

OHIO'S SUTQ

Validation Study Results • February 2017



COMPASS
EVALUATION AND RESEARCH, INC.

SUBMITTED BY:

Compass Evaluation and Research, Inc.

5720 Fayetteville Rd., Suite 202

Durham, NC 27713

Toll Free: 877.652.0189

Phone: 919.544.9004

SUTQ Validation Study Results

December 2016

Submitted to:

Melinda Kowalski, Project Officer
Ohio's Department of Job and Family Services
Columbus Ohio

Submitted by:

Compass Evaluation and Research, in partnership with Westat, Inc.

Sarah Heinemeier and Anne D'Agostino
Compass Evaluation and Research

Jennifer Hamilton, Kwang Kim, and Marianne Winglee
Westat, Inc.

With contributions from Susan Yonk (Compass) and Valerie Orellana and Kim Standing (Westat)

Executive Summary

The study of Ohio's Step Up To Quality (SUTQ) Quality Rating and Improvement System (QRIS) examined the validity of SUTQ from five perspectives: face, content, construct, procedural, and predictive validity. For the current study, these perspectives were defined as:

- Face validity: the extent to which SUTQ criteria capture ideas and practices that are meaningful to practitioners and stakeholders in early education and child care.
- Content validity: the extent to which SUTQ criteria are comprehensive and inclusive of evidence-based and accepted ideas and practices.
- Construct validity: the extent to which SUTQ ratings are consistent with measures of high quality early education environments.
- Procedural validity: the extent to which SUTQ uses a reliable process to generate site star rating.
- Predictive validity: the extent to which the SUTQ rating aligns with child developmental progress or outcomes.

The study incorporated a concurrent mixed-methods design, which is a combination of both quantitative and qualitative data collections and analytic techniques. Key aspects of the study's methodology included:

- Use of a stratified random sample of early education and learning sites, some of which are participating and some of which are not participating in SUTQ. The sample included private child care centers, Type A and B homes, and elementary schools with star-rated Early Childhood Education classrooms.
- Completion of a literature review of QRIS validation studies and research related to SUTQ's domains and components.
- Completion of a suite of classroom and site-level assessments at each sampled site, including assessments of:
 - Quality of the early education environment;
 - Teacher-child interactions;
 - Early language and literacy practices;
 - Family and provider (or teacher) relationship quality; and
 - At selected sites, the quality of administrative practices.
- Assessment of language and literacy skills of 3- and 4-year old children at sampled centers and Early Childhood Education classrooms.
- Analysis of Kindergarten Readiness Assessment scores from across the state.

A total of 81 sites consented to participate in the study, representing more than 190 classrooms. The sample was distributed over both urban and non-urban settings and across star ratings, including sites that do not participate in SUTQ (and are considered non-rated, as opposed to sites with a rating of

“0”). In addition, the site extracted data from Ohio’s Departments of Job and Family Services and Education for more than 100,000 students with Kindergarten Readiness Assessment (KRA) scores in either the 2014-2015 or 2015-2016 school years.

Face Validity

Face validity was assessed through the use of a Site Questionnaire, Child and Family Questionnaire, and state stakeholder interviews. Findings suggest that high proportions of participating directors agreed that at least 50% of items on the questionnaire were indicative of quality; the questionnaire items are analogous to criteria by SUTQ to advance in star rating. As the benchmark for agreement was raised to 75% and then 90%, a smaller proportion of directors was found to be in agreement regarding the criteria that represent quality.

The questionnaire areas that received the most agreement included Learning and Development (which incorporated items related to Curriculum and Planning, Child Assessment and Screening, and Interactions and Environment), Administrative and Leadership Practices (particularly items related to Program Administration and Staff Management), Family and Community Partnerships (which incorporated items related to Transitions and Communication and Engagement), and Group Size. The areas that received the least agreement, overall, included Staff Qualifications, Professional Development, and Accreditation.

Parent responses to the Child and Family Questionnaire tended to align with director responses to the Site Questionnaire. Parents also were given a list of quality indicators and reported that the learning activities used with children, provider warmth and nurturing of children, continuous improvement in site quality, and communication with parents all were indicative of site quality. The item with the least amount of support from parents was teacher education (i.e., teachers needed a college education to provide high quality care and education).

Content Validity

The study team conducted a literature review along with a review of QRIS in 44 other states. The results of these reviews indicated:

- Research-based support for the constructs and domains in SUTQ. This stated, there are items that merit further investigation or discussion, such as the relation of teacher education to child outcomes or the ability to tie quality ratings to child outcomes. For the most part, however, SUTQ criteria and standards are aligned with best practices endorsed by national accrediting and professional organizations, supported by research, and used by other states.
- Criteria related to Staff Qualifications are found in all state QRIS. Also common are indicators related to the use of developmentally appropriate curricula, the quality of environment and interactions, and the quality of family partnerships. Least common are indicators related to quality for children with special learning or developmental needs, cultural or linguistic diversity, and community involvement.

The literature review underscores the large amount of existing research, the large volume of emerging research, and the ever-evolving understanding of how to improve and sustain quality in

support of positive outcomes for children. Thus, it is important to consider the degree to which SUTQ can be a flexible system, adapting to new knowledge and standards as their merits are supported.

Construct Validity

With regards to construct validity, the study team wanted to confirm that star rating conveyed meaningful information about a site's quality. In other words, does a 5-star rated site provide a higher level of quality than a 1- or 2-star rated site? To explore this question, the study team conducted independent assessments of classroom and site quality, using a suite of instruments that included the Environment Rating Scales (ERS), the Classroom Assessment Scoring System (CLASS), the Early Language and Literacy Classroom Observation (ELLCO), the Family and Provider/Teacher Relationship Quality Scale (FPTRQ), and the Program Administration Scale (PAS). The study team also collected center education profiles to explore the relationship of teacher education to star rating. The study assessed quality in Type A and B homes with instruments that were analogous to those used with centers and Early Childhood Education classrooms: the Family Child Care Environment Rating Scale-Revised (FCCERS-R), the Caregiver Interaction Scale (CIS), the Child Home Early Language and Literacy Observation (CHELLO), the FPTRQ, and the Business Administration Scale (BAS).

There were several findings of note. First, higher rated centers/Early Childhood Education classrooms tended to earn higher scores on observations of classroom quality, but were not necessarily the highest rated by teachers or parents with regard to family relationships. Thus, although not statistically significant, there does appear to be a difference in quality among centers/Early Childhood Education classrooms with different star ratings, wherein centers/Early Childhood Education classrooms with higher star rating, on average, provide a higher quality of classroom practices (as assessed using the ERS, CLASS, and ELLCO). Second, there appears to be a benefit to any level of participation in SUTQ. There was a difference of 1 point or more on many of the observations of classroom quality, between non-rated centers and centers with any level of star rating. In short, participating in SUTQ and receiving a star rating appears to be associated with higher quality classroom practices, compared to sites that are not participating in SUTQ. Third, areas in which teachers struggled in classroom observations (e.g., Instructional Support) are consistent with areas of difficulty documented in other studies. These are areas in which teachers may benefit from additional support and assistance. Fourth, there does not appear to be a trend for higher rated sites to have better family partnerships or strategies for engaging and working with parents. This is another area in which sites might benefit from additional support and assistance. Fifth, there does not appear to be a consistent relationship between teacher education and observed quality. Sixth, higher rated sites tend to invest in better administrative and management practices. Seventh, Type A and B homes perform on par with their center and Early Childhood Education classroom counterparts, an encouraging finding.

Procedural Validity

The study team created ratings for participating centers using individual assessments and composite scores, all based on observed levels of quality generated by the ERS, CLASS, and ELLCO. While the choice of methodology influenced the findings (e.g., a shift in cut scores may affect levels of agreement, etc.), there was consistently strong agreement between the study team's rating of 3- and 4-star sites (based on observed levels of quality) and, to a lesser extent, 2-star rated sites. There was

weaker agreement between study-generated rating and SUTQ star rating for 1-star and 5-star sites, with the suggestion that quality in 1-star sites is relatively underestimated and quality in 5-star sites is relatively overestimated. Finally, study-generated ratings for non-rated sites suggest that these sites would earn ratings of low to moderate quality, based on observations of classrooms practices.

Predictive Validity

The study team extracted child-level data (that was securely communicated and from which identifying variables were removed) from the Departments of Job and Family Services and Education. Datasets were merged using a unique student identifier; the merged dataset was used to examine the relationship between Early Childhood Education experiences and KRA scores.

For the overall datasets, there were significant differences in mean scores associated with socioeconomic status, race, and disability status. As regards participation in Early Childhood Education programming, children who participated in early learning and development programs sponsored by Education displayed mean scores second to those of students who were not considered economically disadvantaged. Further, children who participated in publicly funded child care in their pre-kindergarten year had higher scores, on average, than students with disabilities or students that were considered economically disadvantaged (but for whom pre-kindergarten data was not available). There were not, however, significant differences among mean scores associated with star rating. There appears to be a benefit to participation in SUTQ, as well as participation in higher rated programs (3-star or higher). However, there aren't distinct patterns in student mean scores that align with each of the five rating tiers. Further, weeks of attendance in publicly funded child care was found to be a significant predictor of student mean score, which should be taken into account when considering the influence of star rating upon child outcomes.

Conclusion

In conclusion, the study team was able to validate many aspects of SUTQ, including its face validity, content validity, and aspects of its construct, procedural, and predictive validity. It is important to keep the study's limitations in mind when interpreting the study findings. Notably, the study had a relatively short window for data collection, which affected the total sample size. In addition, the study was quasi-experimental in design, as the study team could not implement random assignment to "treatment" or "control" groups. Thus, the study team cannot generate conclusions about the causal relationship of SUTQ star rating and observed quality or SUTQ star rating and child outcomes. Finally, the study was conducted less than three full years after the transition from a 3-star to 5-star approach. Thus, from the study team's perspective, SUTQ is progressing towards maturity. The study team was encouraged by the shorter-term outcomes observed in participating sites (namely, higher observed quality in higher rated sites) and encourages Ohio to continue to track site progress and, eventually, child outcomes as the system comes to scale and becomes established practice.

Table of Contents

Study Purpose and Questions	1
Overview.....	1
SUTQ.....	1
The SUTQ Rating Process.....	2
Overview of QRIS Validation Studies.....	3
Overview of the Report	4
Methodology	5
General Methodology.....	5
Description of the Samples.....	5
Institutional Review Approval.....	7
Instrumentation.....	8
Data Collector Training and Quality Control.....	10
Data Review and Verification	10
Face Validity	11
Site Questionnaire	11
Intra-Construct Agreement.....	12
STAKEHOLDER INTERVIEWS AND CHILD AND FAMILY QUESTIONNAIRES.....	17
Content Validity	19
Highlights	19
Learning and Development	19
Curriculum and Planning.....	19
Child Screening and Assessment	20
Interactions and Environment.....	21
Summary.....	22
Administrative and Leadership Practices	22
Staff Supports	22
Program Administration (PAS).....	23
Staff Management	24
Summary.....	24
Staff Education and Professional Development.....	25
Staff Education.....	25
Professional Development.....	26
Summary	26
Family and Community Partnerships	27
Transitions	27
Community and Engagement	28
Summary.....	28
Group Size and Accreditation	29
Group Size and Ratios.....	29
Accreditation.....	29

Construct Validity	31
HIGHLIGHTS	31
Highlights	31
Findings for Centers and Early Childhood Education Classrooms	31
Environment Rating Scales (ERS)	32
Classroom Assessment Scoring System (CLASS)	35
Early language and literacy classroom observation (ELLCO)	39
Family Provider/Teacher Relationship Quality Scale (FPTRQ)	40
Importance of Education	43
Administrative Practices	46
Findings for Type A and B Homes.....	48
Family Child Care Environment Rating Scale-Revised (FCCERS-R).....	49
Caregiver Interaction Scale (CIS).....	49
Child Home Early Language and Literacy Observation (CHELLO)	50
Family Provider/Teacher Relationship Quality Scale-Director Measure	51
Administration Practices.....	51
Procedural Validity	54
Highlights	54
.....	54
Environment Rating Scales (ERS)	55
Classroom Assessment Scoring System (CLASS)	55
Early Language and Literacy Classroom Observation (ELLCO).....	56
Site Composite Ratings	57
Classroom Composite Scores.....	59
Predictive Validity.....	65
Highlights	65
Limitations.....	72
Discussion.....	73
Conclusions	77
Appendices.....	79
Appendix A: References.....	79
Appendix B: SUTQ Building Blocks—Centers.....	87
Appendix C: SUTQ Building Blocks—Homes.....	91
Appendix D: Methodology.....	95
Consent documents	95
Sampling Approach	101
Data Extraction	102
Site and participant recruitment efforts.....	102
Data Collector Training	103
Instrument Statistics	103
Data Compilation and Examination	105
Data Analysis.....	106

