and MA QRIS ERS cut-offs

The above data represent the Inadequate, Minimal and Good cut-off points as determined by the authors of the ERS, and provide a source for national comparison among MA QRIS programs and programs participating in other research. As indicated above, the MA QRIS ERS criteria by level are distinct from the ERS benchmarks.

The current MA requirements for the ERS portion of the QRIS system are presented below (Table 112). The same ERS cut-off points exist across all types of programs (preschool, infant and toddler, family child care, and afterschool) participating in the MA QRIS. It is important to underscore that the MA QRIS system is a block system. As such, in addition to meeting level-specific ERS standards, programs are **also** required to meet the level-specific criterion of the Curriculum and Learning; Safe, Health Indoor and Outdoor Environments; Workforce Development and Professional Qualifications; Family and Community Engagement; and Leadership, Administration, and Management MA QRIS Quality Standards.

Table 112. MA QRIS ERS cut-offs for each level

	Level 1	Level 2	Level 3	Level 4
Space and Furnishings	Licensing Standards	2	3	4
Personal Care Routines		2	3	4
Language Talking/Language Reasoning		3	4	5
Interactions		3	4	5
Activities		3	4	5
Program Structure		3	4	5
Parents and Staff		3	3	4
Overall Average Item Score		3.5	4.5	5.5

As can be seen in Table 110, the Overall Average Item Score is higher than the requirements for individual subscales. As such, if a program meets only the MA QRIS minimal subscale score requirements across all subscales, it will not meet the requirement for the Overall Average Item Score.

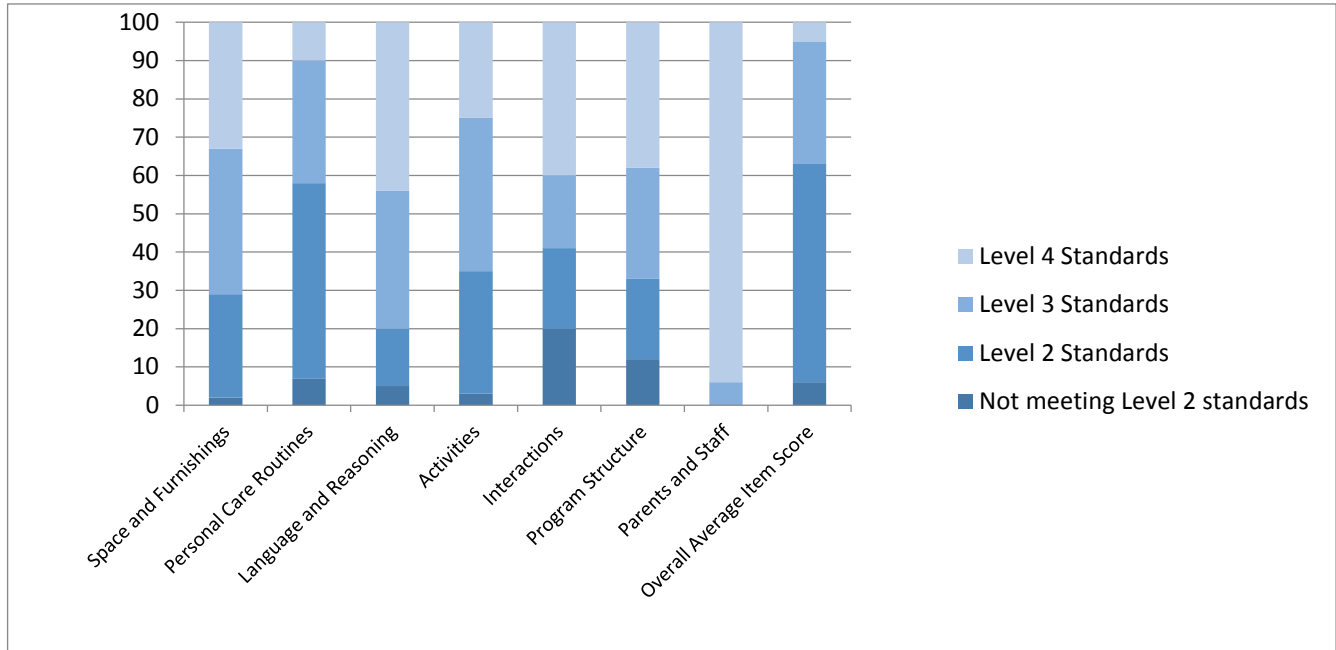
Observed preschool quality and MA QRIS cut-offs

The MA QRIS is a voluntary system that is a block system. As such, programs’ MA QRIS Level is not solely dependent on ERS scores. The ERS scores, however, provide important insight regarding the MA QRIS verification process, the quality of programs at Level 1 (entry point into system), and the needs of programs at all levels in order to progress within the system.

Figure 13 presents the percentage of programs meeting the MA QRIS benchmarks of quality, regardless of the programs’ assigned level. Since the ERS scores are not required for MA QRIS Level 1 programs and observations did not include a complete licensing review, programs that did not meet the ERS thresholds of the MA QRIS Level 2 were designated in the figure as “Not Meeting Level 2 Standards.” Hard cut-off points were taken to determine if programs met standards; observed quality scores were rounded to the nearest hundredth. As such, to meet a score of 3, programs had to achieve a score of at least 2.95; to meet a cut-off point of 3.5, programs had to achieve a score of 3.46; to meet the cut-off for a 4, programs had to receive an observed score of at least 3.95; to meet the cut-off for a 4.5, programs had to achieve a score of 4.46; to reach cut-off for a 5, programs had to achieve a score of at least 4.95; and to meet the cut-off for 5.5, programs had to achieve a score of 5.46.

The data presented in Table 113 represent the percent of programs meeting the MA QRIS ERS cut-offs for the entire sample of preschool classrooms participating in the study, regardless of level.

Figure 13. Preschool programs meeting MA QRIS cut-offs regardless of assigned level



Overall, about 39 percent of programs met or exceeded the MA QRIS Level 3 Overall Average Item Score cut-off for overall classroom quality. Of concern, is the large percent (20 percent) of programs not meeting the MA QRIS Level 2 standards for Interactions (ERS score of 2.95-3.94). Table 114 explores the percent of programs meeting MA QRIS ERS cut-off standards by the programs' MA QRIS verified level. The individual subscales of the ECERS-R are discussed below.

Table 113. Percent of preschool programs at MA QRIS cut-off points

Subscale	Assigned MA QRIS Level	Not Meeting Level 2	Meeting MA QRIS Level 2	Meeting MA QRIS Level 3	Meeting MA QRIS Level 4
Space and Furnishings	Level 1	5	33	49	13
	Level 2	0	36	28	36
	Level 3	0	17	39	44
Personal Care Routines	Level 1	8	64	23	5
	Level 2	13	56	26	5
	Level 3	2	34	44	20
Language and Reasoning	Level 1	8	31	33	28
	Level 2	5	10	54	31
	Level 3	2	7	27	64
Activities	Level 1	8	49	39	5
	Level 2	3	33	39	26
	Level 3	0	15	44	42
Interactions	Level 1	28	23	23	26
	Level 2	23	31	10	37
	Level 3	12	10	24	54
Program Structure	Level 1	18	39	18	26
	Level 2	18	13	33	36
	Level 3	2	12	37	49
Parents and Staff	Level 1	0	0	18	82
	Level 2	0	0	0	100
	Level 3	0	0	0	100
Overall	Level 1	15	69	15	0
	Level 2	3	62	33	3
	Level 3	0	44	42	15

Space and Furnishings. The MA QRIS requirement for Space and Furnishings is lower than for other subscales. As such, there is a disparity in the percentage of programs considered Inadequate and Minimal based on ERS benchmarks and the percentage of programs “Not meeting Level 2 standards.” Based on the MA QRIS cut-offs, all Level 2 programs met or exceeded Level 2 standards, and the majority (83 percent) of Level 3 programs met or exceeded Level 3 standards. Regardless, it is noteworthy that 17 percent of Level 3 programs did not meet Level 3 standards for Space and Furnishings.

Personal Care Routines. Similar to the Space and Furnishings Subscale, the standards for Personal Care Routines is lower than for other subscales, with Level 3 programs required to receive a score of a 3 (Minimal ERS score) or more. About 64 percent of Level 3 programs met or exceeded the Level 3 requirements, with 20 percent of these programs reaching Level 4 standards (average ERS score of 4, between Minimal and Good

quality). Most of the programs at Levels 1 and 2 met Level 2 standards. It is important to emphasize that when considering program needs for TA and professional development, Level 2 standards are below the standards for Minimal Quality as defined by the ERS. Also, a significant portion of programs at Level 3 are only meeting Level 2 standards (below Minimal standards for the ERS). As such, despite the lowered QRIS standards for Personal Care Routines, programs continue to face challenges in meeting Minimal standards for this subscale (see Item analyses below for more detail). It should be noted that this finding is not unique to Massachusetts' programs with multiple studies reporting low scores for Personal Care Routines.^{46, 48}

Language and Reasoning. By and large, most Level 3 programs met or exceeded the Level 3 standards of Language and Reasoning for the MA QRIS. Additionally, most programs at Level 1 and Level 2 met or exceeded MA QRIS Level 3 standards for this subscale. Despite this, Level 1 had approximately 8 percent of programs "Not meeting Level 2 standards" and about one-third (31 percent) meeting MA QRIS Level 2 standards, representing Minimal quality on this subscale as defined by the ERS.

Activities. The vast majority of programs (86 percent) at Level 3 met or exceeded standards for Activities. Most (49 percent) of the programs at Level 1 are meeting standards for Level 2 (Minimal Quality on the ERS), with 8 percent "Not meeting Level 2 Standards." At Level 2, the majority of programs are meeting or exceeding Level 2 standards (97 percent).

Interactions. Although most programs (78 percent) at Level 3 reached or surpassed Level 3 standards for Interactions, over one-fifth of Level 3 programs (22 percent) did not meet the MA QRIS Level 3 standards, with approximately 12 percent falling below Level 2 requirements (below Minimal as defined by the ERS). Over one-quarter of programs at Level 1 did "Not meet Standards for Level 2" (ERS scores in Inadequate) and approximately 23 percent of programs at Level 2 did "Not meet Level 2 standards" (ERS score in Inadequate range). This is significant given that research has identified teacher-child Interactions as important predictors of child outcomes.^{49, 50} It is important to note that Interactions Subscale had the greatest variability in scores among classrooms. Item analyses indicated large variability among all the individual items that comprise this subscale with standard deviations ranging from 1.79 to 2.09, across items.

Program Structure. Most of the Level 3 programs and Level 2 programs met or exceeded the Level 3 benchmarks for Program Structure for the MA QRIS. About 18 percent of Level 1 and Level 2 programs did "Not meet the Level 2 standards" for Program Structure.

Parents and Staff. The Parents and Staff subscale provided little distinctions for quality among levels. Similar to Space and Furnishings and Personal Care Routines, to receive a MA QRIS rating of a 3, programs need to average a 3 on the Parents and Staff Subscale of the ERS, which is the ERS Minimal range for quality.

Overall Average Item Score. Fifty-seven percent of programs at Level 3 obtained or exceeded Level 3 standards. Forty-four percent of Level 3 programs did not meet the Overall Average Item Score requirements, with scores for Personal Care Routines, in particular, bringing down item averages for Level 3. Level 1 had the greatest number of programs "Not meeting Level 2 QRIS Standards" with 15 percent not reaching an Overall Average Item Score of 3.5.

A review of program data indicated that of the Level 3 classrooms that did not meet Level 3 standards across the different subscales, many were the same program continually not meeting MA QRIS standards. Specifically,

⁴⁶ Nancy L. Marshall, Cindy L. Creps, Nancy R. Burstein, Frederic B. Glantz, Wendy Wagner Robeson, Steve Barnett. 2001. The Cost and Quality of Full Day, Year-round Early Care and Education in Massachusetts: Preschool Classrooms. Executive Summary. Wellesley Centers for Women and Abt Associates Inc.

⁴⁷ Helburn, Suzanne W. (Ed). (1995). Cost, Quality and Child Outcomes in Child Care Centers. Technical Report, Public Report, and Executive Summary: Colorado Univ., Denver. Dept. of Economics, Denver, Co.

⁴⁸ Early Child Care Research Network (2005). Child Care and Child Development: Results from the NICHD Study of Early Child Care and Youth Development, Guilford Press: New York, NY.

⁴⁹ Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. Child Development, 72(2), 625-638.

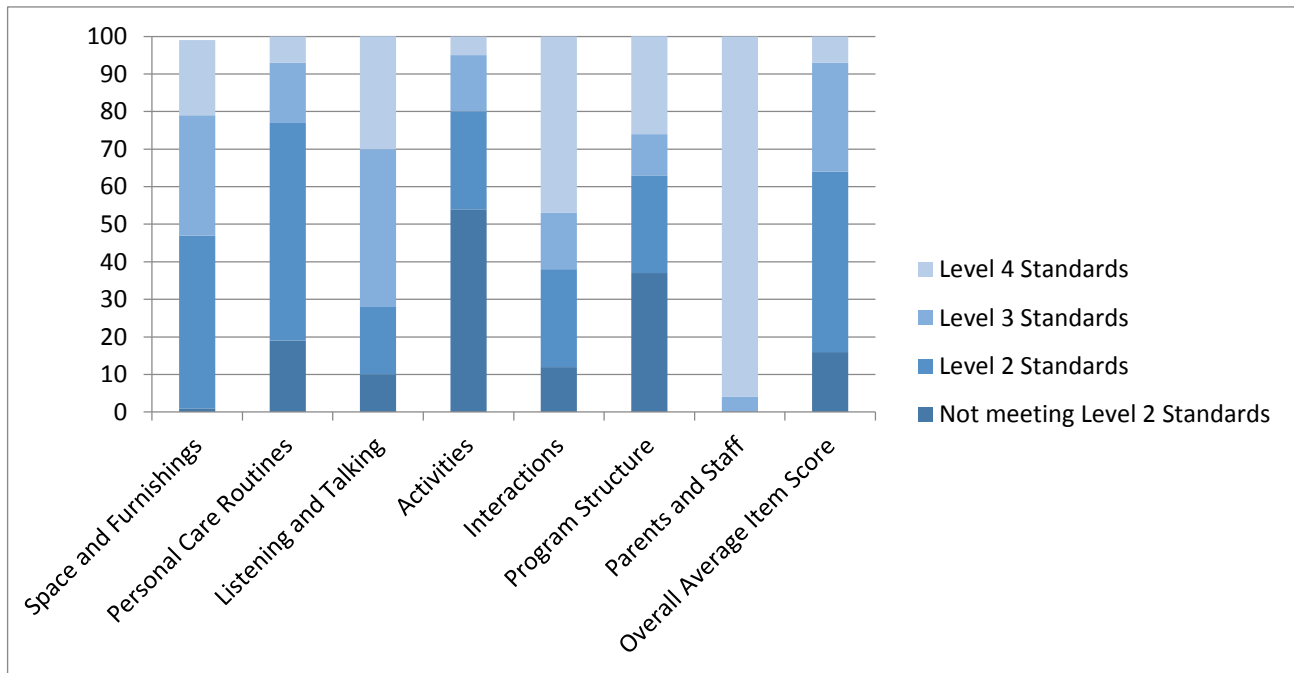
⁵⁰ Mashburn et al. 2008.

about 70 percent of the Level 3 programs who did not meet Level 3 requirements for one subscale also failed to meet Level 3 standards for another subscale. Of the 30 percent of Level 3 programs that did not meet Level 3 standards on only one subscale, the standard these programs consistently did not meet was the Personal Care Routine Subscale (see Item and indicator analyses: Strengths, challenges and licensing below for more detail about Personal Care Routines). This suggests that discrepancies between MA QRIS Level 3 requirements and the observed quality of Level 3 programs as measured in the study may represent the slide downwards of scores for specific programs. It is important to note that about half of the of the Level 3 programs had classrooms verified with data that were over 2 years old, and were subsequently verified with different and/or less defined verification and inter-rater reliability protocols, which have since been refined. Additionally, of the Level 3 programs that did not meet Level 3 standards, the programs were somewhat unevenly distributed among regions, with one region having a higher portion of Level 3 programs consistently not meeting Level 3 standards.

Observed infant and toddler quality and MA QRIS cut-offs.

Similar to preschool classrooms, researchers examined the percent of infant and toddler programs meeting MA QRIS quality standards, regardless of assigned level. Figure 14 presents the percent of infant and toddler programs that met MA QRIS benchmarks, regardless of assigned level.

Figure 14. Infant and toddler programs meeting MA QRIS cut-offs regardless of assigned level



As can be seen from Figure 14, observed program quality varied across all MA QRIS Levels, with the exception of the Parents and Staff subscale in which all programs were within the MA QRIS Level 3 and Level 4 range. Across all subscales, significant portions of programs did “Not meet Level 2 QRIS standards” for the Activities Subscale and the Program Structure Subscale. A detailed review of the data and ERS items found that a significant numbers of programs did not meet the Activities requirements for material accessibility for “much of the day.” A review of data and items for Program Structure found that programs were not regularly going outside for play and that they were not regularly meeting the requirements for free play taking place for “much of the day.” Not meeting “much of the day” reflects practices in which children are spending significant time in forced group activities that do not permit children to leave; or a significant amount of time transitioning, in contained spaces (e.g., exersaucers, swings) and/or waiting. As such, these programs do not reflect Developmentally Appropriate Practices for Infants and Toddlers which calls for children to have free choice; a limited amount of time in

restrictive and confined spaces; and a minimal amount of time in forced group and limited periods of waiting/transitioning. The table below (Table 114) presents the percentage of programs meeting MA QRIS ERS cut-offs for each subscale by the QRIS programs' assigned level.

Table 114. Percent of infant and toddler programs at MA QRIS cut-off points

Subscale	Assigned MA QRIS Level 1	Not Meeting Level 2	Meeting MA QRIS Level 2	Meeting MA QRIS Level 3	Meeting MA QRIS Level 4
Space and Furnishings	Level 1	4	60	33	4
	Level 2	0	45	23	32
	Level 3	0	27	53	20
Personal Care Routines	Level 1	13	68	15	4
	Level 2	19	55	23	3
	Level 3	33	47	7	13
Listening and Talking	Level 1	19	22	44	15
	Level 2	7	19	36	38
	Level 3	0	7	53	40
Activities	Level 1	74	22	0	4
	Level 2	45	29	19	7
	Level 3	40	27	27	7
Interactions	Level 1	15	37	15	33
	Level 2	13	19	19	48
	Level 3	7	20	7	66
Program Structure	Level 1	52	26	4	19
	Level 2	32	26	13	29
	Level 3	20	27	20	33
Parents and Staff	Level 1	0	0	11	89
	Level 2	0	0	0	100
	Level 3	0	0	0	100
Overall Average Item Score	Level 1	31	46	19	4
	Level 2	13	48	29	10
	Level 3	0	53	40	7

Space and Furnishings. Most of the Level 1 programs met Level 2 standards for Space and Furnishings with significant percentages meeting Level 3 and 4 standards; all Level 2 programs met or exceeded MA QRIS Level 2 standards. Although the majority of Level 3 program (63 percent) met or exceeded MA QRIS Level 3 standards, about 27 percent only met MA QRIS Level 2 requirements.

Personal Care Routines. Of all the subscales, this subscale had the greatest number of programs not meeting standards for their current QRIS levels, with most programs meeting only MA QRIS Level 2 standards and a

significant number of programs falling in the “Not Meeting Level 2 standards” range. It is important to note that similar to Space and Furnishings, the MA QRIS requirements for Personal Care Routines is lower than other subscales, and programs need to achieve a score of a 2 (ERS Inadequate range) to meet MA QRIS Level 2 standards. As such, these data indicate that the vast majority of programs do not meet the ERS Minimal range for Personal Care Routines, which reflects an overarching and systematic challenge with Personal Care Routines, regardless of programs’ level

Listening and Talking. Regarding Listening and Talking, Level 1 programs had the largest percentage of programs (19 percent) “Not meeting Level 2 standards.” Both the majority (74 percent) of Level 2 programs met standards for Level 3 or higher and the vast majority (93 percent) of programs at Level 3 met or exceeded MA QRIS standards for Level 3.

Activities. Unlike the other subscales, the Activities Subscale has a large percentages of programs that were “Not meeting Level 2 standards” across all QRIS levels (52 percent Level 1, 32 percent Level 2, and 20 percent Level 3). As indicated above, review of the data suggests that this pattern of scores is reflective of programs largely not meeting the ITERS-R standards of accessibility of materials for “much of the day” as opposed to a lack of Activities in the classroom.

Interactions. About 73 percent of Level 3 programs met or exceeded Level 3 requirements and 67 percent of Level 2 programs met or exceeded Level 3 standards for Interactions. About 37 percent of Level 1 programs met Level 2 standards (Minimal range on the ERS) and 15 percent for Level 1 programs did “Not meet Level 2 Standards” (Inadequate range on ERS). As noted above, Interactions between educators and children have been found to be particularly influential in promoting children’s outcomes.

Program Structure. Most of Level 1 programs (52 percent) and about one-third of Level 2 programs (32 percent) did “Not meet Level 2 Standards” for Program Structure. There was also large variability in scores among Level 3 programs. About 53 percent of Level 3 programs met or exceeded Level 3 standards. However, about one-quarter (27 percent) Level 3 programs that only met Level 2 standards (Minimal range for ERS) and 20 percent of Level 3 programs did “Not meet Level 2 Standards” (Inadequate range). Similar to the Activities Subscale, this reflects, in part, a high degree of structure and restriction in infant and toddler programs (see Item and indicator analyses: Strengths, challenges and licensing section below)

Parents and Staff. Similar to preschool classrooms, most classrooms exceeded or met standards for this subscale, which does not appear to distinguish between MA QRIS Levels. EEC may want to consider raising MA QRIS standards for this subscale which only requires an average score of a 3 to meet the Level 3 requirements.

Overall Average Item Score. Most (53 percent) of Level 3 programs did not meet Level 3 standards for the Overall Average Item Score. Similar to preschool programs, the Overall Average Item Score was brought down, in part, by the low scores on the Personal Care Routines Subscale (see Item and indicator analyses: Strengths, challenges and licensing section below). It is also noteworthy that about one-third (31 percent) of Level 1 programs did “Not meet QRIS Level 2 Standards” (ERS Inadequate range). This underscores the need for supports for programs at Level 1 in order to progress to Level 2. Additionally, this has implications not only for MA QRIS but for Licensing Standards and Guidelines for Care, as a large percentage of programs are not reaching Minimal standards as defined by the ERS.