Appendix E: Site Questionnaire 107
Appendix F: Additional Findings 114
Appendix G 137
 Ohio’s Early Learning and Development Standards in All Essential Domains of School Readiness
 (Birth – Age 5) 137
 Introduction 137
 Organization of the Standards 138
 An Overview of the Domains 138
 References 141
Appendix H 142

Tables

Table 1. Response rate and co-operation rate6
 Table 2. Number of participating sites by program type and SUTQ rating.....6
 Table 3. Percent of respondents who agreed that questionnaire constructs and components are
 indicative of quality, by program type12
 Table 4. Percent of respondents who agreed that Learning and Development practices are indicative of
 quality, by program type.....13
 Table 5. Percent of respondents who agreed that Administrative and Leadership Practices are indicative
 of quality, by program type13
 Table 6. Percent of respondents who agreed that Staff Qualifications and Professional Development are
 indicative of quality, by program type.....14
 Table 7. Percent of respondents who agreed that Family and Community Partnerships are indicative of
 quality, by program type.....14
 Table 8. Percent of respondents who agreed that Group Size and Accreditation are indicative of quality,
 by program type.....15
 Table 9. Percent of respondents who agreed that questionnaire items are indicative of quality, by star
 rating.....15
 Table 10. Percent of respondents who agreed that questionnaire constructs are indicative of quality, by
 star rating.....16
 Table 11. Agreement between center star rating and ITERS-R scores33
 Table 12. Agreement between center star rating and ECERS—3 scores.....34
 Table 13. Percent of classrooms with low, moderate, and high ERS scores34
 Table 14. Agreement between center star rating and CLASS-Infant scores.....35
 Table 15. Percent of classrooms with low, moderate, and high CLASS-Infant scores36
 Table 16. Agreement between center star rating and CLASS-Toddler scores.....37
 Table 17. Percent of classrooms with low, moderate, and high CLASS Toddler scores37
 Table 18. Agreement between center star rating and CLASS PreK scores38
 Table 19. Percent of classrooms with low, moderate, and high CLASS PreK scores38
 Table 20. Agreement between center star rating and ELLCO scores39

SUTQ Validation Study Results

Table 21. Percent of classrooms with low, moderate, and high ELLCO scores	39
Table 22. Agreement between lower and higher star rating and Director FPTRQ scores.....	41
Table 23. Agreement between star rating and Teacher FPTRQ scores	42
Table 24. Agreement between star rating and Parent FPTRQ scores	43
Table 25. Education, disaggregated by star rating and SUTQ participation	44
Table 26. Alignment of teacher education with low, moderate, and high ERS Scores.....	45
Table 27. Alignment of teacher education with low, moderate, and high CLASS PreK scores.....	45
Table 28. Alignment of teacher education with low, moderate, and high ELLCO Scores	45
Table 29. Agreement between star rating and administrative practices	46
Table 30. Participation of Type A and B homes	48
Table 31. Agreement between star rating and FCCERS scores.....	49
Table 32. Agreement between star rating and CIS scores.....	50
Table 33. Agreement between star rating and CHELLO scores	51
Table 34. Agreement between lower and higher star rating and Director FPTRQ scores.....	51
Table 35. Agreement between star rating and administrative practices	52
Table 36. Agreement between SUTQ star rating and site mean ERS Score	55
Table 37. Agreement between SUTQ star rating and site mean CLASS score.....	56
Table 38. Agreement between SUTQ star rating and site mean ELLCO Score	56
Table 39. Study-generated ratings for sampled sites that do not participate in SUTQ.....	57
Table 40. Agreement between ERS and CLASS composite rating and SUTQ star rating	58
Table 41. Agreement between ERS, CLASS, and ELLCO composite rating and SUTQ star rating	58
Table 42. ERS and CLASS composite ratings for sampled sites that do not participate in SUTQ	59
Table 43. ERS, CLASS, and ELLCO composite ratings for sampled sites that do not participate in SUTQ	59
Table 44. Agreement between ERS and CLASS composite rating and classroom rating.....	60
Table 45. Agreement between ERS, CLASS, and ELLCO composite rating and classroom rating.....	60
Table 46. ERS and CLASS composite ratings for classrooms in sampled sites that do not participate in SUTQ	61
Table 47. ERS, CLASS, and ELLCO composite ratings for classrooms in sampled sites that do not participate in SUTQ	61
Table 48. Infrastructure ranked ratings, by respondent type.....	63
Table 49. Director/Owner feedback on helpful or necessary infrastructure	64
Table 50. KRA mean scores for ECE, PSE, and publicly funded students, disaggregated by lower versus higher star rating: 2014-2015 and 2015-2016 KRA data	67
Table 51. 2014-2015 KRA mean scores for publicly funded children, disaggregated by star rating	68
Table 52. 2015-2016 KRA mean scores for ECE, PSE, and publicly funded children, disaggregated by participation in star rated care	69

Appendix D Tables

Table D. 1. Population used for sampling	101
Table D. 2. Internal Consistency of the Site Questionnaire	104

Table D. 3. Cronbach’s alpha of the provider/teacher measure overall..... 105
 Table D. 4. Cronbach’s alpha of the parent measure overall 105

Appendix F Tables

Table F. 1. Mean ECERS--3 score by type..... 114
 Table F. 2. Mean ECERS-3 score by urban versus non-urban setting 115
 Table F. 3. Mean ITERS-R score..... 116
 Table F. 4. Mean ITERS-R score by urban versus non-urban setting 116
 Table F. 5. Mean CLASS Infant score..... 118
 Table F. 6. Mean CLASS Infant score by urban versus non-urban setting 118
 Table F. 7. Mean CLASS-Toddler score..... 119
 Table F. 8. Mean CLASS-Toddler score by urban versus non-urban setting 119
 Table F. 9. Mean CLASS-PreK score..... 120
 Table F. 10. Mean CLASS-PreK score by urban versus non-urban setting 120
 Table F. 11. Mean ELLCO score by type 121
 Table F. 12. Mean ELLCO score by urban versus non-urban setting 121
 Table F. 13. Mean FPTRQ (Director Measure) scores by type 122
 Table F. 14. Mean FPTRQ (Director Measure) scores by urban versus non-urban setting 122
 Table F. 15. Mean FPTRQ (Teacher Measure) scores by type 123
 Table F. 16. Mean FPTRQ (Teacher Measure) scores by urban versus non-urban setting..... 123
 Table F. 17. Mean FPTRQ (Parent Measure) scores by type..... 124
 Table F. 18. Mean FPTRQ (Parent Measure) scores by urban versus non-urban setting 124
 Table F. 19. Mean FCCERS-R score by type..... 125
 Table F. 20. Mean FCCERS-R score by urban versus non-urban setting 125
 Table F. 21. Mean CIS score by type 126
 Table F. 22. Mean CIS score by urban versus non-urban setting..... 126
 Table F. 23. Mean CHELLO score by type..... 127
 Table F. 24. Mean CHELLO score by urban versus non-urban setting 127
 Table F. 25. Mean PAS score by type 128
 Table F. 26. Mean PAS score by urban versus non-urban setting 129
 Table F. 27. Mean BAS score by type 130
 Table F. 28. Mean BAS score by urban versus non-urban setting 131
 Table F. 29. Overall teacher education by type 132
 Table F. 30. Teacher education by urban versus non-urban setting 132
 Table F. 31. KRA mean scores, disaggregated by pre-kindergarten experience: 2014-2015 133
 Table F. 32. KRA mean scores, disaggregated by pre-kindergarten experience: 2015-2016 134
 Table F. 33. Differences in KRA scores, by participation in ECE, PSE, or publicly funded child care 134
 Table F. 34. KRA mean scores for publicly funded children, disaggregated by race/ethnicity: 2014-2015
 135
 Table F. 35. KRA mean scores for publicly funded children, disaggregated by race/ethnicity: 2015-2016
 135

Table F. 36. Differences in KRA scores, among publicly funded children, by race136
Table F. 37. KRA mean scores for publicly funded children, disaggregated by disability status136

Exhibits

Exhibit 1. Distribution of participating sites across the state 7
Exhibit 2. Parent survey items related to indicators of quality 17
Exhibit 3. Participant Consent and Permission Form.....95
Exhibit 4. Agreement to Participate with Informed Consent and Permission.....98
Exhibit 5. Parent and Child Consent and Permission Form.....99
Exhibit 6. Parent and Child Consent and Permission Form.....100

Study Purpose and Questions

Overview

A Quality Rating and Improvement System (QRIS) is a framework for a standards-based early childhood education system that provides “a systemic approach to assess, improve, and communicate the quality of early care and education programs” (National Center on Early Childhood Quality Assurance¹). The QRIS Network² describes a typical QRIS as one that (a) creates and aligns program and professional standards, (b) monitors and assesses programs based on the standards; (c) includes a process for supporting quality improvement; (d) provides financial incentives; and (e) disseminates information to parents and the public about program quality. The federal government also developed benchmarks for a Tiered Quality Rating and Improvement System (TQRIS)³, which included:

(a) tiered Program Standards with multiple rating categories that clearly and meaningfully differentiate program quality levels; (b) monitoring to evaluate program quality based on the Program Standards; (c) supports to help programs meet progressively higher standards (e.g., through training, technical assistance, financial support); and (d) program quality ratings that are publically available; and includes a process for validating the system.

The current study was designed to validate Step Up to Quality (SUTQ), Ohio’s QRIS.

SUTQ

Ohio’s Step Up To Quality (SUTQ) was conceptualized in 1999, when stakeholders began development of program standards, which were based on national research, for early education environments. Implementation of the three-tiered system began in 2005 in a sample of eight counties. The pilot was followed by a statewide scale-up in 2006. Both the Department of Job and Family Services (ODJFS) and Department of Education (ODE) license early learning and development programs: ODJFS licenses child care settings such as private child care centers and Type A or B homes, while ODE licenses programs in educational settings such as elementary schools.

In 2012, as a component of the Race to the Top Early Learning Challenge grant, stakeholders from a variety of agencies began a revision of SUTQ. The revised system has five tiers and organizes its program standards into four domains:

- **Learning and Development**, which consists of Curriculum & Planning, Child Screening & Assessment, and Interactions & Environment criteria.
- **Administrative and Leadership Practices**, which consists of Staff Supports, Program Administration, and Staff Management criteria.
- **Staff Qualifications and Professional Development**, which consists of Staff Education and Professional Development criteria.

¹ <https://eclkc.ohs.acf.hhs.gov/hslc/tta-system/pd/pds>

² <http://qrisnetwork.org/our-framework>

³ Race to the Top—Early Learning Challenge Application August 2011

- **Family and Community Partnerships**, which consists of Transitions and Communication & Engagement criteria.

The five-tiered SUTQ, which began implementation in October 2013, can be considered a hybrid of the building blocks and point system approaches, two common conceptualizations of QRIS. To advance in ratings from a 1-star to a 2-star or 3-star, early education programs must meet all requirements for the desired and the previous tiers (the building blocks approach). However, once a 3-star rating is achieved, programs can advance in tiers by accruing points for meeting additional quality criteria (the points approach). To advance to a 4- or 5-star rating, a program must document the accrual of at least one point in each of the four domains. For example, points may be accrued by documenting the use of intentional activities in all learning domains, the use of learning and exploration opportunities in the daily curriculum, or the ongoing, assessment-driven, evaluation of child progress. Programs also can earn points for maintaining low teacher-child ratios (i.e., fewer children per adult) or for achieving accreditation through one of the national accrediting bodies (e.g., National Association for the Education of Young Children).

Moving forward, SUTQ will become standard practice for an increasing number and proportion of early education programs. To wit, all Early Childhood Education and Preschool Special Education programs licensed by ODE must participate in SUTQ and be rated. As of 2020, all privately-owned and -operated child care facilities that provide services to children receiving public funding to support enrollment also must participate in SUTQ and be rated (including ODE-licensed programs that receive public funding through Publicly Funded Child Care).

SUTQ is a cross-agency, collaborative initiative, with active participation and guidance from ODJFS and ODE. Both agencies rate facilities (ODJFS) or programs (ODE), and provide support for participating sites, throughout the state, in the form of technical assistance, training, and a wide range of supporting documentation (www.earlychildhoodohio.org).

Currently, over 1,800 ODJFS facilities and over 600 ODE programs (including sites licensed by ODJFS) are rated in SUTQ. A high proportion of sites (81%) serve children receiving public funding to support their enrollment. In addition, sites are making progress through the star ratings. A review of changes in star rating over the past few years indicates that 61% of participating sites are advancing in star rating (while 33% are maintaining a star rating). This stated, it is important to note that the current study was conducted less than three full years after the transition from a three-tier to five-tier approach, and the transition incorporated the automatic migration of some sites to a higher star rating. In addition, Ohio's system for early learning and development is comprehensive in scope and has elements that were not targeted in the current study (such as the developmental progress of infants and toddlers or the quality of care and education for school-aged students). Findings should be interpreted in this context.

The SUTQ Rating Process

An early learning and development program begins participation in SUTQ by submitting an application to either the Bureau of Child Care Licensing and Monitoring at ODJFS or the Office for Early Learning and School Readiness at the ODE).

Sites applying for a rating through ODJFS

Licensing Specialists within the BCCLM review the application for eligibility and confirms the site is eligible to participate in SUTQ (e.g., is licensed, does not have a history of Serious Risk Non-Compliance violations). If the site application is for a 1- or 2-star rating, the Licensing Specialist will begin a desk review of the site's documents, including a review of staff educational attainment. Sites also receive an on-site verification visit from a Licensing Specialist. The on-site visit includes staff interviews, observations of group size and teacher-child ratios, and a staff and child file review. Sites applying for a 3-star or higher rating also receive classroom observations using the Ohio Classroom Observation Tool (OCOT).

Following the complete review of documentation, verification of staff education and qualifications, and completion of an on-site visit, the Licensing Specialist determines the site's star rating and notifies the site.

The specific criteria for each tier are codified in Ohio Administrative Code 5101: 2-17-01 and accompanying appendices. Appendices B and C of this report contain brief guidance documents for sites describing the criteria for specific star ratings.

Sites applying for a rating through ODE

Early Childhood Education and Preschool Special Education programs are required to have a 3-star or higher rating. The rating process begins with enrollment and assignment of roles to staff, within Ohio's Education Directory and Child Licensing and Quality Systems. Individual programs complete and submit registration documents and enter staff information into Ohio's Professional Registry (a database of early learning and development staff employed in programs across the state).

Staff within the Ohio Child Licensing and Quality System (OCLQS) complete an initial desk review of the documents submitted by the program. Any revisions or requests for additional information are sent back to the program, for its review and response. After the desk review is completed, an OCLQS consultant will complete an unannounced on-site visit to the program. A program's completed application, supporting documents, and star rating are approved by OCLQS and the program is informed of its rating.

Overview of QRIS Validation Studies

QRIS validation studies are performed to determine the extent to which the rating and improvement system is functioning as intended. A question of interest is whether the ratings used in the QRIS are practical and meaningful markers of the early educational quality experienced by children. Zellman and Feine (2012) outline four inter-related approaches to conducting QRIS validation studies:

- Examine the validity of key underlying concepts of basic QRIS quality components and standards;
- Examine the measurement strategy as well as the psychometric properties of the measures used for assessing quality to verify accurate results;

SUTQ Validation Study Results

- Assess the outputs of the rating process to ensure that ratings are functioning as intended; and
- Examine how ratings are associated with children’s outcomes.

For several years, QRIS studies have focused on verifying the degree to which quality standards and measurements result in accurate and meaningful ratings (Karoly, 2014; Tout & Starr, 2013; Zellman & Feine, 2012). More recently, studies also have included assessments of children’s developmental progress (Karoly, 2014; Thornburg et al., 2009; Tout et al., 2009; Zellman, Perlman, Le & Setodji, 2008).

The current validation study of Ohio’s SUTQ rating system combines the approaches identified above. In doing so, the study examines the extent to which SUTQ is associated with meaningful concepts and practices in high quality early education learning environments and addresses five types of validity:

- Face validity: the extent to which SUTQ criteria capture ideas and practices that are meaningful to practitioners and stakeholders in early education and child care.
- Content validity: the extent to which SUTQ criteria are comprehensive and inclusive of evidence-based and accepted ideas and practices.
- Construct validity: the extent to which SUTQ ratings are consistent with high quality early education environments.
- Procedural validity: the extent to which SUTQ uses a reliable process to generate site star rating.
- Predictive validity: the extent to which the SUTQ rating predicts child developmental progress or outcomes.

Overview of the Report

This report provides the findings for each of these five perspectives on validity. The study’s methodologies are presented first (chapter 2), followed by study findings (chapter 3), and study limitations and conclusions (chapter 4). The appendices attached to this report contain additional information about the study methodologies and findings.

Methodology

General Methodology

The Ohio SUTQ validation study incorporated a concurrent mixed-methods design, which is a combination of both quantitative and qualitative data collections and analytic techniques. In brief, the study:

- Created a stratified random sample of early learning and development sites, some of which are participating and some of which are not participating in SUTQ. The sample included private child care centers, Type A and B homes, and elementary schools with star-rated Early Childhood Education classrooms. A total of 81 sites, and over 190 classrooms, received on-site visits and observations.
- Completed a literature review of QRIS validation studies and research related to SUTQ's domains and components.
- Conducted a suite of classroom and site-level assessments at each sampled site, including assessments of:
 - Quality of the early education environment;
 - Teacher-child interactions;
 - Early language and literacy practices;
 - Family and provider (or teacher) relationship quality; and
 - At selected sites, the quality of administrative practices.
- Assessed language and literacy skills of 3- and 4-year old children at sampled centers and Early Childhood Education classrooms.
- Analyzed the kindergarten readiness of publicly funded children from across the state.

Additional details on study methodology are presented in Appendix D.

DESCRIPTION OF THE SAMPLES

Sample 1: Stratified Random Sample of SUTQ Participating and Non-Participating Sites

The sample of sites in the Ohio SUTQ study is a probability sample of state-registered early childhood sites. Ohio's Department of Job and Family Services and Ohio Department of Education provided lists of these sites with utilization data and star quality rating as of February 2016. Sites that had preschool age enrollment were included for sampling, while sites without preschool enrollment and day camps were excluded. The sites first were sorted into stratified groups following the random block study design and an equal number of sites were allocated per group. The strata included star rating and location (urban versus non-urban)⁴.

Table 1 shows the response rate of the eligible sites and the co-operation rate of all sites that were included. For each site in the initial sample, up to two replacement sites were identified based on

⁴ Appendix D provides further details on the sampling methods for selecting child care centers, elementary schools, and homes.

SUTQ Validation Study Results

the sample selection order. These replacement or reserved sites were only contacted if and when the initial sampled site chose not to participate. As a result, the study team contacted a total of 290 sites; the overall co-operation rate (agreement to participate in the study) was 28%. Using replacement sites as a proxy, the response rate among the eligible sites is 58%, a rate that is fairly consistent across program type: 57% for centers, 60% for elementary schools, and 61% for homes. The final sample is presented in Table 2; a total of 81 sites participated in the study.

Table 1. Response rate and co-operation rate

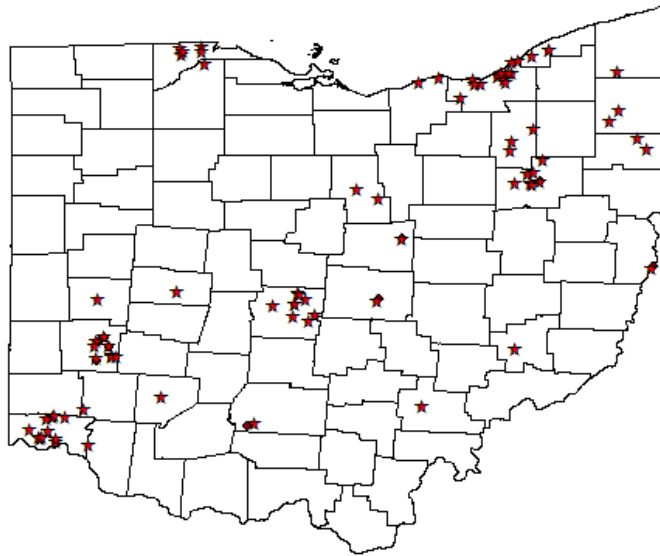
	Number Agreed	Number Eligible	Response Rate	Number Included	Co-operation Rate
Child care centers	55	96	57%	205	27%
ODE-licensed classrooms	9	15	60%	28	32%
Type A & Type B Home	17	28	61%	57	30%
Total sites worked	81	139	58%	290	28%

Table 2. Number of participating sites by program type and SUTQ rating

	Non-rated ⁵	1-star	2-stars	3-stars	4-stars	5-stars	Grand Total
Child Care center	8	9	10	11	10	7	55
Early Childhood Education classrooms in elementary schools				1	3	5	9
Type A & Type B Home	1	3	3	7	1	2	17
Total Participating Sites	9	12	13	19	14	14	81

⁵ Non-rated refers to sites (and classrooms within sites) that do not participate in SUTQ.

Exhibit 1. Distribution of participating sites across the state



Sample 2: Extant Data Extraction

The study team worked with Ohio’s Departments of Job and Family Services and Education to identify and extract KRA scores and pre-kindergarten participation data for children who received publicly funded services (including programs offered through ODE). KRA data were available for the 2014-2015 and 2015-2016 school years. Pre-kindergarten participation data were available for the 2013-2014 and 2014-2015 fiscal years. The study team extracted and analyzed variables specific to KRA scores, including the total score and several subscales: Language and Literacy, Mathematics, Social Foundations, and Physical Well-Being and Motor Development. Child demographic data and disability status also were available. Additional variables included total weeks of attendance in pre-kindergarten programming and pre-kindergarten site star rating.

INSTITUTIONAL REVIEW APPROVAL

The study team applied for and received Institutional Review Board (IRB⁶) approval for its study design, including the documents used to provide and receive informed consent from participating sites and the parents of participating children. Copies of these documents are provided in Appendix D.

⁶ Approval was granted by Westat, Inc. Institutional Review Board, IRB Registration # 695; Federalwide Assurance (FWA) #: 5551.

INSTRUMENTATION

The study team used a suite of instruments to conduct observations, gather feedback and input from site directors, teachers, and parents, gather feedback from state stakeholders, and assess language and literacy skills in participating children. The instruments and tools used are presented below.

Environment Rating Scales (ERS). The ERS commonly are used for assessing the process quality of early education programs. Three of the scales were incorporated into the study: the Infant/Toddler Environment Rating Scale-Revised (ITERS-R, Harms et.al. 2006), the Early Childhood Environment Rating Scale-3 (ECERS-3, Harms et.al. 2015), and the Family Child Care Environment Rating Scale-Revised (FCCERS-R, Harms et.al. 2007). The ERS are comprised of multiple subscales, including Space and Furnishings, Personal Care Routines, Language and Reasoning, Activities, Interactions, Program Structure, and Parents and Staff. The ERS, in conjunction with the CLASS (described below) have been used in at least 10 other validation studies nationally. In addition, the ERS were used in Ohio's 2005-2007 and 2009-2011 SUTQ validation studies.

Classroom Assessment Scoring System (CLASS). The CLASS commonly is used for the assessment of the quality of interactions between teacher and children. The CLASS, which is available for center-based classrooms, was used to assess interactions in domains such as Responsive Caregiving (CLASS Infant; Hamre, et.al. 2014), Emotional and Behavioral Support and Engaged Support for Learning (CLASS Toddler; Le Paro et.al. 2012), and Emotional Support, Classroom Organization, and Instructional Support (CLASS PreK; Pianta et.al. 2008). The CLASS also was used in Ohio's 2009-2011 SUTQ validation study.