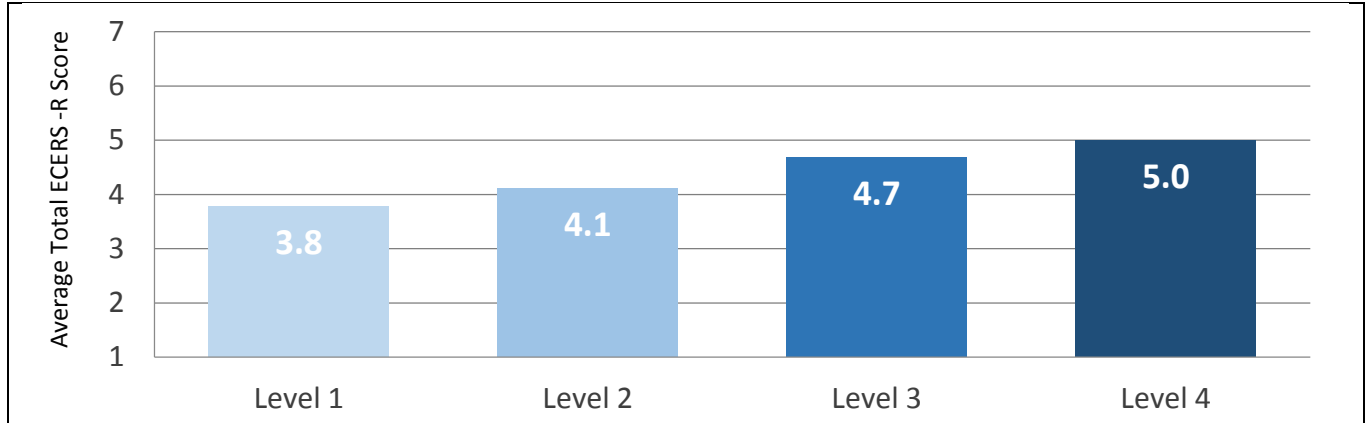
Significant differences in quality by MA QRIS Levels

Analyses indicated a significant difference in quality by verified MA QRIS Level, across both preschool and infant and toddler classrooms.

Preschool differences in quality by MA QRIS Levels

Figure 15 depicts the difference in average Overall Average Item Scores by assigned MA QRIS Levels for preschool classrooms.

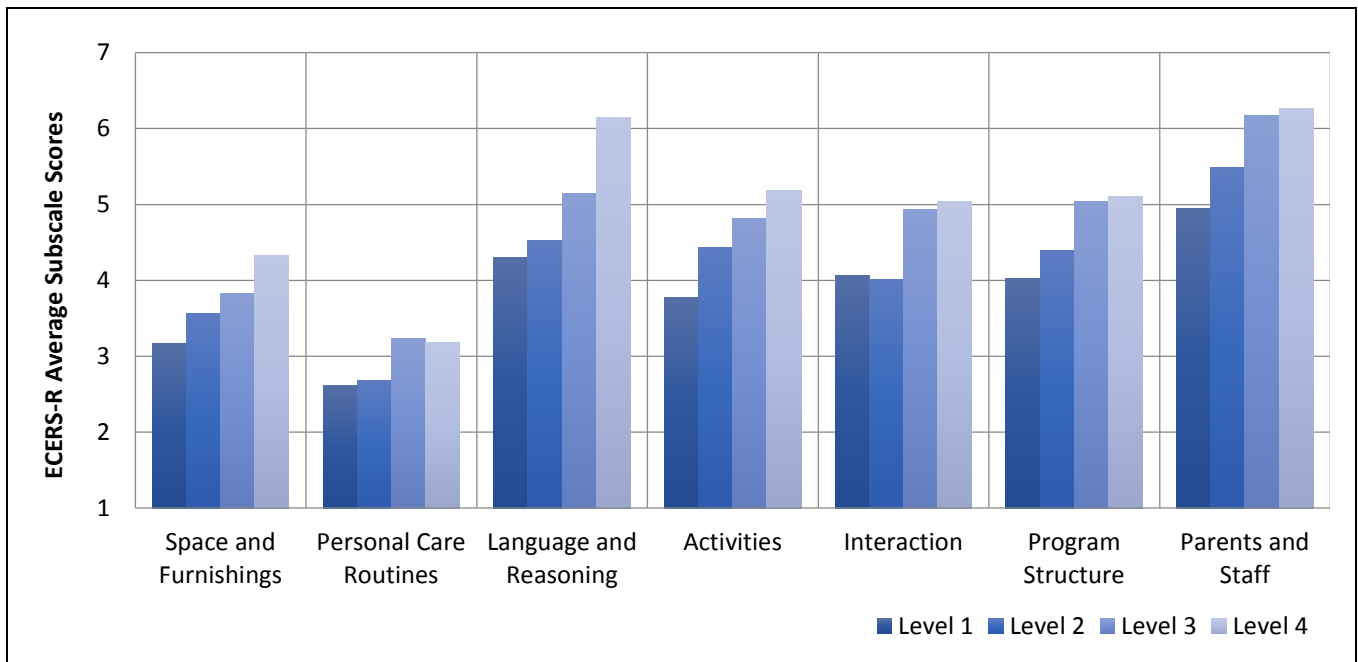
Figure 15. Overall Average item Score for ECERS-R by MA QRIS Levels



Level 4 is included in the figure as descriptive information but was excluded from statistical testing due to low sample size (N=5).

Figure 16 depicts the differences in quality by the seven subscales among the MA QRIS Levels. As can be seen from the figure, there is a fairly consistent linear pattern of increased quality across average ECERS-R scores as programs move from Level 1 to Level 4.

Figure 16. Average ECERS-R subscale scores by MA QRIS Levels



To further inform policy, the Table 115 displays the averages and standard deviations by MA QRIS Level of the individual subscales and overall average item score. Level 4 was not included in this Table as the standard deviations are not comparable to other levels due to the small sample size of programs.

Table 115. Means and standard deviations for ERS subscales for preschool by MA QRIS Level

Subscale	Mean L1	Standard Deviation L1	Mean L2	Standard Deviation L2	Mean L3	Standard Deviation L3
Space and Furnishings	3.17	.81	3.56	.85	3.82	.88
Personal Care Routines	2.61	.78	2.67	.67	3.23	.79
Language and Reasoning	4.30	1.10	4.52	1.10	5.15	.96
Activities	3.77	.70	4.43	.93	4.82	.84
Interactions	4.06	1.58	4.01	1.65	4.93	1.61
Program Structure	4.02	1.38	4.39	1.44	5.04	1.18
Parents and Staff	4.95	1.06	5.49	.77	6.17	.43
Overall Average Item Score	3.77	.78	4.12	.76	4.68	.68

Using a one-way ANOVA (see Table 116) and a Tukey-HSD Post Hoc Test, data were examined to determine if statistically significant differences in observed quality existed among the various MA QRIS Levels.

Table 116. ANOVA table of significant differences for preschool classrooms by MA QRIS Level

Subscale		df	Mean Square	F	Significance
Space and Furnishings	Between groups	3	4.23	5.80	p<.01
	Within groups	120	.73		
Personal Care Routines	Between groups	3	3.28	5.87	p<.01
	Within groups	120	.56		
Language-Reasoning	Between groups	3	8.77	8.18	p<.01
	Within groups	120	1.08		
Activities	Between groups	3	8.59	12.06	p<.01
	Within groups	120	.71		
Interactions	Between groups	3	7.94	3.11	p<.05
	Within groups	120	2.55		
Program Structure	Between groups	3	7.67	4.31	p<.01
	Within groups	120	1.78		
Parents and Staff	Between groups	3	10.73	17.71	p<.01
	Within groups	120	.61		
Overall Average Item Score	Between groups	3	6.63	12.32	p<.01
	Within groups	120	.54		

Post hoc analyses using the Tukey HSD Test identified the significant difference among preschool classrooms by MA QRIS Levels. Significant differences found among levels are presented below (see Table 117).

Table 117. Post hoc analyses of significant differences among levels for preschool classrooms

Subscale	Levels with significant differences	Significance
Space and Furnishings	Level 1 and Level 3	p<.01
Personal Care Routines	Level 1 and Level 3	p<.01
	Level 2 and Level 3	p<.01
Language-Reasoning	Level 1 and Level 3	p<.01
	Level 2 and Level 3	p<.05
Activities	Level 1 and Level 2	p<.01
	Level 1 and Level 3	p<.01
Interactions	Level 1 and Level 3	p<.01
	Level 2 and Level 3	P<.10
Program Structure	Level 1 and Level 3	p<.01
Parents and Staff	Level 1 and Level 2	p<.05
	Level 1 and Level 3	p<.01
	Level 2 and Level 3	p<.01
Overall Average Item Score	Level 1 and Level 3	p<.01
	Level 2 and Level 3	p<.01

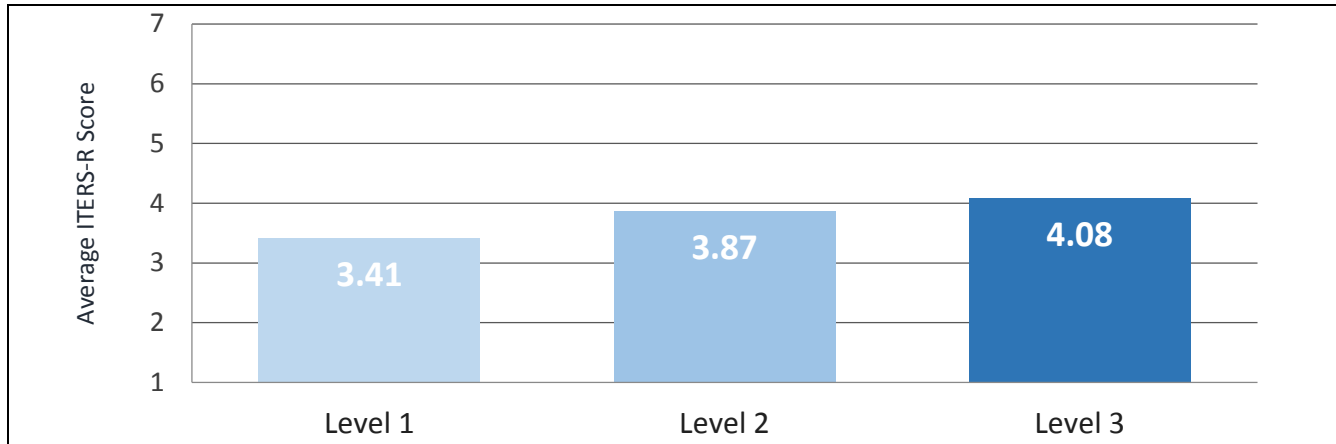
Across all subscales and the Overall Average Item Score, significant differences in quality were found between Level 1 and Level 3 ($p<.01$). These results highlight the significant differences in classrooms quality at the entrance point into the MA QRIS versus Level 3 which represents active engagement in the system as well as the point in which classroom quality has been vetted by a MA QRIS Program Quality Specialist. Additionally, the data indicate significant differences in observed quality between Levels 2 and 3 on the Personal Care Routines Subscale ($p<.01$), the Language and Reasoning Subscale ($p<.05$), the Parents and Staff Subscale ($p<.01$), and the Overall Average Item Score ($p<.01$). The difference between Level 2 and Level 3 for the Interactions Subscale was approaching significance ($p<.10$). It is important to stress that at Level 2 ERS quality is verified through program self-assessment, whereas at Level 3 classroom quality as determined by the ERS is verified through a site visit with TA (as needed) conducted by the MA QRIS Program Quality Specialists. Although causality cannot be determined, these findings of significant differences in observed quality between Level 2 and Level 3 and Level 1 and Levels 3 suggest that the Level 3 process of verification and TA by the Program Quality Specialist is effective and that the MA QRIS Level 3 is functioning well in terms of promoting quality classroom practices.

Fewer significant differences were noted between Level 1 and Level 2, in comparison to differences found among other MA QRIS Levels. Although Level 2 programs had higher observed quality scores overall and across six of the seven subscales in comparison to Level 1, significant differences between levels were only found for the Activities and Parents and Staff Subscales. As noted above, classroom quality, as measured by the ERS, at Level 2 is self-assessed. As such, there is no process for outside verification of classroom practices or compulsory TA support built into the system. Additionally, the MA QRIS does not require professional development or training on the ERS at any level. As such, it is possible that some Level 2 programs have not received guidance on the specifics of the ERS in order to self-assess sufficiently. In general, data suggest that additional supports for Level 2 programs are needed in order to promote quality classroom practices.

Infant and toddler ERS differences by MA QRIS level

A similar overall relationship among MA QRIS assigned levels and observed classroom quality was evident for infant and toddler classrooms (Figure 17). Classrooms in MA QRIS Level 1 received an ITERS-R Overall Average Item Score of 3.4 while classrooms in Level 3 programs received an ITERS-R Overall Average Item Score of 4.1.

Figure 17. Infant and toddler classrooms ITERS-R Overall Average Item Score by MA QRIS Level



As can be seen from the table, a gradual and progressive increase in quality is evident among the MA QRIS Levels for infant and toddler classrooms. To further understand differences among MA QRIS Levels, means and standard deviations by subscale and MA QRIS Levels are presented below (Table 118).

Table 118. Means and standard deviations on the ITERS-R by MA QRIS Level

Subscale	Mean L1	Standard Deviation L1	Mean L2	Standard Deviation L2	Mean L3	Standard Deviation L3
Space and Furnishings	2.85	.68	3.44	1.02	3.28	.75
Personal Care Routines	2.42	.68	2.51	.71	2.34	.90
Listening and Talking	3.86	1.08	4.46	1.05	4.80	.88
Activities	2.65	.91	3.22	1.04	3.38	1.14
Interactions	4.26	1.42	4.75	1.51	5.25	1.30
Program Structure	3.26	1.44	3.73	1.53	3.96	1.10
Parents and Staff	4.97	.87	5.43	.65	6.06	.38
Overall Item Average	3.41	.71	3.87	.77	4.08	.58

Similar to preschool classrooms, one-way ANOVAs were conducted to determine if significant differences in means scores existed among levels across the various subscales of the ITERS-R (Table 119).

Table 119. ANOVA table of significant differences for infant and toddler on the ITERS-R by MA QRIS Levels

Subscale		Df	Mean Square	F	Significance
Space and Furnishings	Between groups	2	2.56	3.51	p<.05
	Within groups	70	.73		
Personal Care Routines	Between groups	2	.15	.27	N/A
	Within groups	70	.55		
Talking and Listening	Between groups	2	4.84	4.56	p<.05
	Within groups	70	1.06		
Activities	Between groups	2	3.39	3.29	p<.05
	Within groups	70	1.03		
Interactions	Between groups	2	4.92	2.37	p=.10
	Within groups	70	2.08		
Program Structure	Between groups	2	2.83	1.40	N/A
	Within groups	70	2.02		
Parents and Staff	Between groups	2	5.79	11.82	p<.01
	Within groups	70	.49		
Overall Average Item Score	Between groups	2	2.60	5.08	P<.01
	Within groups	70	.51		

Results of the ANOVA indicated that there were no significant differences found on the Personal Care Routine and Program Structure Subscales by MA QRIS Level. As such, these subscales are not distinguishing quality by level. Review of the data above (Table 120) indicates that mean scores are low, including for Level 3 ($\bar{x}=3.96$). This suggests added PD is needed related to Program Structure for Infants and Toddlers (Item Analyses are presented below). Similarly, the means for the Health Practices Subscales are low, with Level 3 programs (mean=2.34), on average, not meeting MA QRIS standards for Level 3. Post Hoc Analyses using the Tukey HSD Test identified the significant difference among infant and toddler classrooms by individual QRIS levels. Significant differences found among levels are presented below (Table 120).

Table 120. Post hoc analyses of significant differences among infant and toddler classrooms by MA QRIS Levels

Subscale	Levels with Significant Differences	Significance
Space and Furnishings	Level 1 and Level 2	$p < .05$
Listening and Talking	Level 1 and Level 2	$p < .10$
	Level 1 and Level 3	$p < .05$
Activities	Level 1 and Level 2	$p < .10$
	Level 1 and Level 3	$p < .10$
Interactions	Level 1 and Level 3	$p < .10$
Parents and Staff	Level 1 and Level 2	$p < .05$
	Level 1 and Level 3	$p < .01$
	Level 2 and Level 3	$p < .05$
Overall Average Item Score	Level 1 and Level 2	$p < .05$
	Level 1 and Level 3	$p < .05$

Varying levels of significance were found regarding quality among the MA QRIS Levels for infant and toddler classrooms. Significant differences were found between Level 1 and Level 3 programs on the following subscales: Listening and Talking, and Parents and Staff. Differences between Level 1 and 3 were approaching significance ($p < .10$) for the Activities subscale. There were significant differences found between Levels 1 and 2 on the Space Furnishings, Listening and Talking, Activities, and Parents and Staff Subscales. Significant differences were found between Levels 2 and 3 in Parents and Staff. Interestingly, there are more significant differences between Levels 1 and 2 for infant and toddler classrooms versus the number of significant differences found between Levels 1 and Levels 2 for preschool classrooms. This is in part, driven by the low level of scores at Level 1 across multiple subscales for Infants and Toddlers as well as, the lower scores noted for infant and toddler Classrooms at Level 3, which on average, did not meet the MA QRIS Level 3 Overall Average Item Score requirements of a 4.5.

Item and indicator analyses: Strengths, challenges and licensing

Analyses of items and the indicators that dictate individual ERS item scores were conducted to explore areas of Challenge and Strength for programs.

Preschool item and indicator analyses

To inform policy and TA, researchers examined items that were points of strengths (average in the ERS Good-Excellent range, regardless of MA QRIS Level) and items that seemed to present challenges for preschool programs (average in the ERS Inadequate range, regardless of MA QRIS Level). Items that were identified as areas of strengths for preschool programs are presented below (Table 121).

Table 121. Items of strength on ECERS-R for preschool programs

Subscale	Item	Mean for sample	Mean Level 1	Mean Level 2	Mean Level 3	Mean Level 4
Personal Care Routine	Greeting and Departing	6.23	6.23	6.00	6.41	6.40
Language and Reasoning	Encouraging Children To Communicate	5.55	5.13	5.41	5.95	6.60
Activities	TV/Video and/or Computer	5.69	6.03	5.44	5.51	6.60
Program Structure	Provisions For Children With Disabilities	6.86	6.82	7.03	6.73	7.00
Parents and Staff	Provisions For Parents	5.86	5.15	5.64	6.61	7.00
	Staff Interaction And Cooperation	6.23	6.36	5.87	6.39	6.80
	Supervision And Education Of Staff	6.43	5.90	6.41	6.88	7.00
	Opportunities For Professional Growth	5.94	5.08	5.87	6.71	7.00

As indicated in the table, Strength Items were found across five of the seven subscales. In Personal Care Routines on the Greeting and Departing Item, programs averaged above a 6 at all MA QRIS Levels. This indicates that teachers are consistently talking to adults at drop off and are helping children get involved in activities upon arrival. Additionally, programs, on average, met the Good Benchmark for Encouraging Children to Communicate at all MA QRIS Levels. This indicates that appropriate communication activities take place during both free play and group times. Additionally, materials that encourage children to communicate were accessible in a variety of interest centers. In terms of the Activities subscale, Use of the TV/Video and /or Computer was an area of Strength with an average above a 5 across all MA QRIS Levels. Programs limited computer activities to those considered “good for children”. Most of the TV/Video and /or Computer materials encouraged active engagement and educators were actively involved in their use. In the Program Structure Subscale, Provisions for Children with Disabilities was an area of strength. As such, staff consistently followed through with activities and interactions recommended by other professionals. Modifications were made in the environment, so children with special needs could participate in many activities with other children. Parents were also frequently involved in the sharing of information with staff, setting goals, and giving feedback about how the program is meeting a child’s needs. Lastly, four Items from the Parents and Staff Subscale were seen as strengths. These include practices such as: parents being able to observe children’s group before enrollment; much sharing of child-related information between parents and staff; a thorough orientation for new staff including interaction with children and parents; and monthly staff meetings that include staff development activities.

In addition to Strength Items, Challenge Items in which average scores were in the ERS Inadequate range, regardless of MA QRIS Level, were also identified. Results are presented in Table 122.

Table 122. Items of challenge on ECERS-R for preschool programs

Subscale	Item	Mean for sample	Mean Level 1	Mean Level 2	Mean Level 3	Mean Level 4
Space and Furnishings	Furniture for care, play and learning	2.57	2.49	2.41	2.88	2.00
	Space for gross motor	2.12	2.08	2.00	2.22	2.60
Personal Care Routines	Meals/snacks	2.15	1.87	1.85	2.63	2.60
	Toileting/Diapering	2.06	1.82	2.00	2.39	1.80
	Health Practices	2.20	1.77	2.03	2.63	3.40
	Safety Practices	2.00	1.87	2.41	2.49	2.00
Parents and Staff	Provisions for Personal Needs of Staff*	3.79	3.44	3.79	4.22	3.60

*Falls in the Minimum range but is significantly lower than other items in Parents and Staff Subscale

As can be seen from the Table 122, most of the challenge items are from the Personal Care Routines Subscale. Additionally, programs are also struggling with Furniture for Care, Play and Learning as well as Space for Gross Motor. The Provisions for the Personal Needs of Staff item was included in this list because this particular item is an outlier among the Parents and Staff subscale, suggesting that it is an area of particular challenge for programs.

As the table indicates, there is little difference among MA QRIS Levels regarding these particular items. As such, these challenges appear to be more universal issues found across diverse types of programs. To further understand Challenge Items, researchers examined the specific indicators among these items with high fail rates to determine areas of practice that may be driving these low scores and hand coded observation notes for individual indicators to identify lapses in practice. It is important to note that the scoring scheme for the ERS requires that programs receive an Inadequate Score, if they do not meet standards for Minimal care, even though classrooms might be engaging in activities and behaviors at the higher ranges of quality. This is because certain activities are considered to be necessary for basic Minimal standards of quality to be met. As such, these indicators represent “barriers” to program receiving a higher score on specified items. Indicator analyses are presented in Table 123.

Table 123. ECERS-R challenge item indicator analysis

Subscale	Item	Indicator	Percent not meeting	Percent Meeting
Space and Furnishings	Furniture for care, play and learning	Enough furniture for routine care	84	16
	Space for gross motor	Gross motor space is generally safe	80	20
Personal Care Routines	Meals/snacks	Sanitary procedures usually maintained (75% of the time)	84	16
	Toileting/Diapering	Sanitary conditions are maintained	80	20
	Health Practices	Adequate hand washing by staff and children takes place after wiping noses, after handling animals, or then otherwise soiled (75% of the time)	84	16
	Health Practices	Staff usually take action to cut down on the spread of germs	44	55
	Safety Practices	No major safety hazards indoors or outdoors	87	13
Parents and Staff	Provisions for Personal Needs of Staff	Morning, afternoon and midday breaks provided	91	9

Space and Furnishings. The primary cause of a low score for Furniture for Routine Care, Play and Learning Item was centered on the criteria for indicator: “Sufficient Furniture for Routine Care, Play and Learning” in which 84 percent of programs were rated as not meeting. Review of the field notes indicated that the most common challenge (90 percent) with this indicator was a lack of storage for children’s personal belongings. In general, children’s personal belonging are not supposed to touch as a means of reducing the spread of lice and scabies. This ERS standard is based on recommendations from the Caring for Our Children: The National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs, Third Edition⁵¹.

Approximately, 80 percent of programs did not meet the criteria for indicator “gross motor space is generally safe.” Although the ERS playground requirements are not perfectly aligned with MA Licensing Standards, overlaps are evident. A review of the field notes indicated that many programs are not meeting safety requirements relevant to both the ERS as well as, to multiple MA EEC Licensing Requirements (see Table 124).

⁵¹ American Academy of Pediatrics, American Public Health Association, National Resource Center for Health and Safety in Child Care and Early Education. 2011. Caring for our children: National health and safety performance standards; Guidelines for early care and education programs. 3rd Edition. Elk Grove Village, IL: American Academy of Pediatrics; Washington, DC: American Public Health Association.