Early Language and Literacy Classroom Observations (ELLCO) and Child Home Early Language and Literacy Observation (CHELLO). The ELLCO (Smith et.al. 2008) was developed to assess environmental and interactional support for language and literacy development. The ELLCO has two primary subscales: General Classroom Environment and Language and Literacy. The CHELLO fulfills a similar purpose, for family home child care providers, and has two interdependent scales: the Literacy Environment Checklist and the Group/Family Observation (used in concert with the Provider Interview).

Family and Provider/Teacher Relationship Quality Scale (FPTRQ). The FPTRQ was developed relatively recently (Kim et.al. 2014). Its purpose is to capture the quality of family and provider/teacher relationships in early learning and development programs. Sponsored by the Administration for Children and Family, the instrument is in the public domain (<https://www.acf.hhs.gov/opre/research/project/development-of-a-measure-of-family-and-provider-teacher-relationship-quality-fptrq>). The current study incorporated the Director, Teacher, and Parent measures, which targeted site and teacher practices for engaging and working with parents.

Program Administration Scale (PAS) and Business Administration Scale (BAS). The Program (and Business) Administration Scales were developed to assess the quality of administrative, management, and leadership practices in centers (PAS; Talan & Bloom, 2011) and homes (BAS; Talan & Bloom, 2009). There are multiple dimensions incorporated into each assessment, capturing concepts such as human resources, budgeting, marketing, and so forth. The current study incorporated the PAS and BAS with a subset of participating sites, based on the sites' interest and willingness to complete the comprehensive assessment.

Brigance Inventory of Early Development III (IED-III). The IED-III (Curriculum Associates, 2016) assesses development in gross and fine motor skills, receptive and expressive language, literacy and mathematics, adaptive behaviors, and interpersonal and self-regulatory skills. The IED-III contains 55 items that are norm-referenced, allowing the comparison of study children with national norms in the cited domains. The current study used the Language Development and Literacy portions of the tool. The instrument was completed only for children with signed parent consent; child names were replaced with ID codes to ensure confidentiality.

SpecialLink Early Childhood Inclusion Scale. The SpecialLink Early Childhood Inclusion Scale was developed to assess the quality of early childhood environments and teaching practices experienced by children with special learning or development needs (Irwin, 2009). The scale incorporates items that assess both practices and principles, including items such as the physical environment for special needs, staff training, and therapies. The scale was intended for use with any inclusion classrooms participating in the study.

Site Questionnaire. The study team developed a Site Questionnaire for completion by all participating site directors or owners. The questionnaire was informed by a review of SUTQ criteria for advancing in star rating, and was designed to prompt participants to identify which of the criteria they agreed were indicative of high quality early learning environments. The questionnaire also allowed for feedback on SUTQ as well as input on the nature and scope of supports that would be helpful for achieving or improving the quality of care and education. The questionnaire was reviewed for content, readability, and ease of use before it was distributed to participants. Internal consistency statistics (i.e., Cronbach's alpha) for the questionnaire were calculated and are reported in Appendix D.

Child and Family Questionnaire. The study team developed a Child and Family Questionnaire to capture information on child and family background, pertinent for an analysis of child developmental status (as assessed using the Brigance IED-III). The questionnaire contained items for child characteristics, parent and family characteristics, and history in child care settings. The questionnaire also prompted parent respondents to indicate what, if any, criteria they agreed suggested high quality early learning environments. The questionnaire was reviewed for content, readability, and ease of use before it was distributed to participants.

Site Education Profile. To further investigate the importance of both educational attainment and experience, the study team developed a Site Education Profile that captured the educational level and field of each site's director and lead teachers. The profile also captured total years of experience in early childhood.

Stakeholder Interview Protocol. The study team developed a standard protocol for interviews conducted with state stakeholders including staff in Job and Family Services, the Department of Education, the Governor's Office, Child Care Resource and Referral agencies, and State Support Teams.

SUTQ Online Survey. The study team developed an online survey to capture feedback on SUTQ from any site in Ohio, regardless of participation in the study. The online survey asked respondents to provide general feedback on SUTQ and to identify challenges (if any) experienced in participating in SUTQ or in operating their site. The online survey was available for the duration of the study.

DATA COLLECTOR TRAINING AND QUALITY CONTROL

The data collection team received training in each instrument. Team members were required to satisfactorily complete the publisher’s CLASS, PAS, and BAS reliability assessments in order to use these tools. For the ERS, team members completed ERS Institute training and were required to obtain 80% or higher inter-rater reliability with an experienced and highly reliable rater, in order to use the scales. Team members were trained by the Curriculum Associates’ Brigance IED-III team on the use of the child assessment tool, and completed assessments under the direction of lead team members at each site. Similarly, team members were trained by the publishers of the ELLCO/CHELLO assessments to publisher specifications. Finally, all team members were required to complete the National Institutes of Health (NIH) “Protecting Human Research Participants” training course and worked under the direction of an Ohio-based field coordinator and a lead team member at each site.

The field coordinator undertook several quality control exercises throughout the data collection period. In addition to ensuring all trainings were completed, to the necessary level of reliability, the field coordinator managed the scheduling of sites and communicated weekly with team members regarding observations and experiences at individual sites. The field coordinator was trained to reliability on each of the instruments, across centers and homes, and reviewed data collections weekly, including the scoring and secure uploading of data into the project-developed databank.

DATA REVIEW AND VERIFICATION

Team members securely uploaded data each week into the project’s databank. Project managers securely accessed and retrieved the data and checked for data entry errors, scoring errors, and overall cross-member consistency and reliability. Raw data were maintained by the project managers and used to verify scoring and data entry.

Additional details about study reliability are presented in Appendix D: Methodology.

Face Validity

In this study face validity was defined as the extent to which SUTQ criteria capture ideas and practices that are meaningful to practitioners and stakeholders in early education and child care. The following study question was of particular interest to the study team:

To what extent do Ohio child care program directors and owners agree that SUTQ captures ideas and practices that are meaningful and indicative of high quality early learning environments?

The study addressed face validity using a Site Questionnaire completed by sampled and enrolled study sites as well as interviews with state stakeholders. The Site Questionnaire was developed using the SUTQ criteria for star rating (Appendices B and C)⁷. Excerpted and annotated SUTQ criteria were included as questionnaire items, which were grouped into five constructs: (1) Learning and Development, (2) Administrative and Leadership Practices, (3) Staff Qualifications and Professional Development, (4) Family and Community Partnerships, and (5) Group Size and Accreditation. Each construct contained at least one sub-construct (or component; each sub-construct/component was composed of at least one item. Participants were asked to indicate which of the items they agreed were indicative of high quality early education.

HIGHLIGHTS

- Participating sites and stakeholders reported that Learning and Development, Administrative and Leadership Practices, and Family and Community Partnerships criteria are indicative of high quality early education.
- There was less agreement among sites and stakeholders as to the levels of staff qualifications, professional development, and accreditation that are indicative of high quality early education.
- Parents value nurturing behaviors and interactions, more so than teacher educational achievement.

SITE QUESTIONNAIRE

The director or owner at each study site (n=81) received a Site Questionnaire. A total of 79 questionnaires were returned (a response rate of 98%) and were incorporated into analyses. For scoring purposes, each questionnaire item targeting SUTQ criteria was worth 1 point⁸. This generated a total score for the questionnaire as well as scores in constructs and components. Questionnaire, construct,

⁷ A copy of the instrument is presented in Appendix E.

⁸ There were two exceptions to the scoring approach. First, in the component “Staff Education”, any items referencing Associates-level education were scored at “1” point. Items referencing Bachelors-level education was scored at “2” points. For each survey returned, which contained affirmative responses in this component, either Associates-level or Bachelor’s level items were scored. Second, in the component “Professional Development”, the item “Ensuring administrators, lead teachers and assistant teachers receive 20 hours of Ohio- approved specialized training every two years” was worth “1” point. The item “Ensuring administrators, lead teachers and assistant teachers receive 25 hours of Ohio- approved specialized training every two years” was worth “2” points, and the item “Ensuring administrators, lead teachers and assistant teachers receive 30 hours of Ohio- approved specialized training every two years” was worth 3 points. The highest ranked item on each survey was scored.

and component scores were used to calculate the percent agreement of survey respondents with SUTQ domains and components. The total possible score, if each item was checked in agreement, is 71.

Overall, 89% of questionnaire respondents agreed that SUTQ domains and components are indicative of high quality early education.

When agreement was defined as a benchmark of **50% agreement** with questionnaire items, 89% of respondents, overall, agreed that questionnaire items were indicative of high quality early education. Agreement at this level was relatively high across site type, as shown in Table 3.

When the definition of agreement was raised to a benchmark of **75%**, the percent agreement fell to an overall rate of 62%, with the highest level of agreement given by respondents in Early Childhood Education classrooms, and the lowest by respondents from Type A or B homes. Finally, when the benchmark was raised to **90%**, the percent agreement fell further, to an overall 22%. Of note, only 6% of Type A or B home respondents indicated agreement at this level of measurement, compared to over half (56%) of the respondents from Early Childhood Education classrooms.

Table 3. Percent of respondents who agreed that questionnaire constructs and components are indicative of quality, by program type

	Aggregate (n=79)	Centers (n=53)	Homes (n=17)	Early Childhood Education Classrooms (n=9)
% of respondents who agree that at least 50% of questionnaire items represent quality	89%	89%	82%	100%
% of respondents who agree that at least 75% of questionnaire items represent quality	62%	66%	35%	89%
% of respondents who agree that at least 90% of questionnaire items represent quality	22%	21%	6%	56%

Data source: Site Questionnaire

These overall levels of agreement suggest that SUTQ foundations are robust and consistent with provider and teacher views regarding quality early education. The findings also suggest that some items and components are less important for providers and teachers. These are explored in more detail, below.

INTRA-CONSTRUCT AGREEMENT

Intra-construct agreement was examined to explore patterns in participant responses. The benchmark of agreement with at least 50% of items was again used to determine agreement with the construct and components.

There was relatively high agreement (95% or higher) with the construct Learning and Development (Table 4). The construct's components represent the quality of planned programming, including (a)

Overall, 95% of questionnaire respondents agreed that Learning and Development components are indicative of high quality early education.

curriculum and planning tasks and activities; (b) the use of child screenings and assessments to inform educational planning and communications with families; and (c) a continuous improvement process to ensure high quality interactions and environment. As can be seen, agreement was robust across component and respondent type.

Table 4. Percent of respondents who agreed that Learning and Development practices are indicative of quality, by program type

Agreement with at least 50% of construct items	Aggregate (n=79)	Centers (n=53)	Homes (n=17)	Early Childhood Education Classrooms (n=9)
Learning and Development	95%	94%	94%	100%
<i>Curriculum and Planning</i>	94%	94%	88%	100%
<i>Child Screening and Assessment</i>	92%	93%	88%	100%
<i>Interactions and Environment</i>	94%	93%	94%	100%

Data source: Site Questionnaire

For the construct Administrative and Leadership Practices (Table 5), there was relatively high agreement (using agreement with at least 50% of items as a benchmark) with the construct, with respondents from Early Childhood Education classrooms reporting the highest level of agreement and respondents from Type A and B homes reported the lowest level of agreement.

One component exhibited considerably lower levels of agreement among respondents from Type A and B homes: Staff Supports 2. This section of the questionnaire targeted the number and nature of staff benefits, including (a) paid leave; (b) health and retirement benefits; and (c) discounts and reimbursements for continuing education, among others. This response pattern is consistent with the struggles some Type A and B homes may experience in budgeting and providing for staff benefits and incentives.

Overall, 91% of questionnaire respondents agreed that Administrative and Leadership Practices are indicative of high quality early education.

Table 5. Percent of respondents who agreed that Administrative and Leadership Practices are indicative of quality, by program type

Agreement with at least 50% of construct items	Aggregate (n=79)	Centers (n=53)	Homes (n=17)	Early Childhood Education Classrooms (n=9)
Administrative and Leadership Practices	91%	93%	71%	100%
<i>Staff Supports 1</i>	86%	93%	59%	100%
<i>Staff Supports 2</i>	70%	77%	35%	89%
<i>Program Administration</i>	90%	94%	94%	100%
<i>Staff Management</i>	89%	91%	77%	100%

Data source: Site Questionnaire

SUTQ Validation Study Results

The overall level of agreement for the components Staff Qualifications and Professional Development was considerably lower, as shown in Table 6. Only 58% of respondents agreed that at least 50% of items related to Staff Qualifications (and 47% of respondents for Professional Development) were indicative of high quality early education. Consistently, respondents from Early Childhood Education classrooms reported higher levels of agreement while respondents from Type A and B homes tended to report lower levels of agreement (note the exception for Professional Development, in which the lowest level of agreement was reported by respondents from child care centers).