Table 124. Percent of programs with safety issues related to gross motor space

Safety Issue/ Requirements	Description	Percent of programs with Safety Issue
Fencing	Fence height, gate height (ERS and MA EEC Licensing have identical standards) , incomplete fencing, sharp fencing or exposed bolts	62%
Inappropriate Play Space	Parking lots, areas outside of fenced-in playgrounds (e.g, walkways, sidewalks, unfenced fields and yards), indoor gross motor space with major safety issues (e.g, loose cords, exposed outlets)	38%
Insufficient cushioning	Cushioning requirements under climbing equipment (ERS and MA EEC Licensing have the same requirements)	22%
Inappropriate Equipment	Equipment not appropriate for age of children, equipment presents tipping hazard, equipment broken, equipment too high, entrapment hazards, no equipment	33%

Personal Care Routines. Similar to Space and Furnishings, researchers reviewed field notes to highlight specific areas in which programs face challenges in terms of health practices. Many of the issues flagged by classroom observers coincided with many requirements for MA EEC Licensing. For reference the following was taken from the MA EEC 2010 Regulations for Family, Group and School Age Child Care Program: “The licensee must ensure that educators and children wash their hands with liquid soap and running water, using friction, in accordance with DPH guidelines. Hands must be dried with individual or disposable towels or automatic hand blow-dryers. The use of common towels is prohibited. Educators and children must wash their hands at least at the following times: 1. before and after water play; 2. before eating or handling food; 3. after toileting or diapering; 4. after coming into contact with bodily fluids or discharges (including sneezes, coughing); and 5. after handling caged animals or their equipment. In addition, educators must wash their hands: 1. before and after administration of medication; 2. after performing cleaning tasks, handling trash or using cleaning products.” The table below highlights overlaps between gaps observed in health practices and MA EEC Licensing Standards (see Table 125).

Table 125. Percent of preschool programs with issues related to care routines

	Care Routine Requirement	Description	Percent of programs <u>NOT</u> meeting Care Routine Requirements
Meals	Washing hands before eating/meals	Children wash hands with soap and water immediately before eating, at least 75% of the time	20%
	Washing hands after eating, when eating finger foods	Children and/or educators who eat foods with their fingers, must wash hands after eating	46%
	Educator washes hands before meal prep/serving food*	Educator must wash hands before meal prep or serving meals, even if gloves are worn	27% (most teachers only wore gloves, no hand washing observed)

	Care Routine Requirement	Description	Percent of programs NOT meeting Care Routine Requirements
Toileting/Diapering	Bathroom sinks that have a dual use must be sanitized between diverse uses	Programs must sanitize sink if used for both bathroom and other hand washing needs	75%
	Teachers must wash hands when assisting with toileting and toileting related activities*	Teachers must wash hands with soap and water after toileting or assisting with toileting, even if gloves are worn	23% (most teacher only wore gloves, no hand washing observed)
	Children wash hands after toileting	Children must wash hands with soap and water after toileting.	11%
Health Practices	Hand washing upon arrival and/or re-entry into classroom	Children and teacher must wash hands upon arrival and upon reentry from outside play	61%
	Hand washing before and after water play	Children must wash hands before and after water play with soap and water, 75% of the time	29% (based on all classroom, regardless if water play was observed)
	Hand washing after dealing with bodily fluids	Children and/or adults must wash hands after dealing with bodily fluids unrelated to meals and toileting such as nose wiping and sneezing, 75% of the time	54%
	Hand washing after touching contaminated objects (trash can lid)	Educators and children should not touch trash can lid when throwing away paper towels or other objects. Handless trash can lids required, if lids are used.	59% (all related to touching trash can lids when throwing items away)
	No gross motor hazards (either indoors or outdoors gross motor space)	No major gross motor hazards present (cushioning, fall zones, fencing, space, equipment).*	75%
	No safety hazard in classrooms	No safety hazards in classroom (glass objects that could break or shatter, spraying bleach where children can inhale fumes, unsecured cords and chemicals not properly stored)	49%

*Please note that bollards required for play areas within 30 feet of a street or parking lot are not considered in this percentage, as it is not related to MA Licensing Requirements.

As can be seen from the chart, multiple requirements for health, safety practices and gross motor space that parallel MA EEC Licensing Standards are not consistently being met. This poses particular challenges to the MA QRIS, as this appears to be one area in which programs are entering the system at a level below licensing standards (Level 1). Not only does this influence subscale scores for Personal Care Routines but it also impacts ERS Overall Average Item Scores. As such, it is a contributor to the amount of lift required by programs to meet standards to get to the next level of MA QRIS. It also highlights the need for additional licensing supports related to space, health and safety, in general.

Parents and Staff. Among preschool classrooms, programs scored significantly lower on the Provisions for Personal Needs of Staff Item in the Parents and Staff Subscale. Indicator analyses found that about 91 percent of

programs did not meet the requirements for “Program provides Morning, Midday or Afternoon Breaks daily.” Although almost all teachers reported having a break during the day, the majority reported having only a half hour lunch break, which is the minimal requirement for break time as indicated by the MA Department of Labor for full-time employees. It is important to emphasize how critical break time can be to early educators, since they often cannot leave the classroom for personal needs, even momentarily, due to the need to consistently maintain ratios and because of the demanding nature of the work. This can make balancing personal needs (e.g. bathroom breaks, snacks, the need to make a phone call or appointment difficult). Interestingly, despite the importance of break time, break time is not specified as a MA QRIS criteria regarding professional supports or benefit for educators.

Infant and toddler item and indicator analyses

Paralleling the preschool classrooms, similar analyses of items and indicators for infant and toddler classrooms were conducted. Table 126 presents the items in which programs averaged at least a 5 (ERS Good range), regardless of MA QRIS level.

Table 126. ITERS-R items of strengths for infant and toddler classrooms

Subscale	Item	Mean for sample	Mean Level 1	Mean Level 2	Mean Level 3
Parents and Staff	Staff Interaction And Cooperation	5.89	5.64	5.90	6.20
	Staff Continuity	5.78	5.85	5.71	5.73
	Supervision And Education Of Staff	6.30	5.89	6.35	6.87

As can be seen from the chart, items of Strength all came from the Parents and Staff Subscale. Importantly scores for Staff Interaction and Cooperation as well as Staff Continuity were high across all MA QRIS levels. The various criteria that comprise these items are particularly important for classrooms with very young children.^{52,53} In order to provide high quality infant and toddler care, educators need to be consistent and work closely as a team in order to meet the high care demands and unique social emotion needs infants and toddlers.

Researchers also examined Challenge Items in which, the averages of classroom across all MA QRIS levels were rated in the ERS Inadequate range for care. These items are presented below (Table 127).

⁵² Hamre and Pianta 2005;

⁵³ Mashburn et al. 2008

Table 127. ITERS-R items of challenge for infant and toddler classrooms

Subscale	Item	Mean for sample	Mean Level 1	Mean Level 2	Mean Level 3
Space and Furnishings	Furniture for care, play and learning	2.70	2.78	1.4	2.67
Personal Care Routines	Meals/snacks	1.97	1.67	1.15	2.47
	Nap	1.90	1.96	2.04	2.08
	Diapering/Toileting	1.32	1.19	1.35	1.47
	Health Practices	2.36	2.22	2.77	1.80
	Safety Practices	1.97	1.37	2.23	2.40
Listening and talking	Using Books	2.50	1.89	2.90	2.67
Activities	Blocks	2.55	2.15	2.65	2.98
	Active Physical Play	2.18	2.26	1.97	2.33
Parents and Staff	Provisions for Personal Needs of Staff*	3.68	3.26	3.81	4.13

*Falls in the Minimum range but is significantly lower than other items in Parents and Staff Subscale

As can be seen from the data, infant and toddler classrooms, regardless of MA QRIS level, need added supports in several key areas. As such, the data suggest universal field issues may exist beyond the MA QRIS and subsequently a need for added professional development and field supports is required. To further identify areas of needed support and attention, indicator analyses of these items was undertaken. The table (Table 128) below presents the indicator fail rates driving the low scores. These indicators represent lapses in quality and in most cases barriers to higher scores, as many programs were observed exhibiting behaviors representing higher quality levels while simultaneously not meeting the standards for a Minimal level of quality which is required to achieve a higher rating on individual items.

Table 128. ITERS-R challenge indicator analyses

Subscale	Item	Indicator	Percent not meeting	Percent Meeting
Space and furnishings	Furniture for care, play and learning	Enough furniture for routine care	72	28
Personal Care Routines	Meals/snacks	Sanitary Procedures not met at least half the time	73	27
	Meals/snacks	Inadequate Supervision	47	53
	Nap	Healthful provisions	82	15
	Diapering/Toileting	Sanitary conditions are maintained at least half the time	93	7

	Item	Indicator	Percent not meeting	Percent Meeting
Personal Care Routines	Health Practices	Staff cut down on the spread of germs	46	54
	Health Practices	Hands of children and staff washed 75% of the time when needed	76	24
	Safety Practices	No more than 3 safety hazards indoors or outdoors	77	23
Listening and Talking	Using Books	At least 6 appropriate books accessible much of the day	46	54
Activities	Active Physical Play	Some space for outdoor, physical play used at least 3 times a week, year around	58	42
	Blocks (only scored in classrooms with a child over 12 months)	At least one set of blocks accessible daily	26	74
	Blocks (only scored in classrooms with a child over 12 months)	Blocks accessible much of the day	55	45
Parents and Staff	Provisions for Personal Needs of Staff	Morning, afternoon and midday breaks provided	96	4

Space and Furnishings. Similar to preschool classrooms the primary cause of a low score for the Furniture for Routine Care, Play and Learning Item was centered on the criteria for indicator: “Sufficient Furniture for Routine Care, Play and Learning” in which 72 percent of programs were rated as not meeting. Review of the field notes indicated that the most common challenge (88 percent) with this indicator was a lack of storage for children’s personal belongings. Mirroring the requirements for preschool classroom, the personal belongings of infants and toddlers are not supposed to touch, as a means of reducing the spread of lice and scabies.

Personal Care Routines. Many Personal Care Routine items did not meet the Minimum standard for care on the ITERS-R. Field notes from observations were reviewed to determine which practices programs failed to meet requirements. Similar to the preschool classrooms, many of these practices are consistent with MA EEC Licensing Standards and suggest field-wide challenges, outside of the MA QRIS and levels of quality. For review, EEC licensing standards are presented including the stipulation for mouthed toys, which is common in infant and toddler classrooms. MA EEC 2010 Regulations for Family, Group and School Age Child Care Program are: “The licensee must ensure that educators and children wash their hands with liquid soap and running water, using friction, in accordance with DPH guidelines. Hands must be dried with individual or disposable towels or automatic hand blow-dryers. The use of common towels is prohibited. Educators and children must wash their hands at least at the following times: 1. before and after water play; 2. before eating or handling food; 3. after toileting or diapering; 4. after coming into contact with bodily fluids or discharges (including sneezes, coughing); and 5. after handling caged animals or their equipment. In addition, educators must wash their hands: 1. before and after administration of medication; 2. after performing cleaning tasks, handling trash or using cleaning products. Toys mouthed by children must be set aside and stored after each use and may not be used by another

child until they are washed and disinfected.” Table 129 presents the indicators that overlap with MA Licensing and the ERS requirements and the percent of programs **not** meeting each standard.

Table 129. Percent of programs not meeting specific routine care requirements

	Routine Care Requirement	Description	Percent of programs NOT meeting Care Routine Requirements
Meals	Inadequate hand washing before eating/meal/bottle prep	Children and/or educators did not wash hands with soap and water immediately before eating, at least 50% of the time	62%
	Washing hands after eating, when eating finger foods	Children and/or educators who eating foods with their finger, must wash hands after eating	89%
	Inadequate Supervision	Educator <u>must</u> sit with children while they are actively eating	73%
Nap	Healthful Provisions	Cribs/mat must be space 36 inches apart	47%
	Safe Sleep/Inappropriate Sleep practices	Children under 12 months of age should not use blankets; children should not be swaddled; children should be easily to supervised (separate sleep room); inappropriate sleep space (e.g., sleeping in swings); children left in cribs awake for extended periods	20%
Toileting/Diapering	Sanitary Diapering Practices	Must follow ERS procedures as recommended by Caring for Our Children: The National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs, Third Edition (exceeds MA Licensing Standards)	93%
	Teachers and children must wash hands after diapering	Teachers and children must wash hands with soap and water after toileting or assisting with toileting, even if gloves are worn by teacher	35%
Personal Care Routines	Hand washing upon arrival and/or re-entry into classroom	Children and teachers must wash hands upon arrival and upon reentry from outside play	43%
	Hand washing before and after water play	Children must wash hands before and after water play with soap and water, 75% of the time	22% (based on all classrooms, regardless if water play was observed)

	Routine Care Requirement	Description	Percent of programs <u>NOT</u> meeting Care Routine Requirements
Personal Care Routines	Hand washing after touching contaminated objects (trash can lid)	Educators and children should not touch trash can lid when throwing away paper towels or other objects. Handless trash can lids required, if lids are used.	48% (all related to touching trash can lids when throwing items away)
	Mouth toys removed	Provider removes toys that have been mouthed from play area for sanitation purposes.	22%
Safety	Less than 3 safety hazard either indoors or out doors	No more than 3 major safety hazards present either indoors or outdoors	<p>77% (at least 3)</p> <p>Indoors Hazards present: Spraying bleach near children 26%; picking up children by arms instead of torso 49%; choke hazards present 53%; changing table lip less than 6 inches 70%.</p> <p>Outdoor Hazards present: fences and gates too low (same as MA Licensing requirements) 69%; Inappropriate Gross Motor Equipment or Space 25% (poor repair, not age appropriate or entrapment risk)</p>

The above data indicate that programs are not consistently meeting meals, health practices, gross motor and safety requirements, many of which have consistencies with MA EEC licensing standards. As such, the challenges faced by the MA QRIS in these areas of quality are particularly strong, given that the Minimal standards for licensing are not consistently being met by programs.

Listening and Talking & Activities. Scores for Using Books represents a lack of access to these materials, as opposed to a lack of books in the classroom. The score for blocks represents a lack of blocks in (26 percent of classrooms with children over 12 months of age) as well as a lack of access for “much of the day”. Books and block materials must be accessible for “much of the day” to receive a Minimal score. This means when children are ready to play, they should be able to access these materials, as opposed to being forced to remain in group time, required to engage in group activities and/or being told that the materials are not available for use.

In terms of Active Physical Play, 58 percent of the programs did not go outside at least 3 times per week. Similarly, approximately, 66 percent of programs did not meet standards for criterion “free play occurs daily both indoors and outdoors”. For the purpose the study, programs were held to the standards for outside play as defined by the MA EEC weather and temperature recommendations.

Parents and Staff. Similar to preschool classrooms, infant and toddler classrooms scored significantly lower on the Provisions for the Personal Needs of Staff item in comparison to other items in the Parents and Staff Subscale. Review of the criteria indicated that 96 percent of programs were not consistently offering morning, midday and afternoon breaks. Similar to preschool classrooms, break time is necessary to offer teachers some opportunities to attend to personal needs during the day. Because of the high needs of children, particularly infants and toddlers, the opportunity to attend to personal matters such as to make a quick phone call while in the classroom is extremely limited.

The relationship between preschool and infant and toddler classrooms ERS scores

Researchers examined differences between preschool classroom and infant and toddler classrooms in terms of ERS scores. At first look, infant and toddler classrooms and preschool classrooms appear to be different in terms of quality. When treated as independent samples, multiple significant differences are noted.

T-tests were conducted to examine differences in ERS scores between preschool and infant and toddler classrooms. When treating infant and toddler classrooms and preschool classrooms as independent samples, significant differences are noted between the two samples in terms of Space and Furnishings ($t=-2.64$, $p<.01$), Personal Care Routines ($t=-3.66$, $p<.01$), Language and Reasoning ($t=-2.50$, $p<.05$), Activities ($t=-8.87$, $p<.01$), Program Structure ($t=-4.27$, $p<.01$), and Overall Average Item Score ($t=-4.06$, $p<.01$) with preschool classrooms consistently scoring significantly higher.

It is important to note however, that the infant and toddler classrooms came from the **same** programs as the preschool classrooms (about 60 percent of programs in the sample had infant and toddler classrooms which is consistent with the population of programs in MA QRIS). As such, the infant and toddler classroom are not independent of the preschool classrooms but share the same administrative structure, management and potentially the same approach towards providing early education and care. As a result, a paired sample T-tests were conducted with ERS scores for programs that had both infant and toddler and preschool classrooms. Among the programs that had both infant and toddler classrooms and preschool classrooms, we found only one significant difference among the ERS subscales. Preschool classrooms scored significantly higher on Activities in comparison to infant and toddler classrooms at the same program ($t=7.45$, $p<.05$). It should be noted that the Activities requirements for infant and toddler classrooms are different than those of preschool classrooms, with infant and toddler classrooms having more stringent requirements for accessibility. This suggests that although infant and toddler classrooms may have lower levels of quality than preschool classrooms in general, this relationship does not exist within programs that serve both infants and toddlers and preschool-aged children, with quality scores being more consistent across the two different types of care.

The above analyses were also run by level. Since, there were so few programs with infant and toddler classrooms at Level 3, comparisons could only be made between programs with infant and toddler classrooms and preschool classrooms at Level 1 and Level 2. No significant differences were found among programs with both types of classrooms by MA QRIS Level. This suggests that at Level 1 quality for both infant and toddler classrooms and preschool classrooms is relatively similar. The same conclusion can also be drawn for Level 2 programs.

Classroom quality and the Arnett CIS

The Arnett Caregiver Interaction Scale (CIS) is a commonly used tool for measuring the quality of caregivers' interactions with children. The CIS comprises 26 items organized into four distinct subscales: Sensitivity, Harshness, Permissiveness, and Detachment. Although three of the four Arnett CIS subscales use terms associated with negative interactions (Harshness, Detachment, and Permissiveness), all scales were recoded so that scores of 1 represent teacher-child interactions that rarely or never meet standards and scores of 4 represent interactions that consistently meet standards. As such, scores close to 4 would reflect a low frequency of harshness, detachment, or permissiveness, and a high level of sensitivity or overall interaction.

Preschool classroom quality and Arnett

Across the sample of preschool classrooms, overall Arnett CIS scores ranged from 1.98 to 3.90, with a mean of 3.17 out of 4.00 (see Table 130). However, the distribution of scores was skewed toward the positive end of the scale (skewness of -0.61), and most classrooms (67 percent) received a score of 3 or above. Previous studies have found similar patterns in overall Arnett scores, and it has been suggested that, given this general skew, the scale may not be “well suited to distinguish between caregivers who are ‘highly’ versus ‘moderately’ positive in their interactions with children.”⁵⁴ Scores on the four subscales also varied, with study programs receiving the highest mean scores on the harshness subscale (3.41, or low harshness) and the lowest on the sensitivity subscale (2.95). Subscale results also displayed moderate to high degrees of skew toward the positive end of the scale, with skewness varying from -.41 (for the sensitivity scale) to -1.11 (for the harshness scale).

Table 130. Arnett CIS Scores for preschool classrooms, descriptive statistics

Subscale	Minimum	Maximum	Mean	Std. Deviation
Sensitivity	1.50	3.95	2.95	.61
Harshness	1.81	4.00	3.41	.51
Detachment	2.00	4.00	3.35	.47
Permissiveness	1.38	4.00	3.04	.52
Overall Score	1.98	3.90	3.17	.47

Infant and toddler classroom quality and Arnett

Within infant and toddler classrooms, programs received a mean overall Arnett score of 3.40 out of 4.00, with scores ranging from 2.33 to 3.92 (Table 131). Again, scores tended to be skewed towards the high end of the scale (skewness of -.97). In fact, 86 percent of infant and toddler classrooms received a CIS score of 3 or above. With regard to the four subscales, average scores tended to be the highest on the harshness scale (3.81 or low harshness) and were lowest on the permissiveness scale (3.02). Also of note, CIS scores and subscale scores were higher and less varied (lower standard deviation) in infant and toddler classrooms than in preschool classrooms, both overall and for each of the four subscales.

Table 131. Arnett CIS scores for infant and toddler classrooms, descriptive statistics

Subscale	Minimum	Maximum	Mean	Std. Deviation
Sensitivity	2.00	4.00	3.11	.50
Harshness	2.44	4.00	3.81	.30
Detachment	2.75	4.00	3.69	.34
Permissiveness	2.13	3.50	3.02	.23
Overall Score	2.33	3.92	3.40	.33

Relationship between Arnett and ERS

Notably, classroom Arnett CIS scores were positively associated with environment rating scale scores. For preschool classrooms, overall classroom Arnett CIS scores were moderately correlated with Total ECERS-R

⁵⁴ Colwell, N., Gordon, R. A., Fujimoto, K., Kaestner, R., & Korenman, S. (2013). New evidence on the validity of the Arnett Caregiver Interaction Scale: Results from the Early Childhood Longitudinal Study-Birth Cohort. *Early Childhood Research Quarterly*, 28(2), 218–233. <http://doi.org/10.1016/j.ecresq.2012.12.004>.

scores and strongly correlated with ECERS-R Interaction subscale scores.⁵⁵ For infant and toddler classrooms, overall classroom Arnett CIS scores were moderately correlated with both Total ITERS-R scores and ITERS-R Interaction subscale scores.⁵⁶ This means that programs with higher ECERS-R and ITERS-R Total scores and Interaction subscale scores also tended to have higher Arnett CIS scores, while programs with lower Total scores and Interaction subscale scores tended to have lower Arnett CIS scores.

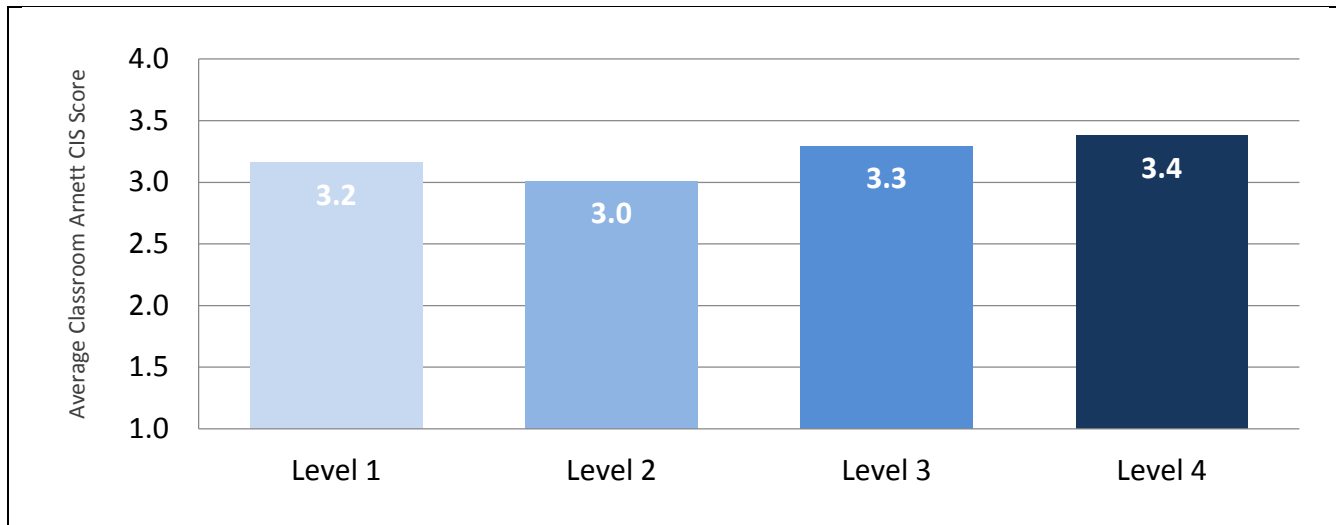
Differences in Arnett quality by MA QRIS levels

Overall, Arnett CIS scores were found to be associated with MA QRIS level for preschool classrooms. No significant association was found between these scores and MA QRIS level for infant and toddler classrooms either overall or for individual subscales.