Overall, 58% of questionnaire respondents agreed that Staff Qualifications and 47% agreed that Professional Development components are indicative of high quality early education.

Table 6. Percent of respondents who agreed that Staff Qualifications and Professional Development are indicative of quality, by program type

Agreement with at least 50% of construct items	Aggregate (n=79)	Centers (n=53)	Homes (n=17)	Early Childhood Education Classrooms (n=9)
Staff Qualifications	58%	60%	35%	89%
Professional Development	47%	42%	59%	56%

Data source: Site Questionnaire

As regards Family and Community Partnerships (Table 7), overall, 89% of respondents reported agreement with the construct, which incorporated items related to practices for transitioning students and families across classrooms or settings and strategies for providing outreach to and working with families.

Table 7. Percent of respondents who agreed that Family and Community Partnerships are indicative of quality, by program type

Agreement with at least 50% of construct items	Aggregate (n=79)	Centers (n=53)	Homes (n=17)	Early Childhood Education Classrooms (n=9)
Family and Community Partnerships	89%	89%	82%	100%
Transitions	82%	83%	77%	89%
Communication and Engagement	84%	83%	77%	100%

Data source: Site Questionnaire

Finally, there was good agreement that group size is indicative of quality (89% agreement, overall) but less agreement regarding the importance of accreditation (51% agreement, overall), as shown in Table 8.

Overall, 89% of questionnaire respondents agreed that Group Size is indicative of high quality early education.

Table 8. Percent of respondents who agreed that Group Size and Accreditation are indicative of quality, by program type

Agreement with at least 50% of construct items	Aggregate (n=79)	Centers (n=53)	Homes (n=17)	Early Childhood Education Classrooms (n=9)
Group Size	89%	91%	77%	100%
Accreditation	51%	51%	53%	44%

Data source: Site Questionnaire

The next step was to disaggregate responses by the star rating of each site and again examine the level of agreement (Table 9). This process revealed several interesting findings:

- At the **50%** benchmark of agreement, 100% of 4- and 5-star rated sites were in agreement that questionnaire items were indicative of high quality early learning
- At the **75%** benchmark, there were increasing levels of agreement by star rating, with the exception of 1-star rated sites, who reported higher levels of agreement than non-participating and 2-star rated sites

There was no discernable pattern to the levels of agreement at the 90% benchmark.

Table 9. Percent of respondents who agreed that questionnaire items are indicative of quality, by star rating

	Non-rated (n=8)	1-star (n=12)	2-stars (n=13)	3-stars (n=19)	4-stars (n=13)	5-stars (n=14)
Overall Scale: at least 50% agreement with questionnaire items	88%	83%	69%	90%	100%	100%
Overall Scale: at least 75% agreement with questionnaire items	25%	58%	54%	58%	77%	86%
Overall Scale: at least 90% agreement with questionnaire items	13%	42%	8%	16%	8%	43%

Data source: Site Questionnaire

Finally, examining findings by construct and star rating confirms that Staff Qualifications and Professional Development are the constructs that consistently receive the lowest levels of agreement (measured against the benchmark of agreement with at least 50% of questionnaire items; Table 10). Further, there were no discernable patterns or trends associated with level of agreement and star rating, especially for the constructs Learning and Development, Administrative and Leadership Practices, and Staff Education and Professional Development.

Respondents across star rated sites consistently gave the least amount of agreement to items related to Staff Qualifications and Professional Development.

Table 10. Percent of respondents who agreed that questionnaire constructs are indicative of quality, by star rating

Agreement with at least 50% of construct items	Non-rated (n=8)	1-star (n=12)	2-stars (n=13)	3-stars (n=19)	4-stars (n=13)	5-stars (n=14)
Learning and Development	100%	92%	85%	95%	100%	100%
Administrative and Leadership Practices	100%	83%	77%	84%	92%	100%
Staff Qualifications	38%	75%	54%	42%	69%	71%
Professional Development	25%	50%	31%	47%	54%	64%
Family and Community Partnerships	88%	75%	77%	90%	100%	100%
Group Size	75%	75%	77%	95%	100%	100%
Accreditation	50%	58%	31%	63%	46%	50%

Data source: Site Questionnaire

Face validity, which can be considered a rough guide to the appropriateness and meaningfulness of SUTQ criteria, is established at a foundational level—meaning that there is relatively high agreement with at least 50-75% of items similar to those used to determine SUTQ rating. However, there is not uniform agreement across all items—Staff Qualifications, Professional Development practices, and Accreditation received the lowest levels of support—suggesting that some criteria are either not associated with quality by practicing providers or teachers or are duplicative or superfluous, when considered in concert with other items.

There also is not uniform agreement when items are examined by star rating but there are few discernable trends in responses. Higher rated sites, for example, tend to generate higher levels of agreement. However, there are exceptions to this trend, especially when examined by construct. The constructs with the highest levels of agreement tend to be practice-oriented items—such as classroom practices that support learning and development, administrative and leadership practices, family outreach and child transition practices, and group size. Practices that received the lowest levels of support tend to be status-related items, such as level of education, accrual of professional development hours, or accreditation. Thus, in defining quality, there appears to be a distinction between what providers, teachers, and directors do with children, families, and staff as opposed to what they have achieved as professionals.

STAKEHOLDER INTERVIEWS AND CHILD AND FAMILY QUESTIONNAIRES

Additional insight can be gained from stakeholder interviews and parent feedback. In general, interview participants agreed that the SUTQ criteria are indicative of and important for high quality early learning, not least because of the comprehensiveness of the initiative and the research supporting each of the domains and components. Interview participants described SUTQ as a blueprint that provides guidance for participating providers, especially those who may need assistance in determining how to improve quality. It also was clear from interviews that some design questions remain, particularly with regard to the level of education teachers should have, the amount of paperwork required for participation in SUTQ, and the translation of quality practices as described on paper to quality practices implemented in the classroom. There was a general consensus that the SUTQ concepts are sound but that some sites may experience challenges in implementing and sustaining high quality practices over time.

Finally, it also was important to hear from parents of children enrolled in participating sites, as to the factors that were indicative of high quality. Parent surveys were received from almost 200 parents of 3- and 4-year old children (parents of children who participated in child assessments). The parent survey presented a curated list of items from the

Site Questionnaire, as shown in Exhibit 2. Parent survey items related to indicators of quality

Exhibit 2. Parent survey items related to indicators of quality

- The provider uses lesson plans that work for the age of my child.
- Teachers follow Ohio guidelines for creating and using lesson plans.
- Teachers use lots of different types of activities to promote child learning.
- Teachers regularly test my child for how well he or she is learning.
- Programs ensure that child are in warm and nurturing classrooms.
- Programs ask for parent feedback when creating learning plans for their child.
- Programs take care of their staff with different types of benefits.
- Programs are always trying to find ways to improve their quality.
- Teachers regularly go to trainings.
- Programs hire teachers that have a college education.
- The director or owner of the child care site has a college education.
- Programs create special activities to help when children start at the site or move into a new classroom.
- Programs reach out to parents and find ways to send information home.
- Programs have learning activities and events for parents.
- Programs like hearing from parents and having parents visit on-site.

The highest levels of parent agreement were reported for:

- Use of different types of activities to promote child learning (99% agreement).
- Program ensures child is in a warm and nurturing environment (99% agreement).
- Programs always trying to improve quality (97% agreement).
- Use of lesson plans that work for the age of my child (95% agreement).
- Sharing information with parents (95% agreement).

Parent survey participants value activities that promote child learning and warm and nurturing environments, as indicators of quality.

SUTQ Validation Study Results

- Use of special activities to assist transitions (e.g., starting at the site or in a new classroom; 92% agreement).

Items with the least agreement from parents included:

- Teachers have a college education (34% agreement).
- Director or owner has a college education (74% agreement).
- Programs provide learning activities and events for parents (74% agreement).
- Teachers use Ohio guidelines for creating and using lesson plans (75% percent).

These findings suggest that at least some parents also are attuned to what providers and teachers are doing, as opposed to what they have achieved educationally. The findings from the current surveys echo those from at least one validation study in another state: a study of Mississippi's Quality Stars program found that parents valued staff that were nurturing, attentive, and passionate about children; two-way communication and open-door policies; and curriculum, school readiness, opportunities for socialization, and classroom materials (DeMarco et.al. 2015). Further, parental choice of child care may be informed by logistical factors such as distance and affordability (as suggested by Shlay, 2010, in a study of parental preferences in Pennsylvania).

The lack of emphasis on education, by both providers and parents, is worth noting. This trend suggests a belief that high quality experiences can be provided without advanced education (four-year degree or higher, for example). However, many within the profession (including, as will be shown below, each state that has invested in a QRIS) consider education to be a critical feature for both advancing the quality of care and the profession.

Content Validity

In the current study, content validity reflects the extent to which SUTQ criteria adequately capture a concept or construct that has meaning for assessing and rating quality early learning environments. To assess content validity, the study team conducted a review of constructs shown to impact quality in extant research as well as a review of components and standards included in QRIS across the United States. The primary objective of this phase of the study was to determine which SUTQ features also are empirically supported and found in other state QRIS initiatives. The questions that were of interest to the study team were:

- ***To what extent does research support the rating criteria used in SUTQ?***
- ***To what extent do other states incorporate the same rating criteria, as used in SUTQ?***

What follows is a brief review of literature⁹ and a comparison of criteria across states, for each rating component (including group size and accreditation).

Learning and Development

CURRICULUM AND PLANNING

Support for the use of an evidence-based, developmentally appropriate curriculum comes from studies such as Schweinhart and Weikart's (1997) examination of the High/Scope preschool curriculum, Domitrovich et al.'s (2007) evaluation of the PATHS socio-emotional curriculum, Clements and Samara's (2008) examination of the Building Blocks mathematics curriculum, and Lonigan et al.'s (2015) study of a comprehensive school readiness curriculum. Attention also should be paid to studies such as that conducted by Justice et al. (2008), which found that the presence of a curriculum, even when faithfully implemented, does not necessarily equate to high quality instruction.

The presence of many different curricula, however, may lead to questions about which curriculum or approach is best. The National Association for the Education of Young Children, in conjunction with the National Association of Early Childhood Specialists in State Departments of

HIGHLIGHTS

- Staff qualifications are a standard in every state QRIS. Environment and interactions also are highly prevalent, as is the use of a curriculum, and family partnerships.
- The least common components across state QRIS are transitions, group size, and accreditation.
- Research findings highlight the ever-evolving understanding of factors that contribute to quality, especially as related to positive child outcomes.

90% of state QRIS include curriculum and planning components in their standards for child care centers.

⁹ The early childhood research base is immense, and growing. What is presented is a brief discussion of topics and support for each component, as applicable.

Education, issued a joint statement in 2003 that emphasized the importance of implementing “... *curriculum that is thoughtfully planned, challenging, engaging, developmentally appropriate, culturally and linguistically responsive, comprehensive, and likely to promote positive outcomes for all young children.*” The two national groups went on to identify the following standards by which the quality of a curriculum could be assessed:

- Children are active and engaged,
- Goals are clear and shared by all,
- Curriculum is evidence-based,
- Valued content is learned through investigation, play, and focused, intentional teaching,
- Curriculum builds on prior learning and experiences,
- Curriculum is comprehensive,
- Professional standards validate the curriculum’s subject-matter content, and
- The curriculum is likely to benefit children.

Thus, it is unsurprising that curricula (and practices such as lesson planning) find their way into a state’s QRIS. Less research has been conducted on lesson plans in early childhood or preschool classrooms, although there are multiple resources guiding the development of lesson plans. Nevertheless, it stands to reason that planning is an important aspect of consistent and high quality instruction.

CHILD SCREENING AND ASSESSMENT

The joint statement issued by the National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education also contained language regarding child assessments. Specifically, the statement recommends that early childhood educators:

Make ethical, appropriate, valid, and reliable assessment a central part of all early childhood programs. To assess young children’s strengths, progress, and needs, use assessment methods that are developmentally appropriate, culturally and linguistically responsive, tied to children’s daily activities, supported by professional development, inclusive of families, and connected to specific, beneficial purposes: (1) making sound decisions about teaching and learning, (2) identifying significant concerns that may require focused intervention for individual children, and (3) helping programs improve their educational and developmental interventions.

75% of state QRIS include child screening and assessment indicators in their standards for child care centers.

The national bodies identify standards for effective child assessment that include:

- Ethical principles guide assessment practices,

- Assessment instruments are used for their intended purposes,
- Assessments are appropriate for ages and other characteristics of children being assessed,
- Assessment instruments are in compliance with professional criteria for quality,
- What is assessed is developmentally and educationally significant,
- Assessment evidence is used to understand and improve learning,
- Assessment evidence is gathered from realistic settings and situations that reflect children’s actual performance,
- Assessments use multiple sources of evidence gathered over time,
- Screening is always linked to follow-up,
- Use of individually administered, norm-referenced tests is limited, and
- Staff and families are knowledgeable about assessment.

It may be helpful to consider the primary purposes that assessments commonly serve in early childhood settings:

- 1) Identify children who may be in need of specialized services;
- 2) Plan instruction for individuals and groups of children;
- 3) Identify program improvement and staff development needs; and
- 4) Evaluate how well a program is meeting goals for children (Schilder & Carolan, 2014).

There are many validated tools that can be used with children, starting in infancy (e.g., Teaching Strategies GOLD, Work Sampling System, Developmental Indicators for the Assessment of Learning, etc.; Schilder & Carolan, 2014). Some states invest in creating their own assessment tools, to align with and support guiding principles for early learning and school readiness. Assessment standards for center-based care are incorporated into 75% of state QRIS, indicating their utility and popularity for improving quality.

Interactions and Environment

There are many studies documenting the importance of high quality early childhood education to positive child (and life) outcomes. Classic studies such as High/Scope Perry Preschool¹⁰ and the Abecedarian Project¹¹ have tracked participating children through their lives and have documented the positive impact of high quality early education on child and life outcomes (such as academic achievement, identification and

90% of state QRIS include interactions and environment indicators in their standards for child care centers.

¹⁰ <http://www.highscope.org/content.asp?contentid=219>

¹¹ <http://abc.fpg.unc.edu/>

placement into special education or learning services, overall health, high school completion, and employment).

In considering what constitutes high quality early education, some authors, such as Manning et.al. (2015), discuss the distinction between “structural” and “process” quality, noting the importance of process quality for child social and academic development. The authors identify teacher-child ratios, group size, and teacher education as structural components and teacher-child interactions and classroom dynamics as process components. The authors also identify several instruments that can be used to assess interactions, including the CLASS and the CIS, which are used in the current study. The ERS are cited as indicators of global quality; three of the ERS scales also are used in the current study.

Recently, Hatfield et.al. (2016) examined the importance of teacher-child interactions, documenting the numerous studies that have linked interactions to positive child outcomes. The authors also posited that threshold levels of quality exist, beyond which child outcomes emerge or accelerate. The authors documented support for theories linking positive outcomes to threshold (and high quality) levels of instruction, as measured by the CLASS. The authors also make the point that this is an active area of research, and that the influence of quality upon outcomes may in fact be curvilinear (instead of linear).

SUMMARY

Ohio's SUTQ integrates Curriculum and Planning, Child Screening and Assessment, and Interactions and Environment components into its Learning and Development domain. To advance in star rating, sites must document advanced use or understanding of best practices in each of these components. Many other states also include these components in their QRIS: According to the QRIS Compendium¹², in 2016, 90% of QRIS include these components in center-based standards (and 92% include them in home-based standards). Further, 90% of QRIS standards across the country include curriculum indicators for centers (82% include curriculum indicators for homes) and 75% include child assessment indicators for centers (72% include assessment indicators for homes).

Administrative and Leadership Practices

STAFF SUPPORTS

Staff supports such as a formal wage structure and benefits are important in that they contribute to staff retention. Studies of wage supplement programs, for example, have linked higher salaries to a decrease in intention to leave or a decrease in staff turnover (Gable et.al. 2007; Child Care Services Association, 2015).

65% of state QRIS include staff support components in their standards for child care centers.

Studies also have documented the impact of staff turnover on classroom quality. Cassidy et.al. (2011), for example, found that ECERS-Revised (ECERS-R) scores dropped after a site experienced staff turnover and transition, while Torquati et.al. (2007) found that

¹² <http://qriscompendium.org/top-ten/question-4>

compensation was associated with global quality. Thus, the retention of qualified and experienced early educators can be important for the achievement and maintenance of high quality classrooms. This stated, it is important to consider the costs of attracting and retaining highly qualified staff, a significant barrier for many independent child care sites.

PROGRAM ADMINISTRATION (PAS)

Program administration is a broad category that can include many operational aspects of a child care or early education facility. As codified in the PAS, this component may include items such as human resources, budgeting and fiscal management, program planning and evaluation, family partnerships, child assessments, staffing, and public relations (among others). Further, high quality management may be linked to long-term sustainability of program quality (Heinemeier and Leonard, 2013).

88% of state QRIS include program administration and management indicators in their standards for child care centers.

SUTQ addresses program administration through standards for self-assessment and continuous improvement. More common in the primary school literature, the processes of self-assessment and development of Continuous Improvement Plans can be important indicators of a program's commitment to and investment in quality. Barnett (2008), for example, links coaching and supervision practices to an ongoing cycle of continuous improvement in teaching practices. Frede (2005) provides a model for continuous improvement that includes setting standards, measuring and assessing progress, analyzing and planning improvements, and providing professional development and technical assistance. In short, research suggests the ongoing provision of supports such as training, technical assistance, coaching, and mentoring can be linked in purposeful ways to the quality of early education classrooms.

63% of state QRIS include continuous quality improvement indicators in their standards for child care centers.

STAFF MANAGEMENT

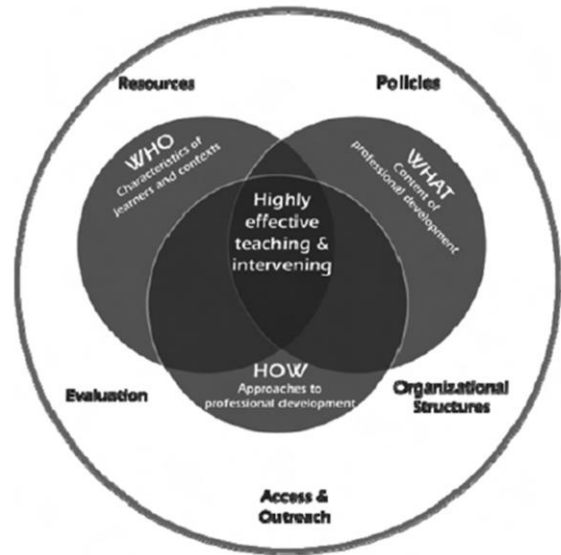
The importance of professional development in general is discussed in the next section. In this domain, however, the importance of classroom assessments, continuous improvement, and professional development plans is addressed. Practically-speaking, the use of tools such as action, continuous improvement, and professional development plans can be important heuristics for achieving and sustaining high quality practices. The North Carolina Institute for Early Childhood Professional Development (2001), for example, identified professional development in general and a plan in particular as a means of advancing up an early childhood career ladder.

There is no one way to construct a professional development plan; the process is perhaps more important than the template. In thinking through the process of constructing a plan, it may be helpful to consider the framework suggested by Buysse et. al. (2009; Exhibit 3), which emphasizes the who, what, and how of professional development. Alternately, process considerations may be informed by authors such as Han (2014) who suggest a five-step process for professional development planning as it relates to the development of child social competence: (1) identifying professional development content based on participants' contextual needs; (2) sequencing professional development aligned with the positive behavior support framework; (3) providing opportunities for participants to receive feedback on their implementation; (4) guiding participants to reflect on their own practices; and (5) embracing socio-cultural perspectives throughout the professional development process. All told, an action or professional development plan, especially one informed by evaluation of teaching practices and needs and aligned with supports for continuous improvement, can be critical for helping teachers advance in professional knowledge and capacities.

SUMMARY

Ohio's SUTQ incorporates Staff Supports, Program Administration, and Staff Management into its Administrative and Leadership Practices domain. These components also are common across states: 88% of state QRIS also include Program Administration, Management, and Leadership components in standards for center-based care, while 82% include them in standards for home-based care. In comparison, only 65% of QRIS include staff supports in standards, and then only for center-based care. Similarly, 63% of QRIS include continuous quality improvement indicators in their standards for center-based care (and 51% include these standards for home-based care) (QRIS Compendium, 2016).

Exhibit 3. Buysse, Winton, and Rous (2009) Conceptual Framework for Professional Development



Staff Education and Professional Development

STAFF EDUCATION

Gomez et.al. (2015) and others (Yoshikawa et.al. 2013, for example) stress the importance of high quality teaching for positive child outcomes. However, despite agreement that high quality teaching is important, there are questions about what levels and types of education and experience are best aligned with high quality teaching and child progress.

All states incorporate professional education, training, and experience into QRIS requirements. All states require some form of professional training, education, and experience to become a lead teacher in a QRIS classroom. However, there is variation across states in the specific standards and requirements (e.g., two-year versus four-year degrees; total hours of professional development each year; Gomez et.al. 2015). Further, some authors have raised questions about the nature and level of qualifications that are important for child outcomes. Early et.al. (2007), for example, conducted an analysis of data from seven preschool studies that had comparable data. As regards the relation of a bachelor's degree to classroom quality, the authors found mixed results—some studies identified a positive association while others did not. As regards the association of a bachelor's degree with child outcomes, the authors again found mixed results, with some studies reporting an association of provider degree and child reading and math skills and other studies failing to find a significant association. The authors questioned whether the lack of clear and consistent associations between education and outcomes was related to how teachers are prepared, how teachers are supported in their classrooms, or market forces affecting recruitment and retention of the most qualified teachers. The authors suggested that high quality teachers are capable of producing the best outcomes—but that education alone may not be a sufficient indicator of high quality. Pianta et.al. (2016) echoed this finding:

100% of state QRIS include staff qualifications indicators in their standards.

The evidence on whether a teacher's degree and certification make a difference is murkier. For lead teachers, credible research supports the hypothesis that a bachelor's degree leads to higher-quality teaching, though it also supports that hypothesis that a BA doesn't ensure effective teaching. Retrospective analyses indicate that state pre-kindergarten program that show promising impacts on student learning in elementary school...all require teachers to have a BA, but this evidence doesn't prove a causal link.

Finally, Kelley & Camilli (2007) conducted a meta-analysis of the relation of teacher education to classroom quality and child outcomes. The authors examined 32 studies and found that child care quality improves with teacher education and that a four-year degreed teacher typically produces the best outcomes. The authors also noted that their calculations generated a relatively small effect size (.15) attached to a bachelor's degree but that in studies in which there was a large effect size, the large effect size was attached to teachers with a bachelor's degree. Like Early et.al. (2007), the authors concluded that it is important to study the actual teaching behaviors associated with the greatest child outcomes and determine if those practices only can be acquired through a four-year degree, or if they

could be taught or incubated through other methods with teachers that hold less than a four-year degree.

PROFESSIONAL DEVELOPMENT

It is important for early childhood educators to have access to and invest in regular professional development; many states require ongoing professional development as a component of their QRIS. Professional development has been identified as critical infrastructure for supporting and promoting high quality early education for the long run (as noted, for example, by Gallagher and Clifford, 2000 and Azzi-Lessing, 2009). Gomez et.al. (2015) identify three pathways for professional development, including access to higher education (i.e., a two-year, four-year, or graduate degree specific to early learning and development), credentialing based on demonstrated competencies, and ongoing professional development provided through technical assistance, onsite coaching or mentoring, periodic trainings and workshops, and Communities of Practice or professional associations. As the authors conclude: *“For many ECE educators, a combination of the three pathways...is needed to ensure that they are armed with the knowledge and skills necessary to work with young children and their families.”*

Gomez et.al. (2015) also note several challenges associated with the provision of professional development, especially as associated with state systems. These include low overall standards or requirements for entry-level teachers, inconsistency in the quality or rigor of ongoing professional development, and a failure to ensure that the material presented in all three pathways described above is consistent and aligned with state early learning and education expectations. The authors note equity issues, and voice concern that all early educators do not have equal access to preparation and quality professional development. As will be shown later in this report, directors from some of the participating sites in the current study agree that professional development is critical infrastructure for achieving and sustaining high quality, statewide.