Preschool classrooms Arnett scores and MA QRIS levels

For preschool classrooms, differences in scores were found to be significant between Levels 2 and 3, indicating that these two levels are differentiating quality with respect to teacher-child interactions. As Figure 18 shows, from Levels 2 through 4, there was a progressive increase in classroom Arnett CIS scores across levels, from 3.0 (Level 2) to 3.4 (Level 4) with higher level programs receiving higher Arnett CIS scores, on average. The average score for Level 1 programs was 3.2, between that of Level 2 and Level 3 programs.

Figure 18. Overall Arnett CIS score by MA QRIS level for preschool classrooms

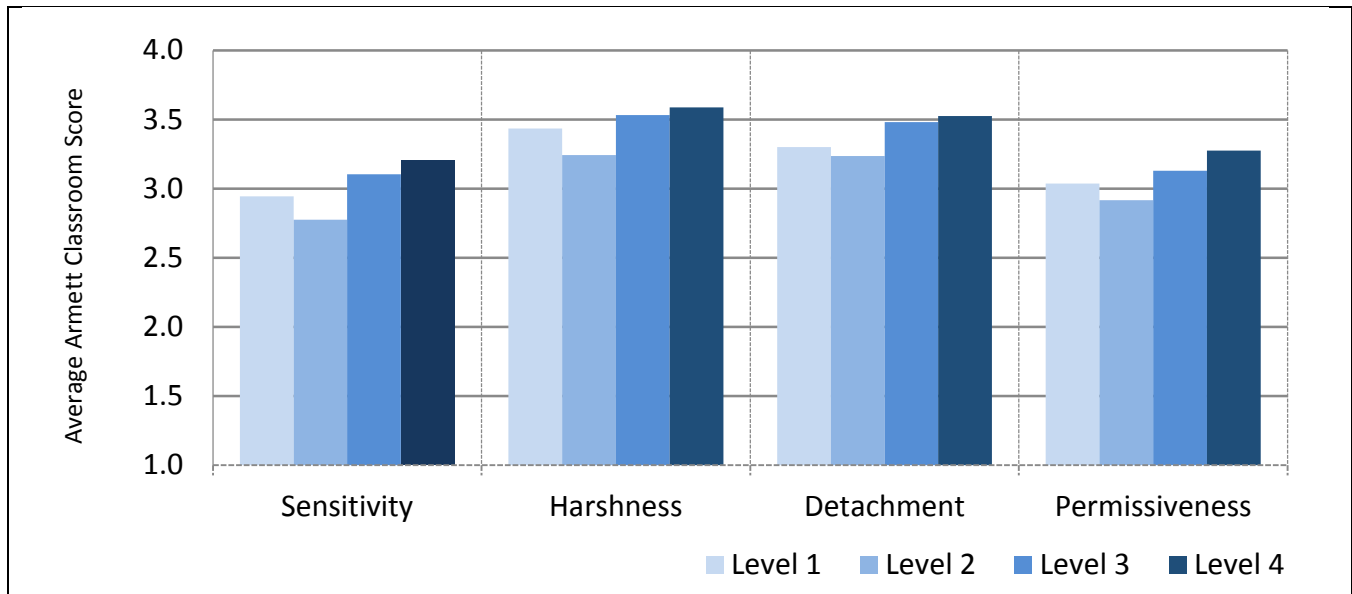


A similar pattern is evident for each of the four Arnett subscales. For all four Arnett CIS subscales, scores increased consistently between Levels 2, 3, and 4, and subscale scores were slightly higher for classrooms at Level 1 compared with Level 2 (see Figure 19).

⁵⁵ For preschool analyses, Pearson correlation coefficients between Arnett CIS scores and Total ECERS-R scores and Arnett CIS scores and ECERS-R Interaction subscale scores were 0.687 (p-value of 0.000) and 0.796 (p-value of 0.000), respectively.

⁵⁶ For infant and toddler analyses, Pearson correlation coefficients between Arnett CIS scores and Total ITERS-R scores and Arnett CIS scores and ITERS-R Interaction subscale scores were 0.445 (p-value of 0.000) and 0.662 (p-value of 0.000), respectively.

Figure 19. Average Arnett CIS subscale scores by MA QRIS Level for preschool classrooms



As further detail, Table 132 shows the mean and standard deviation for overall Arnett CIS scores and for each of the subscales for preschool classrooms by MA QRIS Level. Level 4 was not included in this table, as given the low sample size (n=5), the standard deviation for this level is not comparable to those of other levels.

Table 132. Mean and standard deviation for Arnett CIS by MA QRIS Level for preschool classrooms

	Level 1		Level 2		Level 3	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Sensitivity	2.94	.64	2.78	.60	3.10	.60
Harshness	3.44	.40	3.24	.58	3.53	.51
Detachment	3.30	.55	3.24	.40	3.48	.44
Permissiveness	3.04	.45	2.92	.58	3.13	.54
Overall Score	3.16	.46	3.01	.48	3.30	.47

To determine whether differences in Arnett CIS scores and subscale scores observed across levels are statistically meaningful, a series of one-way ANOVAs were conducted, the results of which are shown in Table 133.

Table 133. ANOVA table of significance differences by Arnett CIS Subscale for preschool classrooms

		df	Mean Square	F	Significance
Sensitivity	Between Groups	3	.819	2.244	p<.10
	Within Groups	120	.365		
Harshness	Between Groups	3	.628	2.534	p<.10
	Within Groups	120	.248		
Detachment	Between Groups	3	.490	2.278	p<.10
	Within Groups	120	.215		
Permissiveness	Between Groups	3	.399	1.484	n/a
	Within Groups	120	.269		
Overall Score	Between Groups	3	.619	2.912	p<.05
	Within Groups	120	.212		

As this table shows, there were significant relationships found between MA QRIS Level and the Arnett CIS overall score (at the 0.05 significance level) and three of the four subscales (Sensitivity, Harshness, and Detachment; at the 0.10 significance level). Post hoc tests were undertaken using Tukey's HSD procedure to identify the specific levels between which these differences occurred (Table 134). In each instance where the overall ANOVA showed a significant relationship between MA QRIS Level and Arnett CIS score, post hoc testing showed the differences between Levels 2 and 3 to be significant. The differences between Level 1 and 2 were found not to be statistically significant in any instance.

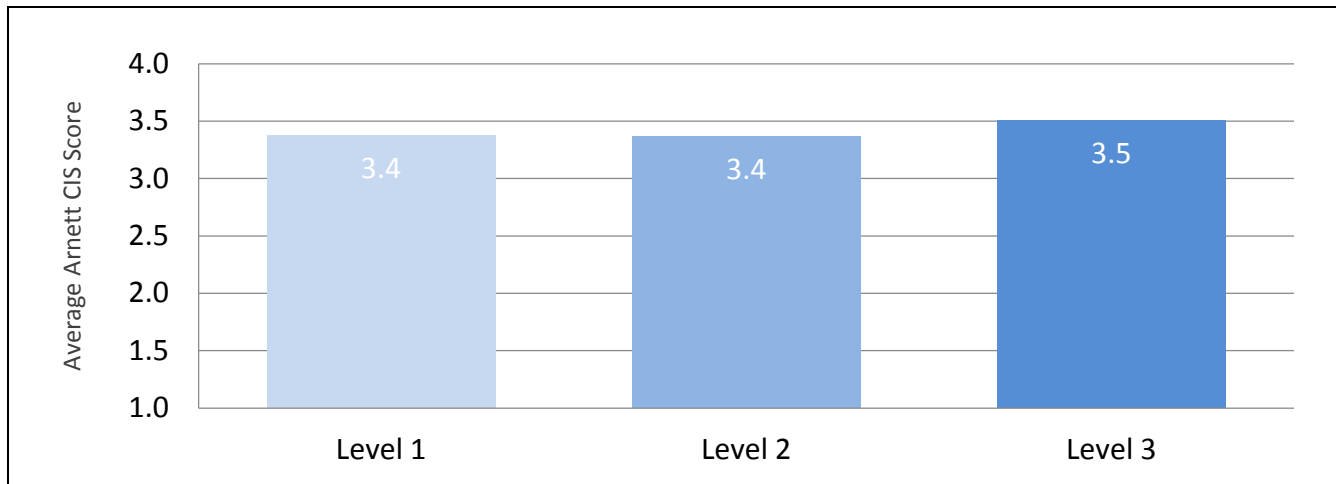
Table 134. Post hoc analysis of significant differences on the Arnett CIS among MA QRIS Levels for preschool classrooms

Subscale	Level Difference	Significance
Sensitivity	Level 2 and Level 3	p<.10
Harshness	Level 2 and Level 3	p<.10
Detachment	Level 2 and Level 3	p<.10
Overall score	Level 2 and Level 3	p<.05

Infant and Toddler Classroom Arnett CIS Scores and MA QRIS Levels

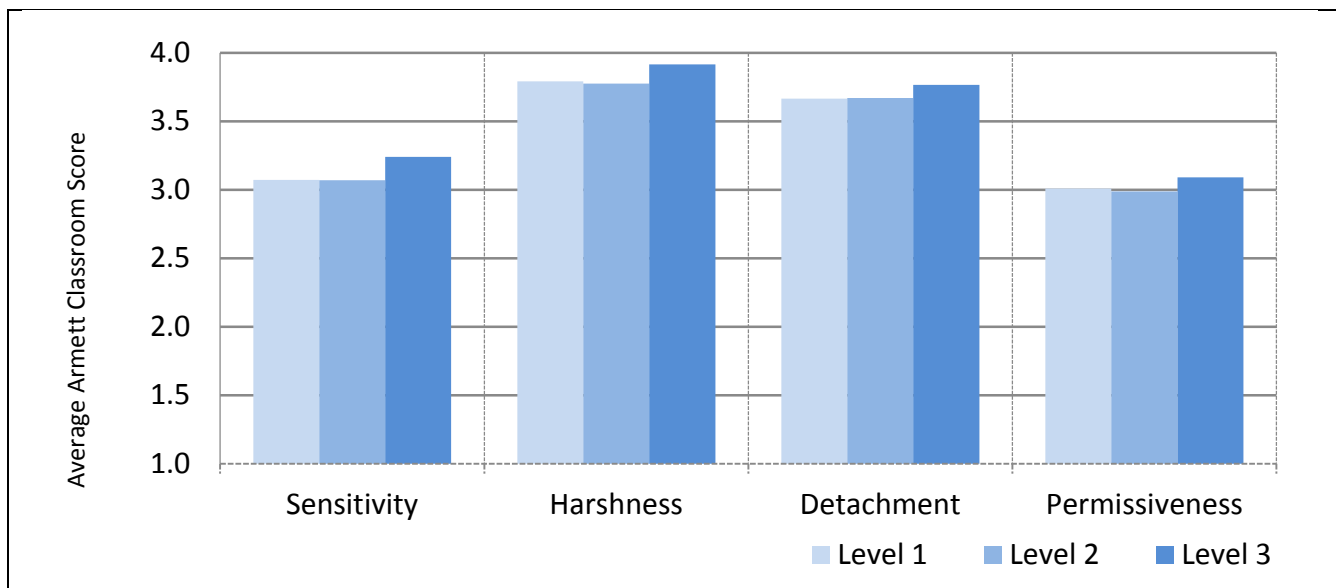
In infant and toddler classrooms, there was little variation in overall average Arnett CIS scores across MA QRIS Levels. For Level 1 and 2 programs, the average Arnett CIS scores for infant and toddler classrooms was identical, or 3.4 out of 4. The average Arnett CIS score was slightly higher, or 3.5, at Level 3 (Figure 20). As mentioned previously, there was somewhat limited variation in CIS scores overall, with most infant and toddler programs receiving scores at the high end of the scale, or above 3.

Figure 20. Average overall Arnett CIS score by MA QRIS level for infant and toddler classrooms



A similar pattern was seen for Arnett CIS subscale scores in infant and toddler classrooms (Figure 21). For each of the four subscales, average scores remained relatively consistent across Levels 1 and 2. Level 3 programs tended to have slightly higher subscale scores, although the increase was generally small.

Figure 21. Average Arnett CIS subscale scores by MA QRIS level for infant and toddler classrooms



As further detail, Table 135 presents the average and standard deviation for overall Arnett CIS scores and the four subscales for infant and toddler classrooms. As with preschool classrooms, Level 4 was not included in this table due to the low sample size (n=1).

Table 135. Mean and standard deviation for Arnett CIS by MA QRIS level for infant/toddler classrooms

	Level 1		Level 2		Level 3	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Sensitivity	3.07	.48	3.07	.54	3.24	.49
Harshness	3.79	.30	3.78	.35	3.92	.15
Detachment	3.67	.35	3.67	.35	3.77	.32
Permissiveness	3.01	.23	2.99	.22	3.09	.23
Overall Score	3.38	.31	3.37	.36	3.51	.29

Not surprisingly in light of the limited degree of difference observed across levels in the descriptive analyses, one-way ANOVA did not reveal any significant relationships between MA QRIS Levels and Arnett CIS scores for infant and toddler classrooms.

Summary and implications

The Massachusetts QRIS Levels appears to distinguish observed classroom quality for center-based providers, with significant differences found in quality among overall levels of quality and across multiple subscales of quality as measured by the ERS and the Arnett CIS. Additionally, the data indicated a progressive system in which incremental gains in observed quality were noted at each MA QRIS Level.

The greatest differences in observed quality as measured by the ERS among preschool classrooms and MA QRIS Levels were noted between Levels 1 and Levels 3. Interestingly, the fewest number of significant differences were noted between Levels 1 and Levels 2. This suggests that further supports for Level 2 programs are needed to promote more efficient and accurate self-assessment. One strategy would be to include ERS training as a requirement for Level 2 and/or required engagement with technical assistance providers to support the self-evaluation process and provide guidance regarding the ERS standards for care and developmentally appropriate practice.

Similar to preschool classrooms, significant differences in observed classroom quality were evident among MA QRIS Levels for infant and toddler classrooms. The number of significant differences found among the diverse subscales of the ERS by MA QRIS Levels was smaller for infant and toddler classrooms in comparison to preschool classrooms. The greatest number of differences was noted between Levels 1 and Levels 2 which may be driven, in part, by the low scores among programs in Level 1. The lack of differences between Levels 2 and Level 3 may, in part, be driven by the low scores among programs at Level 3, which on average scored below MA QRIS Level 3 standards for the ERS.

The MA QRIS Level Verification procedures appear to be functioning, although there is a need to address the process for re-certification and the continued standardization and refinement of verification processes, as there appears to be evidence of drift in program quality with time and some inconsistencies by region. The data suggested that larger the time gap between the quality verification of classroom practices at Level 3 and the study observation, the less likely programs were to be consistently meeting MA QRIS standards for Level 3. Additionally, one region in MA had a disproportional number of Level 3 programs that did not meet Level 3 standards for the study. This is in part, due to the changes in the verification processes that have occurred, since the MA QRIS began. Regardless, this underscores the importance of establishing clear and reliable verification processes and procedures for both MA QRIS Program Quality Specialist as well as providers.

Data indicated that some programs may be “choosing” not to engage in the MA QRIS system, even though they may be meeting ERS standards for the next MA QRIS Level. MA EEC may need to adapt strategies to promote

engagement either through grants or incentives to facilitate movement and the continual quality improvement among programs.

About 18 percent of Level 1 preschool programs and 32 percent of infant and toddler programs were rated as providing Inadequate care, as determined by ERS Overall Average Item Score. This has implications for the MA QRIS in that it suggests that a portion of Level 1 programs are functioning with limited quality and subsequently may need significant supports to progress within the MA QRIS system, including professional development on developmentally appropriate practices, professional development on the MA Early Learning Standards, and TA/coaching either through MA QRIS supports or potentially through mentor programs. Data suggested that some of the health and safety issues raised by the ERS observations appear to be issues that extend beyond the MA QRIS and include MA Licensing Standards. As such, to improve health and safety practices and support the MA QRIS, MA Licensing Standards and their reinforcement need to be included in the approach to refine the MA QRIS. As part of this, MA EEC may need to provide additional licensing supports such as: training and support for the field on MA licensing requirements to ensure educators have a solid of understanding of requirements and regulations; training and reliability standards for licensors to ensure MA quality standards for care are being met; manageable licensing caseloads; and revision of licensing support materials.

ERS strength items for preschool classrooms include: Greeting and Departing; Encouraging Children to Communicate; Use of TV/Video and/or Computer; Provisions for Children with Disabilities; Provisions for Parents; Staff Interaction and Cooperation; Supervision and Education of Staff; Opportunities for Professional Growth. These strengths did not vary by MA QRIS Level and represent more universal strengths of programs beyond the MA QRIS.

ERS Challenge Items for preschool classrooms include: Furniture for Care, Play and Learning; Space for Gross Motor; Meals/Snacks; Toileting/Diapering; Health Practices; Safety Practices; Provisions for Personal Needs of Staff. These items represent universal barriers to quality, regardless of MA QRIS Level.

ERS Strength Items for infant and toddler classrooms include: Staff Interaction and Cooperation; Staff Continuity; Supervision and Evaluation of Staff. Similar to preschool classrooms, these appear to be more universal strengths outside of the MA QRIS.

ERS Challenge Items for infant and toddler classrooms include: Furniture for Care, Play and Learning; Meals/Snacks; Nap; Diapering/Toileting; Health Practices; Safety Practices; Using Books; Active Physical Play; Blocks (only scored in classrooms with a child over 12 months); and Provisions for Personal Needs of Staff. Indicator analyses suggested that TA is needed regarding reducing group time and structured time, encouraging outside play, and training educators on health practices and the identification of safety hazards. Again, these appear to be challenges to the field in general, as opposed to MA QRIS-specific issues.

Data suggest that perhaps MA EEC should consider adding specific MA QRIS criteria related to break time as a means of supporting teachers, particularly since so few educators appear to be getting more than one break/day and many have a limited amount of time (30 minutes only).

Regardless of QRIS level, most preschool classrooms in the sample received ERS scores in the Minimal to Good range of that scale. It should be noted that the mean average score on the ECERS-R is higher than the average score reported in other states.^{57, 58} Scores in infant and toddler classrooms tended to be lower on average, typically falling in the Minimal range. Although scores for preschool classrooms and infant and toddler classrooms appear to be different, when examined at the *program level*, no significant differences were noted between infant and toddler classrooms and preschool classroom quality within programs for Levels 1 and Levels 2. This suggests that the MA QRIS continues to take a *program* approach to TA and quality improvement.

⁵⁷ Lahti, Elickerb, Zellmanc, & Fiene, 2014;

⁵⁸ Tout, Starr, Isner, Cleveland, Albertson-Junkans, Soli & Quinn, 2011.

Section IV: Classroom Quality and Key Program Attributes

In addition to the above analyses, researchers examined associations between observed classroom quality and other key program attributes, as a means of further informing the system and areas for potential focus. These key program variables included: educator turnover, NAEYC accreditation, Head Start affiliation, and curriculum.

Educator turnover

Analyses were done to examine associations between MA QRIS level, observed classroom quality, and educator turnover rates. Administrators were asked to report the number of educators that had left the program in the past 12 months, including both full-time and part-time staff. The percentage of turnover was then calculated by dividing the number of educators who left the program by the total number of educators employed. Table 136 presents the turnover rates for each MA QRIS Level.

Table 136. Percent of educator turnover by MA QRIS Level

	Percent Turn Over	Standard Deviation	Variance
Level 1	25	.24	.06
Level 2	21	.16	.03
Level 3	18	.15	.02
Level 4	17	.21	.05

As the table indicates, turnover for study programs ranged from 17 percent for Level 4 to 24 percent for Level 1. Despite the downward trend by level, an ANOVA indicated no significant differences in turnover among the MA QRIS Levels. As evidenced by the data, there was also limited variability in turnover among programs by MA QRIS Level. Additionally, no significant correlations were found between turnover and observed classroom quality as measured by the ERS.

Curriculum

In validation study interviews, administrators were asked to identify the curriculum the program used, and these data were analyzed on the basis of whether programs were using a commercially vetted, published curriculum or whether they were relying on a self-developed curriculum.

Preschool

The number of programs using a commercially vetted curriculum versus a self-made curriculum was not distributed among MA QRIS Levels, with most of the program using self-made curricula at Levels 1 and 2. Therefore, analyses regarding curriculum by MA QRIS Level was not possible. Analyses did, however, indicate a significant difference in Overall Average Item Quality Score between preschool programs with vetted curricula in comparison to programs with self-developed curricula ($t=-2.95$, $p<.01$). Results also indicated significant differences on ERS Parents and Staff ($t=-3.48$, $p<.01$), Program Structure ($t=-2.00$, $p<.05$), and Activities ($t=-4.93$, $p<.01$) Subscales between programs with vetted curricula and those using their own developed curricula. Interestingly, no significant differences were noted among the Interactions, Space and Furnishings, and Language and Reasoning Subscales, which may be aspects of preschool classroom quality that are less likely to be driven by curriculum.

Infant and toddler

For infant and toddler programs, significant differences were also found among classrooms using vetted curricula in comparison to classrooms utilizing a self-developed curriculum ($t=-3.20$, $p<.01$). In terms of subscales, significant differences were found between infant and toddler programs with vetted curricula and those with self-developed curricula among **all** subscales of the ERS (t ranged -4.03 – 2.11 , $p<.05$).

NAEYC

In general, significant differences were noted between NAEYC-accredited programs and non-NAEYC-accredited programs for the preschool sample ($t=4.85$, $p<.01$) on the Overall Average Item Score and across all ERS subscales (t range= 2.58 - 6.04 , $p<.05$), with the exception of the Personal Care Routines Subscale in which no significant differences were found. To make the MA QRIS system function more effectively and efficiently, EEC may want to consider leveraging outside efforts to promote quality as a means of verifying levels. As such, analyses were conducted to determine the associations between NAEYC accreditation and observed classroom quality by MA QRIS Level.

Preschool

Table 137 presents the descriptive statistics of NAEYC-accredited preschool classroom by MA QRIS level. It should be noted that only two programs at Level 3 were not NAEYC accredited and therefore, were not included in the table below.

Table 137. NAEYC descriptive statistics for preschool classrooms by MA QRIS Level

	Valid N	Mean Overall Average Item Score	Std. Dev.	Minimum	Maximum	Range
Level 1, Not NAEYC Accredited	30	3.63	.79	2.05	5.41	3.36
Level 1, NAEYC Accredited	9	4.20	.59	3.26	5.12	1.86
Level 2, Not NAEYC Accredited	12	3.98	.88	2.38	5.29	2.90
Level 2, NAEYC Accredited	27	4.19	.71	2.98	5.76	2.78
Level 3, NAEYC Accredited	39	4.69	.69	3.17	5.88	2.72

Although the number of NAEYC-accredited preschool programs at Level 1 was small, marginal significant differences were noted in the Overall Average Item Score between NAEYC-accredited programs and non-accredited programs ($t=2.00$, $p=.05$). This marginally significant finding may be due to the relatively low sample size of NAEYC-accredited programs at Level 1.

As can be seen in Table 135, Level 1 NAEYC-accredited programs and Level 2 NAEYC-accredited programs had almost identical means, and as a result no significant differences were noted between Level 1 and Level 2 NAEYC-accredited programs. Additionally at Level 2, no significant differences were noted among NAEYC-accredited programs and non-accredited programs on ERS Overall Average Item Scores and across the various ERS subscales. To further understand this relationship, the associations between Level 2 NAEYC-accredited

programs and Level 3 NAEYC-accredited programs were explored. Results indicated that Level 3 NAEYC-accredited programs scored significantly higher than Level 2 NAEYC-accredited programs for the Overall Average Item Score ($t=-2.86$, $p<.01$). Level 2 NAEYC-accredited programs had a mean that fell between Level 2 non-accredited programs and Level 3 accredited programs. It should be noted that the mean for Level 2 NAEYC-accredited programs meets MA QRIS Level 2 Overall Average Item requirements, but NAEYC-accredited Level 2 programs did not meet the Overall Average Item requirement for MA QRIS Level 3.

Infant and toddler

Similar analyses were replicated with infant and toddler classrooms in the sample. For the sample, NAEYC-accredited infant and toddler classrooms had significantly higher observed quality in comparison to non-NAEYC classrooms ($t=4.48$, $p<.01$). The descriptive data for accredited and non-accredited infant and toddler classrooms by level are presented below (Table 138).

Table 138. NAEYC descriptive statistics for infant and toddler classrooms by MA QRIS Level

	Valid N	Mean Overall Average Item Score	Std. Dev.	Minimum	Maximum	Range
Level 1, Not NAEYC Accredited	8	3.70	.86	2.38	4.31	1.93
Level 1, NAEYC Accredited	13	3.29	.62	2.78	5.03	2.24
Level 2, Not NAEYC Accredited	9	4.09	.54	2.71	4.17	1.46
Level 2, NAEYC Accredited	22	3.34	.76	2.91	5.46	2.54
Level 3, NAEYC Accredited	14	4.11	.59	3.17	5.50	2.33

Similar to preschool classrooms, significant differences in observed quality as measured by the ERS were noted between Level 1 NAEYC-accredited programs and Level 1 NAEYC non-accredited programs ($t=2.01$, $p=.05$). At Level 2, no significant differences were noted among Overall Average Item Scores or among the various subscales of the ERS between NAEYC-accredited and NAEYC non-accredited programs. Similar to preschool classrooms, the sample had too few programs at Level 3 that were not NAEYC accredited and therefore no comparison could be made between NAEYC and non-NAEYC programs at Level 3. Analyses comparing Level 2 NAEYC-accredited programs with Level 3 NAEYC-accredited programs indicated that there were no significant differences between the two QRIS levels on the Overall Average Item Scores. The data, however, indicate that Level 2 and 3 NAEYC-accredited programs on average did not meet ERS standards for MA QRIS Level 3, which requires an overall average item score of a 4.5. This was also true for the sample of infant and toddler programs, at large. As such, the lack of distinction in observed quality between Level 2 and Level 3 NAEYC-accredited programs may be partially due to the lower levels of quality found among infant and toddler classrooms in comparison to the preschool sample.

Head Start

Most of the Head Start programs in the sample were verified at MA QRIS Level 3. Based on the random selection of programs for the study at Levels 1 and Levels 2, there were eight Head Start programs at MA QRIS Level 2 and no Head Start programs at MA QRIS Level 1. Analyses indicated no significant difference in observed

classroom quality, as measured by the ERS Overall Average Item Score, between Head Start programs and community-based non-Head Start Level 3 preschool providers. As such, Head Start programs at Level 3 and community-based Level 3 programs appear to be offering comparable quality. A review of the subscales found only one significant difference between Head Start and non-Head Start programs at MA QRIS Level 3, with non-Head Start programs scoring significantly higher than Head Start programs ($t=3.02$, $p<.05$) on the Personal Care Routines Subscale (see Table 139).

Table 139. Means for Level 3 preschool Head Start and Non-Head Start classrooms

Subscale	Program Type	Valid N	Mean	Std. Dev.
Overall Average Item score	Head Start	23	4.55	.59
	Non Head Start	18	4.84	.76
Space and Furnishings	Head Start	23	3.72	.76
	Non Head Start	18	3.95	1.02
Personal Care Routines	Head Start	23	2.92	.61
	Non Head Start	18	3.61	.84
Language and Reasoning	Head Start	23	5.30	.91
	Non Head Start	18	5.03	1.02
Activities	Head Start	23	4.72	.79
	Non Head Start	18	4.94	.92
Interactions	Head Start	23	5.42	1.68
	Non Head Start	18	4.55	1.41
Program Structure	Head Start	23	4.94	1.06
	Non Head Start	18	5.16	1.34
Parents and Staff	Head Start	23	6.26	.35
	Non Head Start	18	6.04	.50

Because of the limited number of Head Start programs at Level 2 ($n=8$), analyses of significant differences between Level 2 and Level 3 were not conducted. An analysis of descriptives of Level 2 Head Start programs indicated that Level 2 Head Start programs are consistently meeting MA QRIS Level 2 quality standards across all ERS subscales. Level 2 Head Start programs, however, were not consistently meeting ERS quality standards for MA QRIS Level 3. All the Level 2 Head Start programs did meet MA QRIS Level 3 standards of an average ERS score of a 4 for the Language and Reasoning Subscale. On the other subscales approximately 25 percent to 50 percent of Level 2 Head Start programs did not meet MA QRIS Level 3 standards. In an effort to make the MA QRIS system more efficient, data suggest that Head Start programs could potentially enter the MA QRIS at Level 2 based on observed classroom quality data.

Level 4

For the sample, five programs were at Level 4. As noted in the ERS quality section, Level 4 programs' Overall Average Item Score was a 5.0, which is below the current MA QRIS ERS standards of an Overall Average Item Score of 5.5 (Table 140). Because of the small sample size, the data do not lend themselves to analyses. As such, the programs were looked at descriptively.

Table 140. Level 4 ERS Descriptive Statistics for Preschool Classrooms by MA QRIS Level

	Valid N	Mean Overall Item	Std. Dev.	Minimum	Maximum	Range
Overall average item score	5	5.00	.57	4.26	5.81	1.55
Space and Furnishings	5	4.43	.93	3.13	5.75	2.63
Personal Care Routines	5	3.18	.59	2.33	3.83	3.18
Language and Reasoning	5	6.15	.14	6.00	6.25	.25
Activities	5	5.19	1.13	3.89	6.90	5.19
Interactions	5	5.04	.97	3.80	6.00	2.20
Program Structure	5	5.10	1.35	3.25	7.00	3.75
Parents and Staff	5	6.27	.30	6.00	6.67	.67

As indicated, on average, programs met Level 4 standards for all ERS subscales except for the Personal Care Routine Subscale. As noted in the ERS quality section, the minimal Overall Average Item Score requirements exceed the minimal requirements for the individual ERS subscales. As such, if programs only meet minimal standards across the seven subscales, the Overall Average Item requirements will not be met. A review of the programs indicated that two of the five programs that did not consistently meet MA QRIS Level 4 standards were verified based on observation data that exceeded three years. Of the recently verified programs, one program did not meet standards. This program was from the same region, identified in the analyses of the Level 3 classroom data, as having lower scores in comparison to other regions. This further underscores the need for continued standardization in verification as well as, the need for a reassessment process after a MA QRIS Level has been verified to account for the potential drift in practices, changes to the program and curricula, and staff turnover. It should be noted that significant changes have taken place in the verification process and the training of Program Quality Specialists to improve uniformity and standardization across the regions.

Summary and implications

The above data appears to indicate the NAEYC-accredited programs and Head Start programs are, on average, meeting MA QRIS classroom quality standards for Level 2. In the interest of efficiency, sustainability, and reducing the administrative burden on programs and the MA QRIS EEC administrators, data support an entry point of the NAEYC-accredited programs and Head Start programs at MA QRIS Level 2.

Research indicated that curriculum is an important factor related to classroom quality. The MA QRIS requires curriculum training beginning at Level 2, but it does not specify what the curriculum training needs to include or timeframes for the training to ensure it is recent and up to date. MA EEC may want consider developing parameters for curriculum training, particularly for programs that are not utilizing a vetted curriculum with established training protocols. Additionally, programs are not required to use a curriculum that “is aligned with MA guidelines for Preschool Learning Standards and the Infant/ Toddler Learning Guidelines” until Level 4 of the MA QRIS. As a result, MA EEC may want to shift this requirement to a lower MA QRIS Level as a potential mechanism for raising curriculum standards and subsequently observed classroom quality.

Data on Level 4 highlight the need for re-verification of program quality at Levels 3 and 4, as drift was evident over time in quality classroom practices. Additionally, data indicate a need to tightened verification processes and provide standardized verification processes across MA QRIS regions.

Section V: Education, Experience and the Massachusetts QRIS

Analyses were conducted to explore the relations among classroom quality and the professional qualification of educators and supervisors/administrators. Data were collected using a staff roster in which programs were asked to provide information regarding the qualifications of the designated supervisor/administrators as well as of the educators from the observed classroom(s). Programs were also asked to indicate the lead teacher(s), the teacher(s) and/or the assistant teacher(s) for each classroom observed. In the cases where more than one lead teacher was designated or no lead teacher was indicated, the educator with the highest professional qualifications starting with educational level was considered the lead teacher for the purpose of analyses.

Preschool lead teacher education and classroom quality

Descriptive statistics indicated that 70 percent of preschool lead teachers had a degree in ECE (either an Associate's or Bachelor's degree). About 20 percent of lead teachers had only some college experience, with no degree earned. In addition, 87 percent of preschool lead teachers had more than 5 years of experience working at an early care and education program, with lead teachers having an average of 14 years of experience.

To examine the associations between classroom quality and teacher education, an ANOVA ($F=9.65$, $p<.01$) indicated that observed classroom quality as measured by the ECERS-R varied significantly by lead teacher's level of education. A Tukey HSD post hoc analysis indicated that preschool classrooms with lead teachers who had a BA degree had significantly higher quality scores on the Overall Average Item Score of the ECERS-R ($p<.05$) as well as on the Language-Reasoning Subscale ($p<.05$), the Activities Subscale ($p<.05$), Program Structure ($p<.05$), and Parents and Staff Subscale ($p<.05$) (see Table 141 for descriptive data). Differences in overall average item score between lead teachers with an AA and lead teachers with a BA was approaching significance ($p<.06$) and was significant for the Space and Furnishings ($p<.05$) Subscale and the Parents and Staff Subscale ($p<.01$).

Table 141. Preschool classroom quality by lead teacher education

ECERS-R	Degree of lead teacher	N	Average	Standard deviation
Overall Average Item Score	Some College	23	3.68	.75
	AA	21	3.94	.82
	BA	69	4.41	.74
Space and Furnishings	Some College	23	3.31	1.11
	AA	21	3.11	.62
	BA	69	3.76	.84
Personal Care Routines	Some College	23	2.73	.80
	AA	21	2.51	.61
	BA	69	2.97	.81
Language-Reasoning	Some College	23	4.14	1.22
	AA	21	4.16	1.02
	BA	69	4.87	1.03

ECERS-R	Degree of lead teacher	N	Average	Standard deviation
Activities	Some College	23	3.71	.66
	AA	21	4.40	.89
	BA	69	4.49	.91
Interactions	Some College	23	3.68	1.39
	AA	21	3.98	1.70
	BA	69	4.57	1.65
Program Structure	Some College	23	3.89	1.27
	AA	21	4.05	1.39
	BA	69	4.79	1.28
Parents and Staff	Some College	23	4.70	.91
	AA	21	5.29	.99
	BA	69	5.90	.66

Analyses were also run to explore associations among type of degree and classroom quality. T-tests indicated that classrooms with educators who had a BA in ECE or a related field (mean=4.49) versus an AA in ECE or a related field (mean=3.98) scored significantly higher ($t = -2.72, p < .01$) on the Overall Average Item Score for quality as measured by the ECERS-R. This result was not dependent on the inclusion of related fields. If degrees were limited to ECE only, without considering degrees in related fields, significant differences ($t = -4.05, p < .01$) remained in classroom quality (as measured by the Overall Average Item Score of the ECERS-R) between programs with lead teachers with BA degrees in ECE (mean=4.78) and lead teachers with AA degrees in ECE (mean= 3.97).

Infant and toddler lead teacher education and quality

Among lead teachers in infant/toddler classrooms, teachers had an average of 11.9 years of experience in early childhood education, with 24 percent of educators having fewer than 5 years of experience. Infant and toddler lead teachers had 7.3 years of experience working at their current center, with 62 percent having worked at their current center more than 5 years. More than half (54%) of infant/toddler lead teachers reported having a degree in ECE (either AA or BA). Descriptive data on infant and toddler classroom quality by degree held are presented below in Table 142.

Table 142. Infant/toddler classroom quality by lead teacher education

ITERS-R	Degree of lead teacher	N	Average	Standard deviation
Overall Average Item Score	Some College	25	3.47	.68
	AA	14	4.03	.77
	BA	22	3.92	.65
Space and Furnishings	Some College	25	3.01	.71
	AA	14	3.48	1.20
	BA	22	3.21	.86

ITERS-R	Degree of lead teacher	N	Average	Standard deviation
Personal Care Routines	Some College	25	2.40	.65
	AA	14	2.59	.94
	BA	22	2.47	.54
Language-Reasoning	Some College	25	3.93	1.02
	AA	14	4.43	1.35
	BA	22	4.59	.67
Activities	Some College	25	2.79	.89
	AA	14	3.32	.95
	BA	22	3.20	1.43
Interactions	Some College	25	4.14	1.33
	AA	14	5.07	1.57
	BA	22	5.71	1.26
Program Structure	Some College	25	3.47	1.35
	AA	14	3.61	1.51
	BA	22	3.96	1.55
Parents and Staff	Some College	25	4.98	.85
	AA	14	5.87	.49
	BA	22	5.54	.65

An ANOVA indicated significant differences in quality by degree of the lead teacher (BA vs. AA, vs. some college) for the Overall Average Item Score on the ITERS-R ($F=3.88$, $p<.05$). Significant differences by lead teacher degree was also found on the Interaction Subscale ($F=3.80$, $p<.05$); the Parents and Staff Subscale ($F=7.78$, $p<.01$); and approached significance for the Listening and Talking Subscale ($F=2.72$, $p<.10$).

Post hoc analyses using the Tukey HB test indicated a significant difference in the Overall Average Item Score in classrooms in which the lead teacher had some college versus an AA degree ($p<.01$) and a difference that was approaching significance between classrooms in which educators had some college and a BA degree ($p<.10$). Significant differences were also found between classrooms in which the lead teacher had some college and a BA for the Interaction Subscale ($p<.05$). Significant differences in quality on the Parents and Staff subscale were noted among classrooms in which the lead teacher had some college versus an AA degree ($p<.01$) and between classrooms in which the lead teacher had some college and a BA degree ($p<.05$). For the Listening and Talking subscale, the difference in quality between classrooms with lead teachers that had some college in comparison to classrooms with lead teachers that had a BA degree approached significance ($p<.10$).

No significant differences in classroom quality were found between lead teachers with BA or AA degrees in ECE for infant and toddler classrooms. (Note that all teacher with degrees in ECE identified their degrees as being ECE-specific degrees.)

Supervisor/administrator education and quality

Descriptive statistics indicated that the vast majority of supervisors/administrators had earned a BA degree (78%). About 8 percent of administrators had some college with no degree. Among the educators with degrees (AA, BA, or higher), 91 percent had degrees in ECE or a related field. Of the administrators with a BA, almost all of them (98%) had a degree in ECE or a related field. About 84 percent of supervisors/administrators had at least 5 years of experience as supervisors/administrators; on average, supervisors/administrators reported 14 years of experience as an ECE administrator.

Analyses of supervisor/administrator education (supervisor as defined by the MA QRIS as the person with primary responsibility for the supervision of teachers) also found significant relations among supervisor/administrator education level and observed classroom quality. Specifically, t-tests indicated that, for programs with both infant/toddler and preschool classrooms, the average quality, as indicated by the Overall Average Item Scores for both measures (average of infant/toddler ITERS-R and preschool ECERS-R Overall Average Item Scores), was significantly higher for programs with a supervisor/administrator who had a BA in EEC or related field versus programs with a supervisor/administrator who had less than a BA in ECE or related field ($t=3.57, p<.01$). Analyses were limited to a comparison of administrators in ECE or a related field because all administrators with a BA or higher reported having a BA in EEC or a related field.

The same pattern was found when preschool classroom quality scores and infant and toddler classroom quality scores were considered in separate analyses. Preschool classrooms with program administrators/supervisors who had a BA or higher in ECE or a related field had significantly higher ECERS-R scores than programs in which the supervisor/administrator did not have at least a BA in ECE or related field ($t=3.52, p<.01$). Infant and toddler classrooms with program supervisors/administrators who had a BA or higher in ECE or a related field had significantly higher ITERS-R scores than programs in which the supervisors/administrators did not have at least a BA in ECE or related field ($t=3.46, p<.01$; (see Table 143 below).

Table 143. Classroom quality by education of the supervisor/administrator

ERS	Degree of administrator/supervisor	N	Average	Standard deviation
Average of preschool and infant and toddler Overall Average Item Score	No BA in ECE or related field	17	3.37	.67
	BA in ECE or related field	54	3.99	.46
Preschool ECERS-R Overall Average Item Score	No BA in ECE or related field	29	3.77	.80
	BA in ECE or related field	95	4.37	.79
Infant and toddler ITERS-R Overall Average Item Score	No BA in ECE or related field	18	3.29	.48
	BA in ECE or related field	55	3.92	.77

Education, classroom quality, and MA QRIS Level

Analyses were also conducted to determine the associations among educator and supervisor/administrator education and MA QRIS Level by examining differences in observed classroom quality by level with the education level of the lead teacher or supervisor/administrator held constant.

Preschool lead teacher education, classroom quality and MA QRIS level

Because only 10% of educators with degrees at Level 3 had degrees outside of ECE/related fields, the following analyses were restricted to educators with an AA, BA or higher degree in ECE or a related field. Table 144 presents the ECERS-R Overall Average Item scores for the three MA QRIS Levels.

Table 144. Quality among programs in which lead teacher had a degree (AA, BA or higher) in ECE or related filed by MA QRIS Level

ERS	MA QRIS Level	N	Average	Standard deviation
Preschool ECERS-R Overall Item Average Score	Level 1	23	3.91	.78
	Level 2	24	4.36	.75
	Level 3	35	4.72	.65

Among those preschool classrooms with a lead teacher who had a degree in ECE or a related field, significant differences existed in the overall average item score for classroom quality by MA QRIS Level ($F=8.95, p<.01$). Post Hoc analyses using the Tukey HSD showed that significant differences in classroom quality existed between MA QRIS Levels 1 and Levels 3 ($p<.01$).

To further examine differences, analyses were re-run limiting the data to lead teachers who had BA degrees or higher specifically in ECE. Descriptive data are presented below for these groups of educators (see Table 145).

Table 145. Quality among programs in which the lead teachers has a BA or higher ECE (excluding related fields) by MA QRIS Level

ERS	MA QRIS Level	N	Average	Standard deviation
Preschool Overall Item Average Score	Level 1	14	4.16	.76
	Level 2	18	4.50	.72
	Level 3	30	4.72	.68

ANOVA indicated that considering only classrooms with lead teachers with a BA or higher in ECE, significant differences were still evident in classroom quality by MA QRIS Level ($F=2.98, p=.05$). Post hoc analyses indicated a significant difference in quality between classrooms at Level 1 versus Level 3 ($p<.05$).

To summarize, although teacher education was significantly related to observed classroom quality, differences in observed quality for preschool classrooms was still evident by MA QRIS Level when holding teacher education constant. This holds true when analyses were limited to teachers with a BA in ECE (with related fields excluded).

Infant and toddler lead teacher education, classroom quality, and MA QRIS Level

Descriptive data for infant and toddler classroom quality by Level with education limited to only those classrooms in which the lead teacher had a degree in ECE or a related field are presented in Table 146.

Table 146. Teacher had a degree (AA or BA or higher than a BA) in ECE or related field

ERS	MA QRIS Level	N	Average	Standard deviation
Infant and toddler Overall Item Average Score	Level 1	9	3.73	.72
	Level 2	15	4.01	.72
	Level 3	11	4.18	.60

Due to small sample sizes, statistical testing for significance was not possible. However, descriptive data indicated a trend in which quality increases with level for classrooms in which the lead teacher had a degree (either AA or BA) in ECE or a related field. Additional data would be needed to determine the significance of this trend.

Supervisor/ administrator education, classroom quality and MA QRIS Level

Similarly, analyses also examined supervisor/ administrator education in terms of its relation to classroom quality by MA QRIS Levels. Due to a lack of supervisor/ administrator without a BA in ECE or a related field at the upper levels, ANOVAs and Post Hoc analyses focused on comparisons among programs in which the supervisor/ administrator had a BA or higher in ECE or a related field. A series of ANOVAs indicated that, for the Preschool Overall Average Item Scores, significant differences existed among MA QRIS Levels, holding supervisor/administrator education constant (F=6.87, p<.01) Significant differences were also found for the combined Overall Average Item Scores across preschool and infant/toddler classrooms for programs with both types of classroom in the study (F=3.82, p<.05). Post Hoc Analyses were conducted to further examine the differences by level. Results are presented below in Table 147.

Table 147. Significant differences in classroom quality by MA QRIS Level for programs with a supervisor/ administrator that had BA or higher in ECE or related field

Program type	Significant differences between levels	Significance
Average of preschool and infant and toddler Overall Item Average Score (for programs with both infants/toddlers and preschool classrooms)	Level 1 & Levels 3	p<.05
Preschool Overall Item Average Score	Level 1 & Levels 3	p<.05
	Level 2 & Levels 3	p<.01
Infant and toddler Overall Item Average Score	Level 1 & Levels 3	P<.10

As can be seen in the Table 147, significant differences existed in the quality of preschool classrooms between levels 1 & 3 and Levels 2 & 3, among programs in which the administrator has a BA or higher in ECE or related field. For infant and toddler classrooms, the differences were not as strong, with the difference between Levels 1 and Levels 3 approaching significance. It should be noted that the Level 3 programs with infant and toddler classrooms, on average, did not meet the MA QRIS overall ERS item average, which is in part contributing to the lack of disparity among levels. When considering the quality averaged across preschool and infant/toddler classrooms (for programs that had both types of classrooms), significant differences were found between Levels 1 and 3. These analyses indicated that although supervisor/administrator education distinguished quality, significant differences in quality between MA QRIS Levels still remain when holding administrator education constant.

Experience and quality

Analyses also considered the relations among educator and supervisor/administrator experience and classroom quality.

Experience of the lead teacher and preschool classroom observed quality

The associations between preschool lead teacher experience and observed classroom quality were explored through correlational analyses. Results are presented below in Table 148.

Table 148. Correlations between preschool lead teachers' experience and quality measures

Quality Variables	Years of experience as a teacher in an early education and care program	Years at the current ECE program
Overall Average Item Score	.10	.03
Subscale Space and Furnishings	-.08	-.07
Subscale Personal Care Routines	-.01	-.01
Subscale Language and Reasoning	.14	.14
Subscale Activities	3	.13
Subscale Interactions	-.08	-.01
Subscale Program Structure	-.04	-.04
Subscale Parents and Staff	5	.04

^p<.10. *p<.05. **p<.01.

No significant correlations were found between preschool lead teacher years of experience as a teacher in an ECE program (ECE experience) and preschool teachers' years of experience at their current program with observed classroom quality as defined by the ERS. It is important to note, that the sample of lead preschool teachers had many years of experience, with 87% having at least 5 years of experience in an ECE program and 62% having over 10 years of experience at an ECE program.