A recent search for professional development in early childhood resulted in over 1000 articles, many of which focused on professional development for specific curricular areas such as language development, literacy, mathematics and science, or socio-emotional skills. Articles also address strategies for ensuring early educators can receive professional development. In addition to the onsite technical assistance or training opportunities noted above, a variety of video, web-based, or online opportunities are emerging with varying degrees of success (see for example, Early et.al. 2017; Kyzar et.al. 2014; Lee et.al. 2009; Pianta et.al. 2008; Stone-MacDonald & Douglass, 2014). The economic and temporal constraints that many providers experience will continue to drive the development of effective and efficient systems for ensuring the ongoing development of classroom staff.

SUMMARY

Ohio's SUTQ incorporates Staff Education and Professional Development into its Staff Qualifications and Professional Development domain, as do all other states with a QRIS. Specifically, Staff Qualifications (including education, training, and experience) are required in 100% of QRIS across the country and exist for both center- and home-based care.

Family and Community Partnerships

TRANSITIONS

The movement of young children from home or a preschool environment into kindergarten is an example of a transition; there is an emerging literature that documents important practices and strategies for this process (Early et.al. 2014; Hindman et.al. 2013; LoCasale-Crouch et al., 2008; Pianta et.al. 2001; Rous et.al. 2010).

However, transitions also occur on a more regular basis, a concept that is emphasized by the National Center on Parent, Family, and Community Engagement. As stated in their brief “Family Engagement in Transitions: Transition to Kindergarten”:

90% of state QRIS include family partnership indicators in their standards for child care centers.

Children experience many big and small transitions in their early years. Small transitions may include moving from playtime to cleanup, from hand washing to snack time, or from playing outdoors to coming back into the classroom. Big transitions might include moving from home to Early Head Start or Head Start, from being an only child to becoming a big brother or sister, or leaving Head Start to go to kindergarten.

Some transitions are required by law or statute. See for example requirements for transitioning children involved in early intervention services (Lillie & Vakil, 2002) or Head Start (Performance Standards 1302, Sub-Part G, Sections 1302.70, 1302.71, and 1302.72). What appears to be consistent best practice is the involvement of multiple, meaningful, caregivers in transition processes—parents, family, and educators (Brandes et.al. 2007; Early Head Start National Resource Center; Lillie & Vakil, 2002; Puccioni, 2015; Rous and Hallam, 2012).

Some authors have noted the paucity of research on within-program transitions (Rous & Hallam, 2012):

Transition within programs has received very little empirical attention. In particular, as children with disabilities are served in more community-based settings (e.g., child care, Head Start), issues such as staff turnover cause frequent caregiver transitions and frequent classroom changes within the same program, all of which have the potential to negatively impact children and families.

The same authors also recommend a greater focus in these areas, particularly for children with disabilities (Rous & Hallam, 2012). SUTQ addresses transitions through formal and informal mechanisms of transferring information about a child among caregivers. The study team’s review of transition practices across states suggests that approximately half (48%) of extant QRIS include transition practices in their standards.

COMMUNITY AND ENGAGEMENT

More prevalent across states are standards for community and engagement, or the different opportunities that sites provided for family and community communication, outreach, and engagement. The literature on transition practices cited above supports the involvement of parents and family in specific events such as a transition. Other studies support the ongoing engagement of parents and

43% of state QRIS include community involvement indicators in their QRIS standards for child care centers.

family in a child's early education, linking parental involvement to positive child outcomes (Ansari & Gershoff, 2016; Arnold et.al. 2008; Grindal et.al. 2016; Hayakawa et.al. 2016; Powell et.al. 2010) and parent outcomes (Ansari & Gershoff, 2016). In particular, a higher intensity of parental education at a child care site may be linked to stronger outcomes, especially for at risk or vulnerable children

(Grindal et.al. 2016).

Finally, community partnerships or the opportunity to collaborate with community agencies and sponsors can prove beneficial for both the program and the child. Selden et.al. (2006) examined collaboration in early learning and development and cited the linkages among high quality care, subsidies, and welfare reform as examples of how government funding can be used to drive cross-agency alignment and partnership. The authors also cited state and federal regulations that allow funding to be used in both private and public settings of examples of constructive collaboration. Family and community partnerships also are central to Head Start, as noted in the development of the Parent Family and Community Engagement (PFCE) Framework and performance standards tied to engagement. As noted by the National Center on Parent, Family, and Community Engagement, the PFCE is operationalized when *“staff and families collaborate with community, health, mental health, social service, and school partners to build peer networks, link families and children to needed services, and support successful transitions for children and families”* (2011). Examining this framework, there is an emphasis on how early education and child care programs reduce their own isolation so as to become more attuned and responsive to family needs and link families and children to community resources that may be of benefit.

SUMMARY

Ohio's SUTQ identifies Family and Community Partnerships as the fourth domain in which participating programs must document consistent strategies and best practices. This domain includes both Transitions and Community and Engagement components. Ninety percent (90%) of other state QRIS include family partnership in standards for center-based care (and 87% for home-based care) while 43% of QRIS include center-based standards related to community involvement (44% for home-based care). As noted earlier, the study team's review of standards across 44 states suggest that, across QRIS, 48% include standards specific to transition practices.

Group Size and Accreditation

GROUP SIZE AND RATIOS

Manning et.al. (2015) identify small group size (or low teacher-child ratios) as advantageous for children and linked to positive child outcomes such as social competence and academic skills. Small group size allows children to have more individual attention and support and states typically require that the youngest children (i.e., infants) have the best teacher-child ratios. Minimum ratios and group sizes are included in the licensing requirements in all states. However, more than half of states award additional quality points for having lower ratios or group sizes than required.

59% of states include group size components in their QRIS.

ACCREDITATION

Accreditation by a national accrediting body often is considered the apex of quality in an early childhood program. Winterbottom and Jones (2014) performed a direct examination of the difference in quality between accredited and non-accredited programs in Florida. Accredited programs had fewer violations of licensing standards and the authors concluded that accredited programs generally provided a higher quality environment for children. Although Whitebrook et.al. (1997) were critical of some sites that were accredited by the National Association for the Education of Young Children (NAEYC), the NAEYC documents the relation of accreditation to program quality and positive child outcomes (2009) and notes the revision and upgrading of the accreditation process.

55% of states include accreditation components in their QRIS.

Ohio SUTQ sites that have achieved a 3-star rating may advance to a 4- or 5-star rating, in part by earning points related to group size and accreditation. Thus, standards for group size and accreditation are not mandatory for participation in SUTQ. This is consistent with standards in other states: the study team found that 59% of 44 state QRIS that were reviewed included standards for group size and 55% of 44 states included standards for accreditation.

The study team's review of state systems also identified elements that are not currently present in SUTQ. One area that merits discussion is licensing standards, including health and safety or environment standards in early childhood programs. For some states, the achievement of or compliance with basic licensing and health and safety standards contributes to a site's star rating. In others, such as Ohio, these standards are foundational for participation in the QRIS—failure to comply, or the reporting of a violation, can be grounds for suspension or termination.¹³

Components that are emerging across states included provisions for special needs or inclusion children (58% of QRIS for center-based care and 62% of QRIS for home-based care) and attention to cultural or linguistic diversity (50% of QRIS for center-based care and 46% for home-based care). These

¹³ Specifically, licensing standards were incorporated into ratings in 80% (35 of 44) of states reviewed, while health and sanitation standards were found in 59% (26 of 44) of states reviewed.

SUTQ Validation Study Results

standards draw attention to the possibility that quality may be defined differently for different populations of children—specific attention to special learning needs or cultural sensitivity may help ensure that all children experience high quality, relevant for their needs. The recent development of quality scales for assessing inclusion classrooms is evidence of the growing importance of these questions (e.g., SpecialLink Early Childhood Inclusion Quality Scale and the Inclusive Classroom Profile).

Taken together, it is evident that SUTQ includes standards that are evidence-supported and common to other states' QRIS. These two conclusions support the content validity of SUTQ criteria.

Construct Validity

HIGHLIGHTS

Construct validity in the current study refers to the extent to which the five SUTQ star ratings differentiate between levels of quality. Specifically, the study team examined whether or not there appear to be meaningful differences in quality across star ratings. The key question for this phase of the project was:

To what extent do independent observations align with star rating?

If, for example, the study team found that independent assessments of classroom practices could differentiate between classrooms (based upon level of agreement between star rating and observed quality), the study team could conclude that the criteria and benchmarks for that star rating level were sufficient to capture a meaningful distinction in quality.

In completing this phase of the study, the study team conducted independent observations of classroom practices in participating sites, using:

- ERS (ECERS, ITERS, FCCERS);
- CLASS (the CIS was implemented with Type A and B homes); and
- ELLCO (CHELLO was implemented with Type A and B homes)¹⁴.

The study team also collected data from directors, teachers, and parents on practices for engaging and communicating with families, using the Family and Provider/Teacher Relationship Quality scale. Finally, the study team collected PAS data with a sub-set of centers (and BAS data with a sub-set of Type A and B homes).

Findings for Centers and Early Childhood Education Classrooms

The criteria that ODJFS- and ODE-licensed child care centers and Early Childhood Education/Preschool Special Education classrooms need to document or meet to achieve different star ratings are presented in Appendix B. The study team examined the extent to which these requirements translate into quality practices through several assessments of classroom practices. In essence, the study team examined whether a 3-, 4-, or 5-star rating could be associated with higher ratings on observed

HIGHLIGHTS

- Classrooms in sites with higher star ratings tend to demonstrate higher observed quality in practice.
- There are significant differences between classrooms in non-participating sites and classrooms in SUTQ, especially on observations of program environment (using the Environment Rating Scales, or ERS).
- There are fewer meaningful differences among classrooms when progressing by one level from 1-star to 2-star to 3-star, etc. This suggests that quality is accruing through gradual improvements in classrooms practices and interactions.

¹⁴ The team also was interested in examining quality of environment and classroom practices for children with special needs and interests (e.g., developmental delays or special learning needs), utilizing the SpecialLink scale. Too few of the sampled classrooms were inclusion classrooms, however, to be included in reporting.

classroom practices, compared to sites with 1- or 2-star ratings or sites that were not participating in SUTQ.

ENVIRONMENT RATING SCALES (ERS)

As noted earlier, the ERS commonly are used to assess quality in early education programs. Two versions of ERS assessments were used in centers and Early Childhood Education classrooms: the ITERS-R and the ECERS-3.

The ITERS-R contains seven subscales: 1) Space and Furnishings, 2) Personal Care Routines, 3) Listening and Talking, 4) Activities, 5) Interaction, 6) Program Structure, and 7) Parents and Staff. To ensure consistency with the ECERS-3, only the first six subscales were included in calculating a classroom's mean overall rating. Items in each subscale are rated on a seven-point scale, in which a score of "1" indicates "inadequate" and a score of "7" indicates "excellent." For the current study, scores of 1 or 2 were considered "Low" performing, scores of 3-5 were considered "Moderate" performing, and scores of 6 or 7, "High" performing.

The mean ITERS-R score was 3.9 on a seven-point scale (Table 11). In particular, there was a difference of more than 1 point between the mean overall ratings of classrooms in non-participating sites versus classrooms in SUTQ rated sites. As will be shown for ECERS-3 ratings, there appears to be an alignment between participation in SUTQ and higher center-based classroom quality.

The average ITERS-R score was 3.9.

Analysis of Variance indicates a statistically significant association of star rating and overall ITERS-R score ($F_{(5,89)} = 6.856, p < .000$), with specific differences between classrooms in non-participating sites and classrooms in any rating tier of SUTQ. This again indicates an association between participation in SUTQ and higher classroom quality.

ITERS-R subscales also were examined using Analysis of Variance and statistically significant relationships were found between star rating and average subscale scores, as follows:

- In Space and Furnishings, classrooms in 2- to 5-star rated sites showed statistically higher mean scores than classrooms in non-participating sites ($F_{(5,89)} = 7.145, p < .000$).
- In Listening and Talking, classrooms in 2-to 5-star sites showed statistically higher mean scores than classrooms in non-participating sites ($F_{(5,89)} = 6.233, p < .000$).
- In Activities, classrooms in 3- to 5-star rated sites showed statistically higher mean scores than classrooms in non-participating sites ($F_{(5,89)} = 7.536, p < .000$).
- In Interactions, classrooms in 1, and 3- to 5-star rated sites showed statistically higher scores than classrooms in non-participating sites ($F_{(5,89)} = 4.879, p < .001$).
- In Program Structure, classrooms in 3- to 5-star rated sites showed statistically higher scores than classrooms in non-participating sites ($F_{(5,89)} = 4.867, p < .001$).