Experience of the lead teacher and infant toddler classroom observed quality

Similar analyses were undertaken for infant and toddler classrooms (see Table 149). Results indicated that years of ECE experience and years of experience at the current ECE program were significantly related to the Overall Average Item Score on the ITERS-R and several subscales of the ITERS-R. Thus, the data indicate a stronger association between ECE experience and classroom quality for infant and toddler programs in comparison to preschool classrooms. It is important to underscore, that the infant and toddler teachers tended to have less experience than the preschool teacher with about 24% having worked in ECE for less than 5 years. As such, this greater range of variability in experience among infant and toddler lead teachers, particularly at the lower range of experience, may be a contributing factor to the discrepancy in findings between preschool and infant and toddler classrooms.

Table 149. Correlations among infant and toddler classroom quality and lead teachers’ years of experience

Quality Variables	Years of Experience as at an early education and care program	Years at the ECE program
Overall Average Item Score	.29*	.28*
Subscale Space and Furnishings	.37**	.27*
Subscale Personal Care Routines	.08	.12
Subscale Listening and Talking	.25*	.23^
Subscale Activities	.24^	.29*
Subscale Interactions	.14	.07
Subscale Program Structure	.11	.07
Subscale Parents and Staff	.27*	.32*

^p<.10. *p<.05. **p<.01.

Supervisor/administrator experience and quality

When examining associations among supervisor/administrator experience and classroom quality, no significant associations were found between supervisor/administrator experience and the Overall Average Item Scores for preschool classrooms, the Overall Average Item Scores for infant toddler classrooms or the Overall Average Item Scores when averaging the overall quality of the preschool and infant toddler classroom (for programs with both preschool and infant/toddler classrooms) (see Table 150). Additionally, no significant relationships were found between supervisor/administrator experience and the subscales of the ITERS-R or ECERS-R. It is important to note that administrators in the sample were largely experienced with the vast majority (84%) having at least 5 years of experience as a supervisor/administrator.

Table 150. Correlations among quality and years of experience of administrator/supervisor

Quality Variables	Years of experience as an administrator/supervisor of an early education and care program
Average of preschool and infant and toddler overall average item score	-.10
Preschool overall average item score	.03
Infant and toddler overall average item score	-.14

Relations among classroom quality, experience and education

To further explore associations among quality, experience and education, regression analyses were conducted to estimate the amount of variability in classroom quality accounted for by these key variables.

Relations among classroom quality and preschool lead teachers’ experience and education

Regression analyses were conducted to determine the amount of variability in quality accounted for by teacher education, a degree in ECE or related field and years of lead teacher ECE experience. Regression analyses indicated that when considering all three variables, the model was significant with about 20 percent of the variability in classroom quality accounted for by the model (Table 151). When looking at the individual contribution of variables, both having a BA degree and a Degree in ECE were significant predictors in the model.

Table 151. Regression model predicting classroom quality by preschool lead teachers' experience and education

Variables	B	Standard Error	Standardized β
BA degree	.55	.08	.33**
Degree in ECE	.47	.09	.20**
Years of experience	.01	.00	.05
Adjusted R Square	.20 *		

*p<.05. **p<.01.

Relations among classroom quality infant and toddler lead teachers' experience and education

Analyses were repeated for infant and toddler classrooms and a different pattern emerged. The model accounted for about 14 percent of the variability in infant and toddler observed classroom quality (Table 152). Analyses of the individual variables indicated that when considered as a whole, years of experience was the only variable that contributed significantly to the model.

Table 152. Regression model for predicting classroom quality by infant and toddler lead teachers' experience and education

Variables	B	Standard Error	Standardized β
BA degree	.18	.23	.12
Degree in ECE	.28	.23	.18
Years of experience	.03	.01	.28*
Adjusted R Square	.14 *		

*p<.05. **p<.01.

Relations among classroom quality and administrators' experience and education

Similarly, regression analyses indicated the model including the supervisor/administrator having a BA degree, the supervisor/administrator having a degree in ECE and the supervisor/administrator Years of Experience accounted for about 10 percent of the variability in quality for preschool classrooms, with only a BA degree contributing significantly to the model (Table 153).

Table 153. Regression model for predicting preschool classroom overall average item score by administrators' experience and education

Variables	B	Standard error	Standardized β
BA degree or higher	.74	.19	.37**
Degree in ECE	-.17	.27	-.06
Years of experience	.00	.01	-.01
Adjusted R square	.10 *		

*p<.05. **p<.01.

Similar analyses were repeated for infant and toddler classrooms and for the combined Overall Average Item Scores score for preschool classrooms and infant and toddler classrooms when the program had both types of classrooms in the study. Tables 154 and 155 present findings below. Regardless of the quality measure used, the supervisor/administrator having a BA degree was the only significant predictor of classroom quality, as measured by the ERS Overall Average Item Score; having a degree in ECE and years of experience were not significantly associated with quality, after considering level of education.

Table 154. Regression model for predicting infant and toddler Overall Average Item Scores by administrators' experience and education

Variables	B	Standard Error	Standardized β
BA degree	.73	.24	.43**
Degree in ECE	-.20	.35	-.07
Years of experience	-.01	.01	-.15
Adjusted R Square	.13 *		

*p<.05. **p<.01.

Table 155. Regression model for predicting the combined overall average item score (for infant/toddler and preschool classrooms) within a program by administrators' experience and education

Variables	B	Standard Error	Standardized β
BA degree	.75	.22	.47**
Degree in ECE	-.29	.32	-.13
Years of experience	-.01	.01	-.12
Adjusted R Square	.14 *		

*p<.05. **p<.01.

Relations among classroom quality and the education of the lead teacher and support staff

In addition to the above findings, researchers conducted analyses to determine the potential influence of the education level of educators other than the lead teacher. Regression analyses were conducted to determine the amount of variability in the Overall Average Item Score in preschool classrooms that was accounted for by the Lead-teacher BA, lead teacher degree in ECE or related field and lead teacher years of experience, as well as, teacher/assistant teacher BA, teacher/assistant teacher education degree in ECE or related field and teacher/assistant teacher education years of experience for classroom. Data on both the lead teacher and other classroom teachers were available for 90 classrooms. For those classrooms, the model was run using just the lead teacher variables and then re-rerun with the addition of the assistant teacher variables. Results indicated that both models were significant ($p<.01$). When considering just the lead teacher BA, lead teacher degree in ECE or related field and lead teacher years of experience, 13 percent of the variability in the overall Average Item Score was accounted for in the model. When the same variables for the assistant teacher were added to the model, 15 percent of the variability was accounted for by the model (Table 155). As can be seen in Table 156, while adding the assistant teacher variables increases the R Square slightly, the individual assistant teacher variables are not statistically significant, although assistant teacher experience approaches significance

Table 156. Regression model for predicting classroom quality by preschool teachers' experience and education

Variables	B	Standard Error	Standardized β
Lead teacher BA degree	.59	.19	.35*
Lead teacher degree in ECE	.17	.22	.09
Lead teacher years of experience	.01	.01	.07
Teacher/assistant teacher BA degree	.25	.23	.15
Teacher/assistant teacher degree in ECE	-.02	.01	-.01
Teacher/assistant teacher years of experience	.02	.22	.19^
Adjusted R Square	.15**		

^p<.10. *p<.05. **p<.01.

Summary and implications

Data indicated that preschool teacher education level and supervisor/administrator education level were related to classroom quality, with the greatest differences in classroom quality found between those educators who had some college versus those that had earned a BA. Despite this association, differences in quality by MA QRIS Level were still evident when holding lead teacher and supervisor/administrator education constant for both preschool and infant and toddler classrooms. In addition, years of experience in ECE was the only significant predictor of quality in infant and toddler classrooms in a model that included teacher education and experience.

Section VI: Child Outcomes and the Massachusetts QRIS

Validation study findings show that early education and care programs at higher levels of the QRIS typically demonstrate higher observed process quality. A central question remains, however. To what extent is quality, as measured by the QRIS, associated with outcomes for children? This line of inquiry reflects an emerging area of interest for many QRIS systems, driven in part by a federal requirement that RTTT-ELC grantees conduct validation studies that assess the extent to which differences in quality levels are related to children's learning, development, and school readiness. Results from other validation studies have been mixed, with many reporting little or no association between QRIS level and outcomes.⁵⁹ However, differences in state approaches to both QRIS systems and their validation make it difficult to draw generalizations regarding the connection between QRIS and children's outcomes.

To address this important question, the validation study team conducted pre- and post-assessments with randomly selected children in study classrooms with a minimum of 6 months between pre and post. The study used several norm-referenced assessments, and gains were compared across MA QRIS levels with statistical controls for family background and child characteristics. These three features — pre-post design, multiple well-validated assessments, and controls for family background — characterize stronger research designs, as defined by the National Institute for Child Health and Human Development.⁶⁰

Preschool assessment and sample

Preschool children were randomly selected from one randomly selected preschool classroom at a participating program. To account for attrition, 6 children (3 boys and 3 girls) were assessed in the Fall and 4 of these children (2 girls and 2 boys) were assessed in the Spring. In total, 737 preschool children were assessed at pre-test and 482 children were assessed at post-test. The assessment battery was designed to reflect key components of school readiness as identified by prior research. The assessment battery included the Peabody Picture Vocabulary Test- 4th edition (PPVT-4), and two subtests from the Woodcock Johnson Tests of Achievement – 3rd edition (WJ-III TA): Letter Word Identification and Applied Problems. Both the PPVT-4 and the WJ-III TA are well-known and highly regarded assessment tools that have been used in many other QRIS validation studies and have been found to predict children's school readiness and later outcomes. Preschool children's social emotional development was assessed through teacher ratings as measured by Devereux Early Childhood Assessment for Preschoolers, Second Edition (DECA P2) and the Preschool Learning Behavior Scale (PLBS). Refer to Table 157 for a list of measures utilized in the current study and a description of the domains they assess.

⁵⁹ Karoly, L.A. (2014). Validation studies for early learning and care quality rating and improvement systems: a review of the literature. Working Paper Prepared for the Delaware Office of Early Learning. RAND Corporation.

⁶⁰ Does quality of child care affect child outcomes at age 4 1/2? (2003). *Developmental Psychology*, 39(3), 451–469.

Table 157. Child assessment measures and domains

Preschool Direct Child Assessment	Domain
The Peabody Picture Vocabulary Test, 4th edition ⁶¹	Receptive language
Woodcock-Johnson III Form A: Test of Achievement, Letter-Word Identification ⁶²	Pre-literacy skills
Woodcock-Johnson III Form A: ⁶³ Test of Achievement, Applied Mathematics	Early mathematics skills
Preschool Teacher Rating Assessment	Domain
The Devereux Early Childhood Assessment Preschool Program, Second Edition ⁶⁴	Social emotional development, related to relationships and working with others
Preschool Learning Behaviors Scale ⁶⁵	Academic-related social emotional competencies

The direct child assessment took between 25-40 minutes per child and was administered one-on-one in the child's program, in a quiet area with a familiar adult present. Child assessors underwent extensive training and had to demonstrate reliability prior to conducting assessments. Child assessments were conducted in English only. This was done, in part, to provide a standardized protocol for all English language and dual language learners, given that reliable and valid child assessment measures are not consistently available in languages other than Spanish and English, and children from many diverse languages were included in the study. The assessment battery began with the PPVT in order to establish a child's adequate understanding of English. If a basal (baseline) was not established on the PPVT in the Fall, the child was considered un-testable for the study and an alternative child was randomly selected. Pre-and post-assessments of the same children allowed for analysis of children's growth over time, not just end-point benchmarks. Thus, the diverse starting points of children from diverse backgrounds could be taken into account through the process of establishing baseline equivalency and through the emphasis on gains over time.⁶⁶ The sample characteristics and assessment pre-test scores are presented in Table 158.

⁶¹ Dunn, L.M. & Dunn, D. M. (2007).

⁶² Woodcock, R., McGrew, K. S., & Mather, N. (2001).

⁶³ Woodcock, R., McGrew, K. S., & Mather, N. (2001).

⁶⁴ LeBuffe, P.A. & Naglieri, J.A. (2012).

⁶⁵ McDermott, P. A., Green, L. F., Francis, J. M., & Stott, D. H. (2000).

⁶⁶ Zellman, G. L. & Fiene, R. (2012).

Table 158. Descriptive characteristics of participating children by MA QRIS level

	Sample N	Sample Mean	QRIS level		
			Level 1	Level 2	Level 3
Child-Level Characteristics					
% ELL	462	22.3	18.1	27.0	21.7
% Special Education	462	13.0	18.1	13.8	7.5
% Receive tuition subsidy	462	54.8	22.8	67.1	72.7
Direct Child Assessment pre-scores					
Mean Fall PPVT-4 score	462	104.1	107.1	102.7	102.7
Mean Fall WJ TA Letter Word Identification score	462	101.3	100.3	99.4	101.4
Mean Fall WJ TA Applied Problems score	462	104.9	107.1	103.8	103.5
Teacher Rating Scale pre-scores					
Mean Fall Total Protective Factors DECA score	453	50.8	50.6	49.5	52.2
Mean Fall Total PLBS score	442	52.1	51.8	50.7	53.6

In general on average, children in the sample scored above average across measures. The data highlights significant differences in the sample among the MA QRIS Levels. Of particular note, Level 3 has significantly higher percentages of children receiving a subsidy. Although these covariates are controlled for in the analytic models, the data suggest that a significant percentage of children in Level 3 programs represent high-needs populations with other potential risk factors that are not accounted for in the model (e.g. maternal depression, parental education). As such, it is important to note that one potential limitation to the analyses is that the models must rely on a restricted number of predictors and covariates.

Preschool children's gains

Children generally demonstrated significant gains over the course of the study (see Table 159).

Table 159. Developmental pre/post net gain by measure

	Level 1	Level 2	Level 3
	Net gain	Net gain	Net gain
PPVT	2.69***	0.99	3.25***
Letter Word	0.47	1.05	0.34
Applied Problems	3.67***	2.83***	1.89
DECA total	1.98	2.61**	2.81**
Initiative	3.10**	2.75***	1.90*
Self-Regulation	2.02*	1.21	1.21
Attachment	0.11	2.63*	4.25***
Behavioral Problems	-1.41	-0.91	0.51
PBLS	52.36***	51.66^^	51.57***

* p<.05. ** p<.01. *** p<.001.

As indicated in the table, children in programs assigned to Level 1 showed significant gains in scores on the PPVT, Applied Problems, the Initiative Subscale of the DECA, and the Self-Regulation Subscale of the DECA. Children in Level 2 showed significant gains in Applied Problems, the Initiative Subscale of the DECA, the Attachment Subscale of the DECA, and the DECA Total score. Children in Level 3 showed significant gains on the PPVT, Initiative Subscale of the DECA, Attachment Subscale of the DECA, and the Total score of the DECA.

Analytic modeling of preschool gains

Analytic models, using baseline equivalent preschool samples, were used to determine differences in outcomes by MA QRIS Level. An important consideration when comparing gains over time is the fact that children in different groups may start at different points; those different starting points may be associated with different growth trajectories (change over time). As a result, baseline equivalency is critical to determining differences in gains among distinctive groups. For this study, pre-test scores on all measures were compared for children in QRIS levels 1, 2 and 3 (Level 4 programs were not considered in these analyses due to the small sample size). To produce comparison groups that were within 0.25 standard deviations of one another — the limit for pre-intervention differences under What Works Clearinghouse standards when statistical controls are used — the baselines equivalent sample omits children appearing much higher or much lower than the rest of the children on baseline (pre-test) cognitive and social emotional scores.

Analytic models controlled for the child-level co-variates of special education, ELL, and subsidy status. For the social emotional outcomes (initiative, self-regulation, attachment, behavioral problems), age was also used as a control. Pre-test scores were also entered as controls in the model to control for the diverse starting points of children. These characteristics were centered at the mean of the analysis sample so the values in the tables represent the adjusted mean post-test scores for each MA QRIS Level.

To account for the nesting of children within classrooms, tests of differences in child outcomes across MA QRIS Levels were conducted in a multilevel multiple group structural equation modeling framework. This multilevel structural equation modeling allowed researchers to simultaneously model individualized child-level covariates (e.g., ELL, subsidies, special education) in conjunction with the shared influence of classroom-level variables

(classroom QRIS Level). This modeling strategy allows both the child-level (“Within”) and classroom-level (“Between”) of analysis to be estimated concurrently and the impact of QRIS Level to be accounted for at both the child and program QRIS level.

When comparing the adjusted post-test means among MA QRIS Levels using multi-level modeling, there were two significant differences: children in Level 3 showed significantly greater improvement in their PPVT scores over time than did those in Level 2 ($p < .05$) and significantly greater developmental gains in Attachment Subscale scores of the DECA than did those in Level 1 ($p < .05$), see Figures 22 and 23 below. Based on standard deviations of the outcome measures, the differences in gain scores across outcome measures represent a medium effect size.

Figure 22. Comparing adjusted mean post-test scores for PPVT and WJ-III among MA QRIS Levels

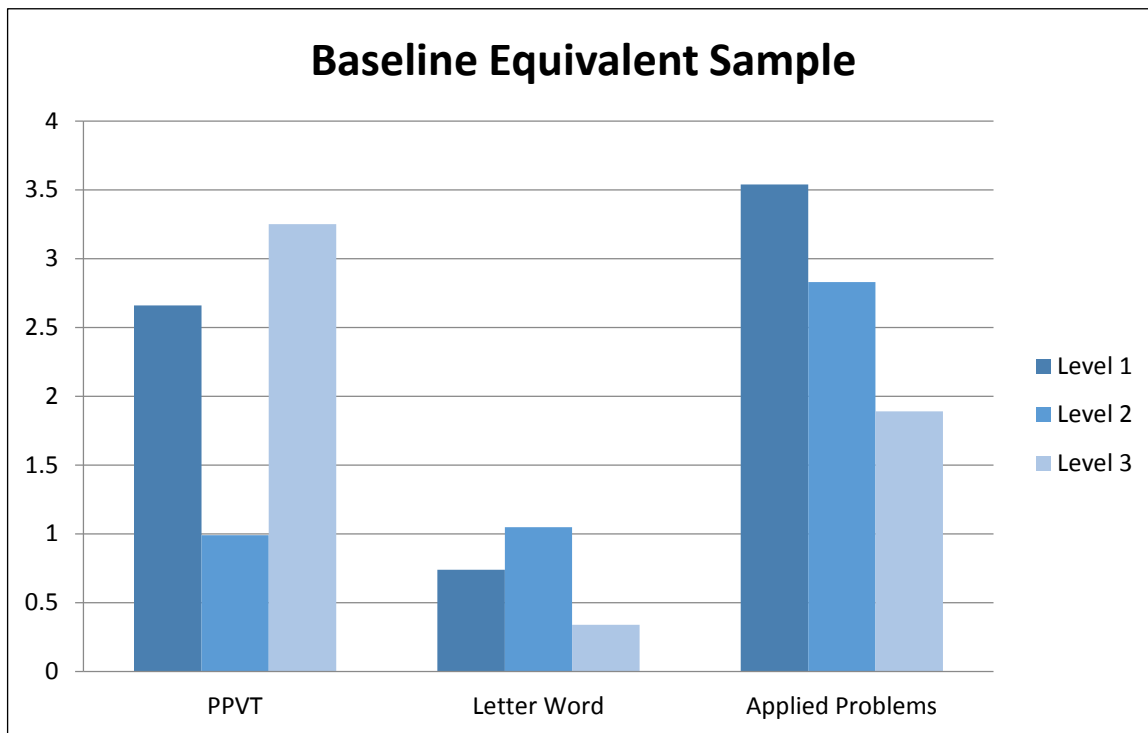
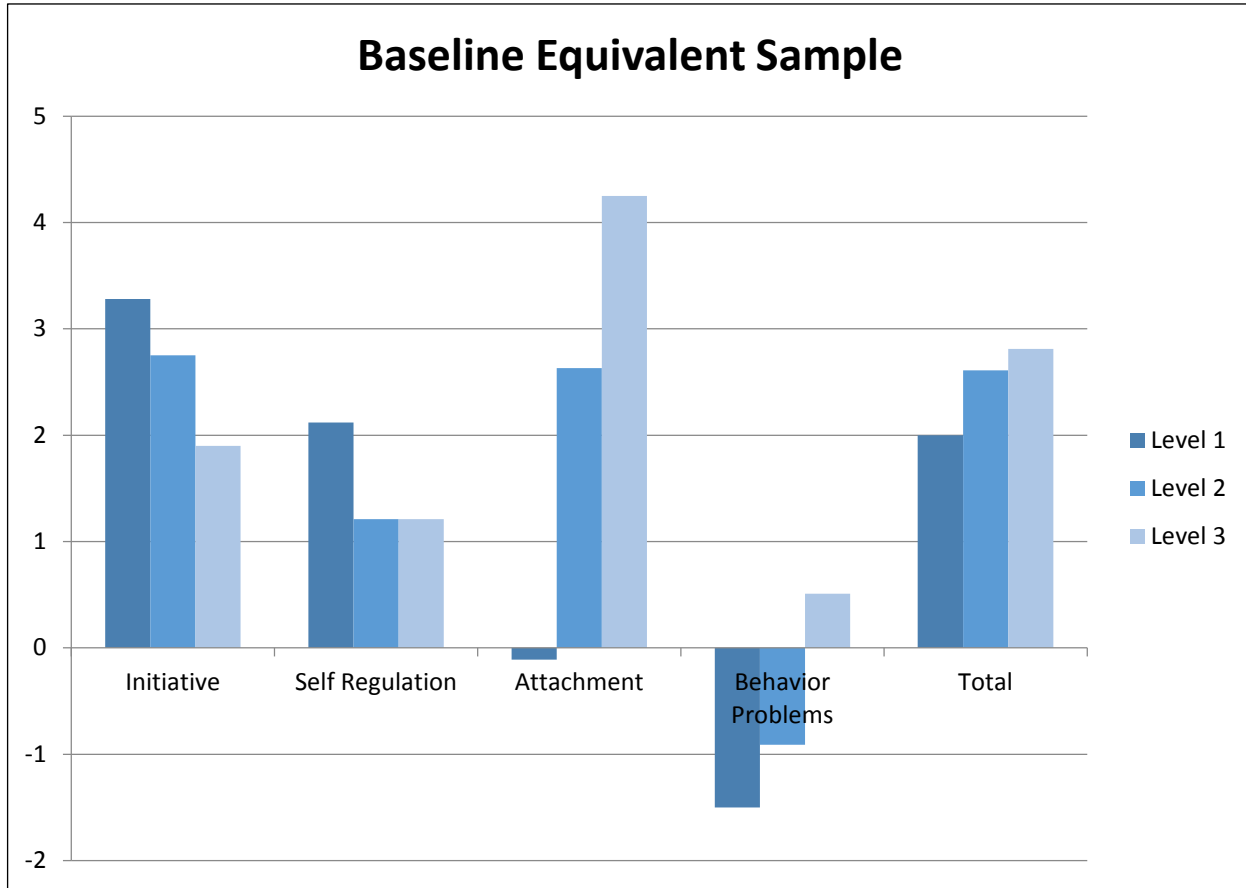


Figure 23. Pre/post outcomes gains for preschoolers and differences in change scores among MA QRIS Levels

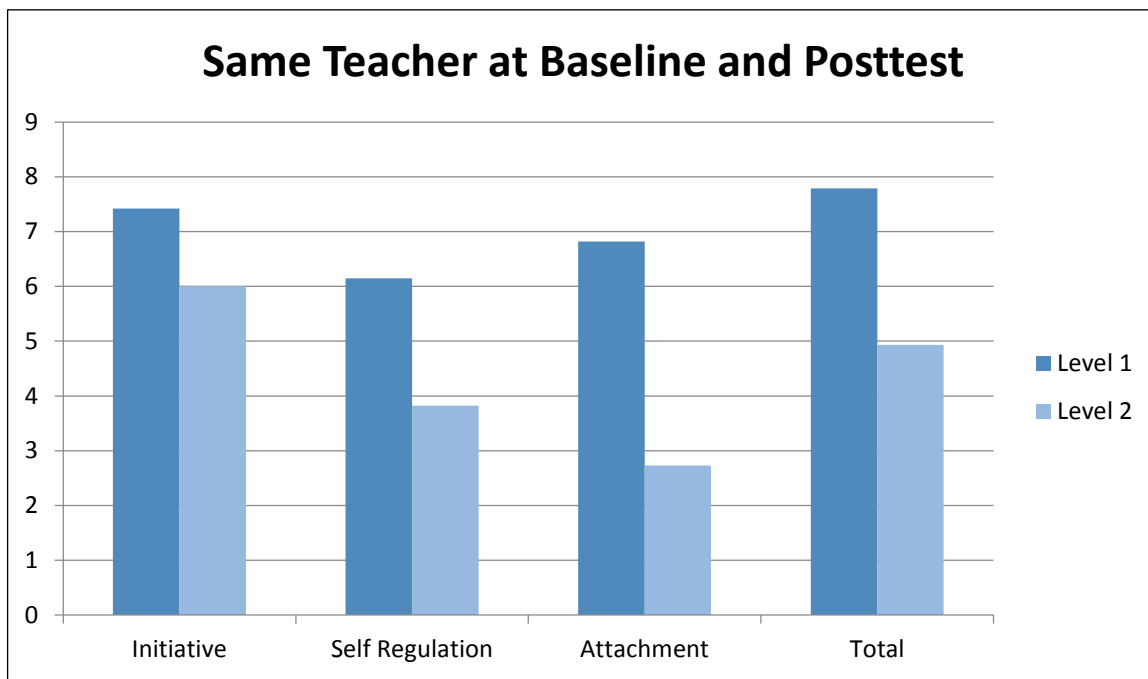


For the PBLs, an analytic approach that assumes there is an underlying continuum for the outcome variable that extends beyond the highest and lowest rating point on the measure was used in the study – that is, this approach assumes that other children, not assessed, might have scored lower or higher than the children in this sample. Values of the baseline scores, along with the demographic characteristics used as statistical controls are used to estimate these “out-of-range” scores. This method of modeling change is limited by the choice of predictors in the model and by the assumption that the relationship between a given predictor and the outcome score can be described by a straight line. In these models, no significant differences in gains for the total PLBS score by MA QRIS Level emerged.

Toddler developmental gains

Toddlers were also assessed pre and post on social emotional development, using the Devereux Early Childhood Assessment for Toddlers. Unlike preschool children, over the six month period between pre- and post-assessments, many toddlers had changes to their classrooms/teachers. As result, toddlers who were scored by different teachers at baseline and posttest were omitted, to reduce potential bias and measurement error. Additionally, some toddlers aged out of the measure during the 6 month period, turning 3 prior to the post assessment period and subsequently had to be excluded from the analyses. For the purpose of analyses, too few toddlers were in Level 3 and 4 programs (n<25) for the baseline equivalent sample to be considered. Toddlers showed significant improvement in all outcomes in both Levels 1 and 2. These change scores were significantly different among levels only for the Attachment Subscale, where children in Level 1 programs showed significantly greater gains (p<.05) in comparison to Level 2 (see Figure 24).

Figure 24. Pre/post outcomes gains for toddlers and differences in change scores between MA QRIS Level 1 and Level 2



Summary and implications

It is important to note that the existing system of voluntary participation and the dynamic changes that have taken place in the verification processes for the MA QRIS since its beginning introduces measurement error into the models and subsequently makes finding associations between MA QRIS Levels and outcomes challenging. Quality data suggested that some programs at Level 1, Level 2, and Level 3 are functioning at a higher MA QRIS Levels or at lower MA QRIS Levels than currently verified. Despite this, there is evidence that child outcomes are significantly different across children in programs assigned to different MA QRIS Levels. Importantly, results indicated that children in Level 3 showed significantly greater improvement in their PPVT scores over time than did those in Level 2 (p<.05) and significantly greater developmental gains in Attachment Subscale scores of the DECA than did those in Level 1 (p<.05) when controlling for subsidy receipt, special education, ELL, and pre-test scores. It is also important to consider that children at Level 3 appear to be a higher risk population (more likely to receive subsidies) in terms of school readiness and that many other potential demographics related to social economic status were not able to be controlled for in analyses.

Section VII: MA QRIS Program Re-Leveling

To further evaluate the MA QRIS, researchers undertook the process of re-Leveling programs based on data collected for the study. This aspect of the study is important given that a program's assigned MA QRIS Level may not be fully reflective of their quality at the time of the study for a number of reasons, including but not limited to the voluntary nature of the MA QRIS in which some high-quality programs may choose not to pursue higher MA QRIS Levels, self-assessment of classroom quality at Level 2, changes in program staffing and policies that may have occurred since initial verification, and changes and differences in QRIS policies and verification procedures over time.

The study used an overall approach to re-Leveling that attempted to closely approximate the EEC verification process using criteria-level data collected during the validation study, as reported in the MA QRIS Standards, Criteria, and Verification section of this report, as well as the programs' Overall Average Item ERS scores. Some criteria, such as those that were unable to be reliably measured, those with high exemption rates, and those regularly not being met, were excluded to maintain variation. Consistent with EEC policy, which allows programs to apply for and receive exemptions from MA QRIS criteria, with the exception of ERS score requirements, programs were given up to two exemptions from MA QRIS criteria, except at Level 4, where there were relatively few criteria.⁶⁷ In total, the re-Leveling process took into account 59 of the 79 QRIS criteria. Where criteria were measured by the ERS, the Overall Average Item Score was used, and programs were required to meet QRIS cut-offs for preschool and infant and toddler classrooms if both age groups were served. Programs were assigned the highest level for which they met the relevant criteria. Programs that did not meet the criteria for Level 2 were assigned to Level 1 as part of the re-Leveling procedure

Using both the Overall Average Item Score and the other MA QRIS criteria, 34 percent of study programs were assigned to Level 1, 49 percent to Level 2, 15 percent to Level 3, and 2 percent to Level 4 (Table 160).

Table 160. Distribution of programs, by Re-Level using both MA QRIS criteria and ERS scores

	N	%
Level 1	43	34%
Level 2	62	49%
Level 3	19	15%
Level 4	2	2%

Overall, the re-Level brought on program movement but the change was largely focused on a one quality level shift, typically downward. The exception was Level 1, where about a quarter (26%) of the programs shifted up to Level 2 (see Table 161). Among Level 2 programs, about half (55 percent) stayed at Level 2, while one-third (33%) moved down to Level 1 and 13 percent moved up to Level 3. The most movement was seen among programs in the upper tiers. Among Level 3 programs, 73 percent moved down (71 percent moved to Level 2 and one program moved to Level 1), and among Level 4 programs, most of the programs moved to Level 3.

⁶⁷ Although EEC requires programs to apply for exemptions by submitting a plan that demonstrates their ability to meet the criteria in the future, for re-Leveling the validation study applied exemptions to any unmet criteria, up to two per level, except for those assessed through ERS requirements.

Table 161. Shifts in centers' levels as a result of re-Leveling including ERS criteria

QRIS granted Level	Distribution of programs at each Level after Re-Level including ERS							
	Level 1		Level 2		Level 3		Level 4	
	#	%	#	%	#	%	#	%
Level 1 (n=39)	29	74	10	26	0	0	0	0
Level 2 (n=40)	13	33	22	55	5	13	0	0
Level 3 (n=40)	1	2	29	71	10	24	1	2
Level 4 (n=6)	0	0	1	17	4	67	1	17

When re-Leveling resulted in a program receiving a lower level than their granted MA QRIS level, analyses showed that these programs were often not meeting ERS Overall Average Item Score cut-offs or criteria related to educator degree attainment, although these were not necessarily the exclusive reasons for downward movement. Among granted Level 2 programs that did not meet Level 2 requirements after re-Leveling, four criteria stood out as the most common barriers to achieving their granted level (Table 162). Approximately three-fifths (62%) of these programs did not meet the Level 2 requirement that half of classrooms have an educator with a bachelor's degree or higher, which was the largest barrier by far in the selected criteria. Other unmet requirements included: ERS Overall Average Item Score cut-offs, administrator professional development specific to supervising adults, and participation in community events (38 percent each).

Table 162. Most commonly unmet criteria for programs granted Level 2 that did not meet Level 2 after Re-Leveling, including ERS

Criteria	Status	#	%
3B.2.2: 50 percent of classrooms have educator(s) with a bachelor's degree or higher who work for the full program day.	Did not meet	8	62
	Met	5	38
ERS Overall Average Item Score requirements (no exceptions allows)	Did not meet	5	38
	Met	8	62
3A.2.5: Program administrator has received professional development in supervision of adults and strategies for working with adults.	Did not meet	5	38
	Met	8	62
4A.2.5: Program participates in community events.	Did not meet	5	38
	Met	8	62

The most common unmet criteria for programs that dropped from Level 3 after re-Leveling are shown in Table 163. Overall, 23 of the 30 Level 3 programs that were re-Leveled downward did not meet the EEC ERS Total Score cut-off for Level 3. Other commonly unmet criteria included training in special diets, allergies, and feeding issues (27 percent); tracking teacher turnover (27 percent); bachelor degree attainment for educators (23 percent); and ensuring access to developmental, mental health, health, and nutrition services (23 percent).

Table 163. Most commonly unmet criteria for programs granted Level 3 that did not meet Level 3 after Re-Leveling, including ERS

Criteria	Status	#	%
ERS Overall Average Item Score requirements (no exceptions allowed)	Did not meet	23	77
	Met	7	23
2A.3.2: Staff are trained in how to work with children with special diets, allergies, and specialized feeding issues.	Did not meet	8	27
	Met	22	73
5A.3.5: Program tracks and monitors teacher turn-over and has a plan for addressing turn over.	Did not meet	8	27
	Met	22	73
3B.3.1: 75 percent of classrooms have educator(s) with a bachelor’s degree or higher who work for the full program day.	Did not meet	7	23
	Met	23	77
4A.3.5: Program ensures young children and their families have access to developmental, mental health, health, and nutrition services either through private pay arrangements OR are offered such services through other programs.	Did not meet	7	23
	Met	23	77

For Level 4 programs, the reason programs were re-Leveled downward was the ERS Overall Average Item Score requirement. All five programs that dropped after re-Leveling did not meet the ERS criteria (Table 164).

Table 164. Most commonly unmet criteria for programs granted Level 4 that did not meet Level 4 after Re-Leveling, including ERS

Criteria	Status	#	%
ERS Overall Average Item Score requirements (no exceptions allowed)	Did not meet	5	100
	Met	0	0

Re-Leveling and the individual QRIS quality standards

In addition to re-Leveling, overall, programs were also re-Leveled within each of the eight QRIS quality standards, using the measurable criteria specifically within that domain, including ERS Overall Average Item Score requirements when these were used by EEC to measure criteria related to those standards. To fully assess the number of programs that met criteria within each quality standard, **no** exemptions were applied for these analyses. By re-Leveling for each standard, it is possible to examine whether particular standards are more difficult to meet than other standards.

For each of the eight quality standards, Table 165 shows the proportion of programs at each level after re-Leveling; and Table 166 displays re-Leveling shifts in comparison to programs MA QRIS verified Level.

Table 165. Programs' Re-Level within individual quality standards

	Level 1	Level 2	Level 3	Level 4
1A: Curriculum, Assessment, and Diversity	14%	59%	4%	23%
1B: Teacher-Child Relationships and Interactions	15%	59%	22%	4%
2A: Safe, Healthy Indoor and Outdoor Environments	28%	54%	15%	3%
3A: Program Administrator Qualifications and PD	38%	10%	5%	47%
3B: Program Staff Qualifications and PD	44%	11%	9%	37%
4A: Family and Community Engagement	35%	48%	14%	2%
5A: Leadership, Management, and Administration	22%	39%	37%	2%
5B: Supervision	3%	10%	21%	66%

Table 166. Programs' Re-Level within individual quality standards compared with EEC granted level

	Lower	The same	Higher
1A: Curriculum, Assessment, and Diversity	23%	36%	41%
1B: Teacher-Child Relationships and Interactions	24%	49%	27%
2A: Safe, Healthy Indoor and Outdoor Environments	35%	44%	21%
3A: Program Administrator Qualifications and PD	21%	30%	49%
3B: Program Staff Qualifications and PD	22%	38%	40%
4A: Family and Community Engagement	40%	41%	18%
5A: Leadership, Management, and Administration	19%	56%	25%
5B: Supervision	2%	13%	85%

As these analyses show, the distribution of programs by MA QRIS Re-Level varies for individual quality standards. For example, for Standard 1A (Curriculum, Assessment, and Diversity), most programs fell into Level 2 (59 percent) or Level 4 (23 percent) after the re-Level, and for Standard 1B, most programs were performing at Level 2 (59 percent) or Level 3 (22 percent). For 3A (Program Administrator Qualifications and PD) and 3B (Program Staff Qualifications and PD), most programs were either Level 1 or Level 4, with relatively fewer programs at Levels 2 and 3, suggesting that programs are either largely meeting quality standards or not at all. For Standard 5B (Supervision), most programs met quality standards that would allow them to attain an upper tier rating for this particular standard, with 21 percent meeting all measured Level 3 criteria and 66 percent meeting all Level 4 criteria. For two standards — 2A (Safe, Healthy Indoor and Outdoor Environments) and 4A (Family and Community Engagement) — most programs were performing at Level 1 or Level 2, with relatively few programs meeting criteria for the upper tiers. The observed shifts downward in Curriculum, Assessment, and Diversity; Teacher-Child Relationships and Interactions; and Safe, Healthy Indoor and Outdoor Environments, are largely driven by programs not meeting ERS Standards, for reasons discussed above as well as in the ERS section.

Re-Leveling and child outcomes

To further assess the MA QRIS system, the relationship between child outcomes and program MA QRIS Re-Level status, as indicated in Table 168, was explored. First, researchers examined gains of children by MA QRIS Re-Levels, with the exception of the PLBS in which pre-post differences were assessed, as opposed to gains. Models controlled for special education, ELL, and subsidy status, as well as pre-test score (covariates). To account for the nesting of children within classrooms, tests of differences in child gains across MA QRIS Re-Levels were conducted in a multilevel multiple group structural equation modeling framework. These characteristics were centered at the mean of the analysis sample so the gain values presented below (Table 167) represent the covariate-adjusted mean gain scores for each Level (post scores for PLBS).

Table 167. Gains from pre-test to post-test on individual measures by Re-Level of program (Covariate-adjusted gain scores)

Measure	Level 1 (n=134 children)	Level 2 (n=209 children)	Level 3 (n=55 children)
PPVT	1.79*	2.03**	3.84**
Woodcock Johnson Letter Word	1.50	.18	1.04
Woodcock Johnson Applied Problems	2.58**	2.34**	4.60***
DECA Total Score	.73	2.71**	5.95***
Initiative Subscale	1.62*	2.59**	4.96**
Self-Regulation Subscale	.87	1.14	3.87*
Attachment Subscale	-0.66	3.26***	6.67***
Negative Behavior Subscale	-1.86^	.80	-2.53*
PLBS Total Score (co-variate adjusted post score)	51.23***	50.68***	57.21***

^p<.10. *p<.05. ** p<.01.

As can be seen in the chart, significant gains from pre- to post-assessment were noted across all MA QRIS Re-Levels. The PPVT and Woodcock Johnson are age-normed measures with gains based on standard scores adjusted by age of the child within 15 days. As such, significant changes from pre- to post-assessments on standard scores suggest that children are doing better in comparison to their age-comparable peers based on normed expectations. At each level, significant gains on the PPVT, Woodcock Johnson Applied Problems, and the DECA (with the exception of Level 1) and the PLBS are evident. The magnitude of the gain is consistently greater for children in Level 3 programs than in Levels 1 or 2.

Analytic models were used to determine if the apparent differences in gains by MA QRIS Re-Leveled assignment were statistically significant. Models controlled for special education, ELL, and subsidy status, as well as pretest scores (covariates). For the PPVT, Woodcock Johnson Letter Word, Woodcock Johnson Applied Problems, and the DECA measures tests of differences in child gains across Levels were conducted in a multilevel multiple group structural equation modeling framework to account for the nesting of children within classrooms. This multilevel structural equation modeling allowed researchers to simultaneously model individualized child-level covariates (e.g., ELL, subsidies, special education, pre-test scores) in conjunction with the shared influence of classroom-level variables (classroom MA QRIS Re-Level). This modeling strategy allows both the child-level (“Within”) and classroom-level (“Between”) of analysis to be estimated concurrently and the impact of the QRIS Re-Level to be accounted for at both the child and program Level.

For the PBLs, a different analytic approach was used that assumes there is an underlying continuum for the outcome variable that extends beyond the highest and lowest rating point on the measure used in the study – that is, this approach assumes that other children, not assessed, might have scored lower or higher than the children in this sample. Values of the baseline scores, along with the demographic characteristics used as statistical controls are used to estimate of these “out-of-range” scores. This method of modeling change is limited by the choice of predictors in the model and by the assumption that the relationship between a given predictor and the outcome score can be described by a straight line.

Results using the multi-level modeling for the MA QRIS Re-Leveled programs are presented below. Some significant differences were noted.

Table 168. Tests of differences among Levels in magnitude of change scores on individual measures (Wald test coefficient)

Measure	Level 1 vs. Level 2	Level 1 vs. Level 3	Level 2 vs. Level 3
PPVT	.15	1.99	1.84
Woodcock- Johnson Letter Word	-1.14	-.034	.80
Woodcock Johnson Applied Problems	-.45	1.89	2.34
DECA Total Score	1.98	5.22**	3.24^
Initiative Subscale	.96	3.34^	2.38^
Self-Regulation Subscale	.28	3.00^	2.73
Attachment Subscale	3.93**	7.33***	3.41^
Negative Behavior Subscale	2.65**	-.67	-3.33*
PLBS Total Score (co-variate adjusted post score)	.056	5.97***	6.53***

^p<.10. *p<.05. **p<.01. *** p<.001

As can be seen in the Table 168, significant differences in social emotional gains as measured by the DECA also were noted by MA QRIS Re-level, with children in higher level programs exhibiting greater social emotional gains. Specifically, children in Re-Level 1 programs had smaller gains in the Attachment Subscale than did children in Levels 2 and 3; the differences on the Attachment Subscale between children in Re-Level 2 programs, compared to children in Re-Level 3 programs, were marginally significant. In addition, children in Re-Level 1 and Re-Level 2 programs showed significantly smaller gains on the Negative Behavior Subscale (with scores reversed so that a higher score indicated positive developmental growth) compared to children in Re-Level 3 programs. Finally, children in Re-Level 1 had significantly smaller gains, and children in Re-Level 2 had marginally smaller gains, in the total DECA score, compared to children in Level 3. Significant differences were also noted on the PBLs by MA QRIS Re-Level, with significant differences found between Levels 1 and Levels 3, and Levels 2 and Levels 3.

However, no statistically significant differences were noted by re-Level for the PPVT, Woodcock Johnson Letter Word, or the Woodcock Johnson Applied Problems. Data on gains presented above in Table 169 do suggest a trend with children exhibiting greater gains at higher re-levels. The lack of statistical significance between levels reflects the fact that variability is greater within levels (among children within the same level) as opposed to between levels (among children at different levels). In addition, it is important to note that the sample sizes are small (particularly for Re-Level 3, where the largest re-Level differences are observed) and conservative data analytic techniques were employed that considered the clustered nature of children, which reduces power and raises the bar on significance testing.

Summary and implications

To summarize, the re-Level exercise indicated significant shifting in programs. As noted previously, drift in quality appears evident at Level 3 and Level 4 regarding ERS scores. This is most likely due to the changing verification processes and requirements of the MA QRIS system, time lapse from verification point to classroom study observation, and potential regional discrepancies in verification. Additionally, data suggest that the Level 2 ERS self-assessment process may need added supports and technical assistance. Lastly, educational requirements appear to be a barrier for many programs at both Level 2 and Level 3.

When looking at the criteria of individual MA QRIS Quality Standards, the distribution of programs meeting standards the Program Administrator Qualifications and PD; and the Program Staff Qualifications Standards indicated very few programs in the middle levels, suggesting a revision of the criteria that is more progressive in requirements may be needed. In terms of the requirements of the Supervision Standard, most of the programs were at Level 4. This suggests that criteria may need to be better defined to promote higher quality practice (e.g., specifying standards for benefits, supervision, and evaluation; focusing on specific PAS items as opposed to overall scores). The observed shifts downward on the Curriculum, Assessment, and Diversity; the Teacher-Child Relationships and Interactions; and the Safe, Healthy Indoor and Outdoor Environments Standards is largely driven by programs not meeting ERS Standards, for reasons discussed above as well as in the ERS section. It is important to note that researchers were unable to verify CEUs for the training and PD requirements, and therefore, were not able to consider CEU obtainment for training in the re-Level. As such, educators were required to have completed training with no CEU or time frames to PD considered.

Importantly, significant differences in social emotional gains as measured by the DECA also were noted using MA QRIS re-Leveled data, with children in higher re-leveled programs exhibiting greater social emotional gains. Specifically, children in Level 1 programs had smaller gains in the Attachment Subscale than did children in Levels 2 and 3; the differences on the Attachment Subscale between children in Level 2 programs, compared to children in Level 3 programs, were marginally significant. In addition, children in Level 1 and Level 2 programs showed significantly smaller gains on the Negative Behavior Subscale (with scores reversed so that a higher score indicated positive developmental growth) compared to children in Level 3 programs. Finally, children in Level 1 had significantly smaller gains, and children in Level 2 had marginally smaller gains, in the total DECA score, compared to children in Level 3. Significant differences were also noted on the PBLs by MA QRIS re-Levels, with significant differences found between Levels 1 and Levels 3 and Levels 2 and Levels 3.

Section VIII: Perceptions of the MA QRIS among Programs and Providers

Reflecting EEC's ongoing commitment to engaging the early education and care community as it builds and strengthens the Massachusetts QRIS, the MA EEC commissioned a survey of early education and care programs and providers. The purpose of the survey was to gather input and feedback about how the MAQRIS is working for participating centers, schools, after school education and care providers and family child care programs, as well as the ways in which the system and support for the system can be improved.⁶⁸ The confidential survey was administered to randomly selected early education programs from March 23, 2016, to April 15, 2016. It reflects the second such administration of a statewide provider survey, the first taking place in 2013.

In total, for 2016, early education and care professionals from more than 550 programs and family child care homes shared their experiences and feedback through the survey. This included 291 community based centers (a 58% response rate), 81 public school-based providers (a 56% response rate), and 181 family child care providers (a 36% response rate). The random selection of programs and providers, coupled with relatively high response rates, suggests that the views expressed on the surveys are likely to be reflective of those in the field more generally. The administration of the survey in 2013 yielded similar response rates.

Key findings

Programs and providers said that the MA QRIS led to changes in their programs and improvement in their overall quality of care. Approximately two-thirds of respondents from centers, 80% of family child care, and 50% of public schools participating in the MA QRIS reported their programs made changes as a result of being in the system. Where changes were made, respondents expressed that these had made a meaningful difference in their overall quality (95% for centers, 97% for family child care, and 91% of public schools). Among all surveyed programs, 75% of centers, 80% of family child care providers, and 56% of public schools indicated that participation in the MA QRIS has led to an improvement in quality.

Most MA QRIS participants planned to advance to higher levels. About three-quarters of respondents from centers, two-thirds of those from family child care, and half of those from public schools planned to advance in the system, a finding that has remained relatively consistent over time. For centers and family child care, quality improvements were the most commonly identified reason for wanting to progress (54% and 67%, respectively), whereas for public schools, funding eligibility was the primary motivation.

Respondents from all provider types commonly cited the time required to complete the self-assessment as one of their most pressing barriers. Education and training requirements were also commonly cited barriers for centers; family care and public schools had trouble finding the time to put together required documents; and family child care providers commonly referenced the cost of materials, services, or trainings.

Many programs participated in both the MA QRIS and another quality accreditation system. More than half of center respondents (56%) and large proportions of public school and family child care respondents (42% and 38%, respectively) participated in a quality accreditation system in addition to the MA QRIS. Among these programs, the vast majority of centers and public schools (86% and 89%, respectively) and a majority of family child care providers (52%) would prefer to use a single quality or accreditation system. Family child care providers preferred that the MA QRIS be the single system used (56%), while respondents from centers and public schools preferred other accreditation systems (58% of centers and 64% of public schools).

Most providers felt that communication about and support for the MA QRIS had improved in the last two years. A majority of center respondents reported that the resources to help programs understand the MA QRIS had

⁶⁸ The survey was originally funded through EEC's Race to the Top Early Learning Challenge grant which was focused on care for children of birth through five years old; as a result after-school and out-of-school-time providers were not able to be included.

improved (66%), as had communication about that MA QRIS (59%) and support available to improve quality (58%). Family child care providers also had positive feelings about those three areas, with over two-thirds reporting improvements. In 2016, approximately one in three respondents from centers received information through monthly technical assistance webinars, a communication vehicle not available at the time of the 2013 survey. Public school providers tended to feel as though these aspects of QRIS had stayed the same; although large minorities indicated that improvement was evident, ranging from 33% (for communication about QRIS) to 40% (for support available to help programs improve).

Other notable changes in response patterns on the 2016 survey compared with those from 2013 include:

- An increase in the proportion of centers reporting an improvement in quality as a result of participating in QRIS (up from 65% in 2013), with the greatest increase in the percentage of programs indicating that quality had improved very much (25% in 2016 compared with 17% in 2013).
- An increase in the importance of quality improvement as centers' and family child care providers' reasons for joining and/or planning to progress in QRIS, and a concurrent decrease in emphasis on funding eligibility and/or requirements from systems, school districts, or umbrella organizations.
- An increase in the proportion of center-based respondents who felt that MA QRIS ratings were an accurate reflection of quality (46% in 2016 compared with 36% in 2013).
- An increase in the proportion of family care providers who felt they understood the support available to help programs move up (53% said quite a bit or completely in 2016, compared with 42% in 2013), and a decrease in the proportion of center-based respondents who felt they understood the MA QRIS application system (56% said quite a bit or completely in 2016, compared with 67% in 2013).

Administrators from centers in the upper tiers (Levels 3 and 4) appeared to reflect an engaged constituency that felt they understand the system, believed it promotes quality, and planned to progress. All or nearly all administrators from upper tier programs said that they understood the system's goals, standards, process for applying for a level, and the requirements of each level. Nearly all administrators from these centers (95%) believed their quality had improved as a result of the MA QRIS. Among these providers, quality was most often described as somewhat improved, whereas it was typically described as a little improved for those in lower levels. Receipt of a MA QRIS Quality Improvement grant also appeared to be associated with more favorable impressions of quality improvement. Nearly all administrators from Level 3 and 4 centers (92%) planned to work towards the next level, as did 85% of those from Level 2 programs. (Among Level 1 programs, 63% planned to progress.)

Public School administrators tended to be the least engaged and have the least favorable opinions of the system. Compared with centers and family child care providers, respondents from public school–based programs were the least likely to feel they understood aspects of QRIS, and they continued to be the most skeptical of the accuracy of MA QRIS ratings (31% said QRIS was accurate). The proportion of school-based administrators planning to advance decreased from 58% in 2013 to 54% in 2016, and the proportion not planning to advance increased from 4% to 15% (the remaining programs were not sure).

Further, respondents suggested a number of possible strategies to help programs and providers overcome barriers and advance in the system. These include:

- Increase funding opportunities and/or offer higher reimbursement rates for providers that reach higher levels.
- Provide more coaching, consultation, training, or mentoring.
- Simplify tools, and require less paperwork and make it clearer.
- Remove or allow for flexibility in relation to some requirements so programs are not unnecessarily held back by one or two of the system's criteria.
- Provide Spanish-speaking family childcare providers additional support in Spanish.

Summary and implications

Although these results offer some positive indications with regard to the MA QRIS and providers' impressions of the system, they also revealed opportunities for continued growth as reflected in the following recommendations:

- Continue to build field-level understanding of the MA QRIS and its requirements, including continued communication with providers of all types — but particular outreach is needed for programs that are currently less engaged, notably those in Levels 1 and 2 and public schools.
- Expand access to supports to help programs progress to higher levels, such as mentoring, coaching, and training, as well as increased funding for quality improvement.
- Review MA QRIS requirements and criteria to eliminate unnecessary barriers, minimize the administrative burden, and, where possible, allow programs flexibility to demonstrate and meet quality benchmarks.

A number of validation studies have been conducted by other states, particularly those that were early adopters of QRIS.⁶⁹ In 2014, Karoly identified and reviewed 14 studies conducted for 12 states or substate areas, finding substantial variation in methods and findings. According to Karoly, studies commonly addressed one of three major questions, namely, whether higher ratings correspond with higher observed quality, whether ratings improve over time, and whether higher rated programs have better developmental outcomes for children. Most studies investigating the relationship between ratings and observed quality did find evidence of a positive relationship, but Karoly notes that this was not necessarily surprising given that the observed quality measure used, often the ERS, was typically part of the rating system. Results related to the correspondence between rating and developmental outcomes were more mixed, with two of four studies classified by the author as having stronger design showing evidence of such a relationship. Karoly concludes that the “lack of robust findings across these studies indicate that QRISs, as currently configured, do not necessarily capture differences in program quality that are predictive of gains in key developmental domains.”

Since that time, 20 states have received federal RTTT-ELC grants, which require validation studies that investigate the relationship between rating and observed quality using a previously validated measure and also between the ratings and developmental outcomes for children. Studies may also incorporate other evaluation- and validation-related questions to support states’ implementation of their tiered quality rating systems, as necessary. According to the Quality Initiatives Research and Evaluation Consortium (INQUIRE), to date, 3 of the 20 RTT-ELC states have published validation studies, including one state that has published part, but not all of its validation study. These states are Minnesota, Wisconsin, and California.

Minnesota’s validation study relied on multiple sources of evidence, including observations of quality in a sample of over 300 programs participating in the state’s QRIS — Parent Aware — and direct assessments of over 1,000 children in pre-kindergarten. The study incorporated all program types participating in the QRIS, including licensed family child care programs and child care centers, Head Start programs, and school-based pre-kindergarten programs. Main study findings provided positive support for the following: Parent Aware ratings appear to be valid in meaningfully differentiating levels of quality related to children’s development; the Accelerated Pathway to Rating process which offers a rating of Four Stars to programs that meet external quality standards aligned with Parent Aware appears to be an effective way to identify programs that engage in practices supportive of school readiness; children in Parent Aware-rated programs made gains from Fall to Spring of their pre-kindergarten year in measurable areas that are critical for school readiness.⁷⁰

Like Minnesota’s study, the Wisconsin Early Child Care Study (WECCS) included both family and group child care providers participating in the state’s YoungStar program. In total, the study included 157 programs and 887 children with the program sample stratified by low (2-Star) and high (3-Star and above) quality level, as well as region (Northeast or Milwaukee). For this study, data collection instruments included surveys to children’s parents, teachers, and program administrators. In addition, children’s school readiness was assessed in a battery of standardized assessments administered by trained research staff. Lastly, classroom quality was measured using the Environment Rating Scales (ERS). Results from this study showed that, on average, children in the study were meeting expected benchmarks for learning, but no relationship was found between quality level and developmental gains. Children in programs of all levels showed similar levels of school readiness across multiple early academic and behavioral outcomes in the spring of the study year. Children’s school readiness also did not differ when variation in rating points or observational quality measures were used. The results of this study provided empirical evidence that observed quality was found to be higher among 3 Star or higher rated programs than for 2 Star programs.

⁶⁹ Karoly, L. (2014). Validation Studies for Early Learning and Care Quality Rating and Improvement Systems. RAND Education Working Paper. Available at: http://www.rand.org/pubs/working_papers/WR1051.html.

⁷⁰ Tout, K., Cleveland, J., Li, W., Starr, R., Soli, M. & Bultinck, E. (2016). The Parent Aware Evaluation: Initial Validation Report. Minneapolis, MN: Child Trends.

California's validation study differs from that of other states because the state has a county-based system for determining program quality, such that the system differs from county to county. The system is administered locally in 27 counties by 23 lead agencies called the Consortia. Participation in the five-level QRIS is voluntary. Participating sites included center-based programs and family child care homes, and are rated on three main categories: 1) Child Development and School Readiness, 2) Teachers and Teaching, and 3) Program and Environment. This system is a hybrid rating structure with blocks at the early levels and points at later levels.⁷¹ At the time of this report, a final report for the CA QRIS has not been published. However, study findings have been published in a 2015 half-term report and as part of the annual BUILD 2016 conference. Researchers reported "some evidence of validity of the system," in the form of relationships between quality tiers and some aspects of the CLASS and PQA, but "no evidence of higher gains in higher tiers," although they were unable to include the two lowest tiers. Findings showed the quality rating element measured in the system to function differently for centers and family child care homes.⁷²

⁷¹ Hawkinson, L., Quick, H., Muenchow, S., Anthony, J., Weinberg, E., Holod, A., Parrish, D., Meakin, J., Lee, D., Tarrant, K., Cannon, J., Zellman, G., Karoly, L., & Faxon-Mills, S. (2015). Independent Evaluation of California's Race to the Top-Early Learning Challenge Quality Rating and Improvement System: Half-Term Report. San Mateo, CA: American Institutes for Research.

⁷² Hawkinson, L., & Quick, H. (2016). Independent Evaluation of California's RTT-ELC QRIS: Validation Study Results (Copyright © 2014 American Institutes for Research). Retrieved from <http://www.qrisnetwork.org/sites/all/files/conference-session/resources/222ValidationStudiesPPT1.pdf>

In general, the results from the MA Validation study suggest that the current MA QRIS system is functioning largely as intended. Analyses indicated that the MA QRIS Level is positively and significantly associated with higher levels of observed classroom quality. Additionally, a series of ANOVAs found significant differences by MA QRIS Level in the proportion of programs meeting individual criterion for the majority of MA QRIS criteria. MA QRIS Level was also associated with greater gains in some child outcomes.

A detailed review of individual MA QRIS criterion and verification requirements indicated that adjustments to the current MA QRIS system could strengthen the effectiveness of the MA QRIS in promoting and maintaining program quality and positive child outcomes. Recommendations for modifications to the system are based on a system-wide approach and incorporate the criteria and the verification processes, as well as consider the interdependence of each level in the enhancement of practice and movement of programs within the MA QRIS.

Recommendations for system-wide levers and supports

Analyses indicated some additional overarching supports and policies are needed to impact the efficacy and functioning of the MA QRIS.

Strengthen licensing and the foundational supports to quality. Although differences in observed classroom quality were noted among programs by MA QRIS Level, the results suggested that many programs face challenges related to health practices and safety regardless of programs' verified MA QRIS Level. Item and indicator analyses of the ERS highlighted an overlap between ERS items related to health and safety practices and the MA Licensing Standards that programs were not consistently meeting. As such, researchers recommend additional support for licensing (e.g., provider training on licensing standards and requirements,; licensur training and reliability; revision of licensing assessment forms to facilitate licensing visits, clarification of licensing standards, and licensing case load reviews) to ensure that basic standards for health and safety are in place, as well as to support the MA QRIS as a foundational level and complementary system for practice and the promotion of quality.

Differential support by program type. Sample characteristics indicated: larger programs that were part of systems or networks tended to be at the upper Levels of the MA QRIS. This suggests that smaller programs may need added supports such as: fiscal incentives and PD supports; a size-specific approach to technical assistance such as access to mentor programs; and curriculum and assessment support grants to facilitate the implementation of a curriculum and formal assessment protocols.

Leverage other verification systems and consider diverse entry points. In order to support the efficiency and sustainability of the MA QRIS, researchers suggest that MA EEC leverage other verification systems to streamline the verification of programs within the MA QRIS. Observed classroom quality data from the study support the entry of Head Start programs (in good standing) and NAEYC-accredited programs at Level 2 of the MA QRIS. Consistent with this, criteria for Level 2 should be largely aligned with the requirements of Head Start and NAEYC.

Remove exemptions and add flexibility and extensions to the system. Data suggests that many programs are typically taking exemptions on some criteria within all MA QRIS Levels. These exemptions appear to have limited requirements for eventual compliance. Researchers believe that this adds to the complexity of the administration of the system as well as, inconsistencies in its implementation. The research team recommends that MA EEC consider removing/modifying criteria that appear to be barriers to progress (high exemption rates); provide more flexibility for criteria requirements (e.g., expand definition of MA QRIS PD requirements beyond CEU only) ;and offer extensions on meeting certain criteria with eventual completion dates established, as opposed to offering blanket exemptions on individual criteria at each MA QRIS Level.

Recommendations for revisions to system related to classroom quality

The research team recommends several changes to the system to specifically promote and maintain the classroom quality of program in the MA QRIS.

Strengthen Self-Assessment of Quality at Level 2. The results indicated that the self-assessment process at Level 2 needs to be strengthened. The validity of the self-assessment at Level 2 influences the amount of support that programs need in order to meet Level 3 and the amount of lift required in terms of changes to programming and practice. Level 2 was conceptualized as an improvement over Level 1 in practices that support quality; and a demonstration of programs' willingness to engage with the MA QRIS. However, Level 2 self-assessments do not currently prepare programs for the investment needed to achieve Level 3, potentially resulting in frustration or perceived barriers. Additional supports that improve the functioning of Level 2 are critical. Specifically, researchers recommend required ERS training to facilitate self-assessment and obligatory training on the MA Early Learning Standards to underscore expectations for practice, either at entrance into the system or as part of Level 2 criteria. Additionally, MA EEC should consider supporting Level 2 programs through compulsory mentoring or technical assistance as a means of promoting self-assessment and developmentally appropriate practices. This is in-line with educators' preference for coaching and TA supports to facilitate progress within the MA QRIS, as indicated in the *Field Experiences with and Impressions of QRIS: Findings and Recommendations from the 2016 Provider Survey*.

Implement re-verification guidelines at Levels 3 and 4. Currently the MA QRIS has no active protocol for re-verification once a program has been granted a specific MA QRIS Level. Analyses found a negative association between the time frame from when a program was verified and the study observation; with programs verified at Levels 3 and 4 more than 2 years prior to the study observation more likely to have observed quality below the MA QRIS observed quality standards. As such, it is recommended that MA EEC implement a process of re-verification at Levels 3 and Levels 4 to ensure quality is maintained and facilitate the movement of programs to the next level of the MA QRIS.

Recommendations for system revisions related to criteria and verification

Currently, the MA QRIS system has 89 individual criteria. Analyses indicated significant differences in the proportion of programs that met criteria at different MA QRIS Levels for the vast majority of criteria. For criteria showing no proportional difference by MA QRIS Level, researchers believe that they should be dropped from the system. Additionally, researchers recommend modification to the remaining criteria to clarify and strengthen requirements. A detail of recommendations for each criterion is provided in the Descriptive Analyses of MA QRIS Standards, Criteria, and Verification Section of the Report. General themes for these modifications are presented below.

Increase consistency between standards and verification. Researchers noted a lack of consistency among standards and verification practices. Notably, many criteria are currently written as evaluative statements or beliefs, which are difficult to verify. Moreover, multiple criteria identify specific practices but call for more generalized verification. As such, researchers recommend more consistent alignment of the language used for the criteria and the verification processes.

Greater distinction among verification for criteria. Currently, many different criteria have the same verification requirements, which, therefore, limit the ability of individual criteria to make a unique contribution to the promotion of quality programming and practice. Researchers believe that each distinct criterion needs distinct verification requirements that are *specific* to each criterion.

Reduce MOU requirements. Although the data indicated much collaboration between programs and other direct service organizations, most programs reported not having formal MOUs or collaborative agreements in place. In interviews, administrators often noted challenges in getting partner agencies to commit to signing formal agreements. This was particularly true regarding establishing MOUs with Early Intervention, the public schools,

or large social service agencies for services. Therefore, this appears to be a barrier within the ECE field at large. Researchers believe that criteria requiring MOUs should be relaxed to proof of collaboration versus a formal MOU or collaborative agreement.

Limitation of the use of overall scales scores for verification. Additionally, it is recommended that the MA QRIS limit the use of overall scale scores for verification and increase the focus on relevant subscales or items within a scale, particularly for the use of the ERS, BAS and PAS. Overall scale scores take into account multiple aspects of quality to produce an aggregate score. Thus, it is possible for programs to meet verification standards but not have in place the specific practices or policies referenced in the criteria being verified, even if the tool addresses the practice at some level. Focusing on specific items or subscales allows for an increased focus on specific practices of quality and ensures that they are being met and implemented.

Incorporate Continuous Quality Improvement Plans (CQIPs) as self-assessment criteria for quality standards. Currently CQIPs are used as a part of the verification process for many criteria of the MA QRIS. The CQIP is a required document generated by a program that describes the needs and plans for future quality improvements. This document is submitted by programs at each level and verified by the PQS. Therefore, the CQIP functions like a criterion not a *verification process* of the MA QRIS. As a result, researchers believe that the CQIP should be integrated into the system as a *criterion* within each MA QRIS Quality Standard, as opposed to simply being incorporated as a piece of the verification process. This improves the internal consistency within the system regarding the functioning of criteria versus a verification process, underscores the importance of creating a CQIP as a dynamic document for improving quality, and emphasizes the need for programs to focus on each individual standard for quality.

Recommendations for system revisions to support the workforce

The MA QRIS system incorporates many PD requirements and educator supports to build practice and promote quality. The results of the study suggest that some modifications to criteria related to the workforce are needed to strengthen requirements and refine the system.

Reduce compound QRIS Professional Development criteria. Many of the MA QRIS PD requirements are written as compound criteria, in which multiple PD requirements are listed as part of one criterion. Analyses found that, in the case of multiple-PD criteria, programs typically met some but not all of the requirements. In fact, as the lists of requirements got longer in a specific criterion, fewer programs were able to meet that criterion; requirements in the middle and end of the list typically had the lowest percentages of compliance. There also appears to be some redundancy of requirements across compound criteria, further underscoring the need to streamline. As such, it is recommended that the compound requirements be split into individual criteria as a means of clarifying expectations, increasing compliance and reducing redundancy.

Establish timeframes for professional development. Data indicated large ranges in timeframes regarding the completion of required PD reported by educators. Currently, the MA QRIS does not place time parameters on PD, and consequently, trainings taken at any time are considered to meet MA QRIS criteria. To reinforce best practices, researchers recommend that EEC establish timeframes for some key MA QRIS PD criteria to ensure that educators are current with the curriculum being implemented, assessment protocols being used, and the Massachusetts Early Learning Guidelines.

Clarify continuing education unit (CEU) requirements and educate the field. The MA QRIS system requires that teachers receive continuing education units (CEUs) or college credit for PD and training. Despite multiple efforts, reliable data regarding CEUs could not be obtained for the study, as most teachers and administrators were unable to identify whether CEUs for professional development were received. Findings suggested a lack of understanding of CEUs and CEU requirements in the field. Researchers recommend that the MA QRIS build awareness of the CEU requirements as an important aspect of the system, starting with explicitly referencing the CEU or credit requirements as part of individual MA QRIS criteria to reduce ambiguity and emphasize the need for CEU-approved PD. Based on feedback from the field and the provider survey, the research team also suggests the EEC consider expanding their PD requirements to include some formal PD opportunities that are not CEU-bearing trainings, such as PD received directly from a curriculum or assessment developer.

Greater consistency with licensing and credentialing. When considering PD and educational requirements, researchers recommend greater consistency between the MA QRIS and Massachusetts licensing and credentialing. Of particular note are MA QRIS requirements for administrators and the MA director credentialing that are similar but do not directly overlap. Researchers recommend modifying the MA QRIS system to incorporate the licensing credentials for Director I and Director II among the MA QRIS Levels to allow for administrators to have only one set of PD goals and promote the administrator career ladder.

Organize all PD requirements into Professional Development Quality Standard. Currently the MA QRIS has PD requirements distributed across multiple MA QRIS Quality Standards, even though the system includes a separate quality standard entitled: *Program Staff Qualifications and Professional Development Quality Standard*. As such, the MA QRIS PD requirements can be difficult to track and understand. Researchers recommend that all PD requirements be consolidated into the *Program Staff Qualifications and Professional Development Quality Standard* in order to streamline the system and facilitate the tracking and understanding of requirements, as well as the development of the CQIPs and Individual Professional Development Plans for educators.

Revise the MA QRIS Supervision Quality Standard to support educator retention and address staff turnover at all QRIS levels. Data indicated that staff turnover was high across all levels of the MA QRIS. Re-leveled data from the *MA QRIS Supervision Quality Standard*, which includes program supports for educators, suggested that most programs were at Level 2 or 4 with few programs at Level 1 and Level 3. As such, revised criteria are recommended to create more distinction between MA QRIS Levels and requirements, as well as raise the bar in terms of supervision and program supports for educators.

Improve access to break time for educators. Although the Parents and Staff Subscale of the ERS was an area of strength among programs, the Personal Needs of Staff item was an outlier with most programs offering limited break time. Analyses indicated that, for the vast majority of programs, educators received only one 30 minute break period (the minimum amount required). Given the high demands of the job and the limited ability of educators to step away from the classroom due to supervision and ratio requirements, the researchers suggest building break time into the MA QRIS as an added support for educators.

Recommendation for system revisions related to the upper tiers of QRIS

As more programs achieve Level 3 status, the research team believes that MA EEC should consider implementing new standards for Level 4 and 5 to continue to promote quality beyond the aspects of quality established through Level 3 in order to continue to promote child outcomes.

Consider adapting a hybrid system to build upper levels and promote outcomes. Nationally, QRIS systems use one of three structures: a block system, a points system, or a hybrid model. For the block system, providers must meet all standards at one level before moving to next higher level (this reflects the current system). For a point system, points are earned for each standard and added together. For a hybrid system, elements of the block and point systems are combined. Research on other QRIS systems indicates that point systems can inflate quality levels, without necessarily addressing classroom practice and quality. The research team recommends a hybrid model be developed, using the current block system up to Level 3 to ensure the foundational levels of quality. After Level 3, researchers recommend that the MA QRIS utilize a point system that focuses less on foundational elements and more on specific educational practices, by adding observed quality measures that address specific practices (e.g., the CLASS or ELLCO) and more choice for programs in terms professional development. This strategy allows programs to build on the foundational elements of quality established up to Level 3, through the continual the promotion of quality practices and targeted educator supports related to classroom curriculum, interactions and program pedagogy.