Table 11. Agreement between center star rating and ITERS-R scores

		Non-rated (n=15)	1-star (n=18)	2-stars (n=18)	3-stars (n=18)	4-stars (n=14)	5-stars (n=12)
Space and Furnishings	Mean	2.6	3.9	4.0	4.9	4.3	5.2
	SD	1.2	1.4	1.3	1.5	1.3	1.2
Personal Care Routines	Mean	2.3	3.6	3.7	4.2	4.2	4.3
	SD	1.5	1.98	2.1	1.6	1.98	2.1
Listening and Talking	Mean	2.0	3.6	3.8	4.4	4.7	5.0
	SD	1.1	1.7	1.8	1.5	1.7	1.9
Activities	Mean	1.9	3.2	3.0	4.1	4.1	4.3
	SD	.8	1.4	1.4	1.3	1.2	1.3
Interactions	Mean	2.8	5.0	4.5	5.1	5.4	5.6
	SD	1.95	1.9	1.8	1.8	1.2	1.8
Program Structure	Mean	2.3	3.7	3.6	4.7	4.5	5.4
	SD	1.6	1.8	2.3	1.7	1.96	1.8
Mean ITERS Score	Mean	2.3	3.8	3.7	4.5	4.4	4.8
	SD	1.1	1.3	1.5	1.3	1.2	1.4

Data Source: Infant/Toddler Environment Rating Scale

The ECERS-3 contains six subscales: (1) Space and Furnishings, (2) Personal Care Routines, (3) Language and Literacy, (4) Learning Activities, (5) Interaction, and (6) Program Structure. Items were scored and interpreted the same as with the ITERS-R.

Overall, the mean ECERS-3 score for Early Childhood Education classrooms (n=95) was 3.6 on a seven-point scale. Several findings are of interest, as shown in Table 12¹⁵. In particular, and similar to the findings for the ITERS-R, there is a difference of more than 1 point between the mean overall rating of classrooms in non-participating sites versus classrooms in SUTQ rated sites. Thus, for the sites sampled in the study, participation in SUTQ was aligned with higher quality classroom practices.

The average ECERS-3 score was 3.6.

Analysis of Variance, in which ECERS-3 scores and star rating were included as variables, indicate a statistically significant association with score and star rating ($F_{(5,89)} = 4.406, p < .001$). Follow-up analyses revealed significant differences between classrooms in non-participating sites and classrooms in 3-, 4-, and 5-star rated sites.

Each ECERS-3 subscale also was examined, using Analysis of Variance. There were significant differences in multiple subscales:

- In Space and Furnishings, classrooms in 3-, 4-, and 5-star rates sites showed statistically higher scores than classrooms in non-participating sites ($F_{(5,89)} = 5.902, p < .000$).

¹⁵ Additional findings are presented in Appendix F.

SUTQ Validation Study Results

- In Personal Care Routines, classrooms in 5-star rated sites showed statistically higher scores than classrooms in non-participating sites ($F_{(5,89)} = 3.4, p < .007$).
- In Language and Literacy, classrooms in 4- and 5-star rated sites showed statistically higher scores than classrooms in non-participating sites ($F_{(5,89)} = 4.286, p < .002$).
- In Interactions, classrooms in 4- and 5-star rated sites showed statistically higher scores than classrooms in non-participating sites ($F_{(5,89)} = 3.71, p < .004$).
- In Program Structure, classrooms in 4- and 5-star rated sites showed statistically higher scores than classrooms in non-participating sites ($F_{(5,89)} = 4.751, p < .001$).

Table 12. Agreement between center star rating and ECERS—3 scores

		Non-rated (n=14)	1-star (n=15)	2-stars (n=13)	3-stars (n=17)	4-stars (n=19)	5-stars (n=17)
Space and Furnishings	Mean	2.8	4.1	3.6	4.6	4.6	4.8
	SD	1.3	1.2	1.3	1.0	1.2	1.3
Personal Care Routines	Mean	2.9	3.9	3.5	4.3	4.5	5.0
	SD	1.5	1.7	1.7	1.4	1.7	1.5
Language and Literacy	Mean	2.2	3.5	3.1	3.5	4.0	4.4
	SD	.9	1.7	1.7	1.4	1.3	1.5
Learning Activities	Mean	1.9	2.7	2.5	2.7	3.0	3.0
	SD	.7	1.3	1.1	1.2	1.6	1.5
Interactions	Mean	2.7	3.96	4.2	4.4	4.5	5.2
	SD	1.5	1.8	1.8	1.6	1.9	1.3
Program Structure	Mean	2.4	3.7	3.8	3.9	4.2	5.2
	SD	1.2	1.9	1.8	1.6	1.7	1.2
Mean ECERS-3 Score	Mean	2.4	3.5	3.3	3.7	3.97	4.3
	SD	.97	1.4	1.3	1.1	1.4	1.3

Data source: ECERS-3

Using both the ECERS-3 and the ITERS-R, classroom quality was grouped into three categories: low, moderate, and high. The percentage of classrooms in each category are presented in Table 13. As can be seen, over 60% of classrooms were in the moderate level of quality on both the ECERS-3 and the ITERS-R, suggesting that relatively few classrooms scored at the highest levels possible. Thus, although classrooms in the higher-rated (3-, 4-, and 5-star rated) sites tended to have the highest relative performance, there still is room for improvement, in the consistent achievement of observed scores in the high range of performance on these two instruments.

Table 13. Percent of classrooms with low, moderate, and high ERS scores

	ECERS-3	ITERS-R
Low (scores of 1 or 2)	21%	22%
Moderate (scores of 3 to 5)	66%	60%
High (scores of 6 or 7)	13%	18%

Data Source: ECERS-3; ITERS-R

The findings from the current study are similar to those found in other studies. In their reviews, Karoly (2014) and Karoly et al. (2016) found that in 12 states with completed and published validation study reports, ERS were used as measures of quality. Although there was some variation in ERS scores within rating levels, validation studies in 11 of the 12 states found that ERS scores were positively correlated with QRIS ratings and that higher scores were associated with higher quality ratings. The only exception was in Minnesota where some programs received no formal assessment but were rated in the top tier (Tout et al., 2011).

CLASSROOM ASSESSMENT SCORING SYSTEM (CLASS)

The CLASS commonly is used to assess the quality of teacher-child interactions, but at the current time only is available for center-based classrooms. Three versions of the CLASS were used in the study: the CLASS PreK, the CLASS Toddler, and the CLASS Infant.

CLASS Infant

The CLASS Infant captures teacher-child interactions as a single domain, Responsive Caregiving, divided into four dimensions: Relational Climate, Teacher Sensitivity, Facilitated Exploration, and Early Language Support.

The average Responsive Caregiving score was 4.6.

Similar to the ERS, CLASS items in each subscale are rated on a seven-point scale. In the current study, scores of 1 or 2 were considered “Low” performance, scores of 3-5 were considered “Moderate” performance, and scores of 6 or 7 were considered “High” performance.

A total of 43 infant classrooms participating in the current study received the CLASS Infant; the mean Responsive Caregiving rating across all classrooms was 4.6 on a seven-point scale. Because of small sample sizes in some rating categories, classrooms were grouped into non-participating, lower rated (1- to 2-star rated) and higher rated (3- to 5-star rated) categories. As regards the alignment of star rating and CLASS Infant scores, the findings presented in Table 14 indicate that the lowest ratings occurred in sites that are not participating in SUTQ. Of the sites that are star rated, there is a general trend for CLASS Infant scores to increase along with star rating, with lower scores occurring in classrooms in 1- and 2-star rated sites and higher scores occurring in classrooms in 3- to 5-star rated sites.

There were statistically significant differences among these categories ($F_{(2,40)} = 9.921, p < .000$) with significant differences between classrooms in non-participating sites and classrooms in lower and higher rated sites.

Table 14. Agreement between center star rating and CLASS-Infant scores

		Non-rated (n=7)	1-to 2-star classrooms (n=18)	3-to 5-star classrooms (n=18)
Responsive Caregiving	Mean	2.8	4.6	5.4
	SD	.7	1.6	1.3

Data Source: CLASS Infant

As shown in Table 15, just over half the classrooms assessed demonstrated scores in the moderate range, with almost 40% of remaining classrooms demonstrating scores in the high range. These findings suggest that, overall, the quality of Infant classroom teacher-child interactions is relatively high.

Table 15. Percent of classrooms with low, moderate, and high CLASS-Infant scores

	Responsive Caregiving
Low (scores of 1 or 2)	9%
Moderate (scores of 3 to 5)	51%
High (scores of 6 or 7)	40%

Data Source: CLASS Infant

It is important to address the quality of classroom environments across age groupings. In states such as North Carolina, for example, there is evidence that infants and toddlers experience lower overall quality than older children. Similarly, providers and educators in infant and toddler classrooms tend to have fewer educational qualifications or achievements, compared to teachers in Early Childhood Education classrooms (North Carolina Performance Based Incentive System findings, 2015). Thus, the relatively high scores on assessments such as the CLASS Infant are especially encouraging, as they suggest high quality exists as a uniform concept, across participating sites.

CLASS Toddler

The CLASS Toddler contains two subscales: Emotional and Behavioral Support and Engaged

The average Emotional and Behavioral Support score was 5.3. The average Engaged Support for Learning score was 3.1.

Support for Learning. Emotional and Behavioral Support captures factors such as positive or negative climate, teacher sensitivity, regard for child perspectives, and behavior guidance. The Engaged Support for Learning domain encompasses the facilitation of learning and development, quality of feedback given to children, and teacher use of language modeling.

Items in each subscale are rated on a seven-point scale wherein a score of 1 or 2 is considered “Low” performance, scores of 3-5 are considered “Moderate” performance, and scores of 6 or 7 are considered “High” performance.

A total of 52 toddler classrooms received the CLASS Toddler. The mean Emotional and Behavioral Support score was 5.3, while the mean Engaged Support for Learning score was 3.1. It is worth noting the differences between mean scores on these two subscales. The findings from the current study suggest that teachers may struggle with their facilitation of learning and development, and in particular with language modeling.

As with the CLASS Infant, due to smaller sample size in some categories, classrooms were combined into non-participating, lower rated, and higher rated categories (Table 16). The differences among mean scores were not statistically significant, however.

Table 16. Agreement between center star rating and CLASS-Toddler scores

		Non-rated (n=7)	1-2-star classrooms (n=19)	3-5-star classrooms (n=26)
Emotional and Behavioral Support	Mean	4.8	5.3	5.4
	SD	1.3	.9	.97
Engaged Support for Learning	Mean	2.9	3.3	3.0
	SD	1.3	1.4	1.1

Data Source: CLASS Toddler

The mean subscale scores observed in Table 16 are further illustrated in Table 17, which shows all Toddler classrooms scoring either in the moderate or high range for Emotional and Behavioral Support (note the higher mean scores in this subscale, presented in Table 16), while only 4% of classrooms scored in the High range for Engaged Support for Learning. These findings suggest the subscale Engaged Support for Learning is an area for support and assistance to Toddler classroom teachers.

Table 17. Percent of classrooms with low, moderate, and high CLASS Toddler scores

	Emotional and Behavioral Support	Engaged Support for Learning
Low (scores of 1 or 2)	0%	40%
Moderate (scores of 3 to 5)	56%	56%
High (scores of 6 or 7)	44%	5%

Data Source: CLASS Toddler

CLASS PreK

The CLASS PreK contains three subscales: Emotional Support, Classroom Organization, and Instructional Support. The Emotional Support domain captures positive and negative climate, teacher sensitivity, and regard for child perspectives. The Classroom Organization scale captures behavior management, productivity, and instructional learning formats. Finally, the Instructional Support section captures concept development, quality of feedback given to children, and language modeling.

The average Emotional Support score was 5.5. The average Classroom Organization score was 4.8. The average Instructional Support score was 2.8.

Items in each subscale are rated on a seven-point scale wherein a score of 1 or 2 is considered “Low” performance, scores of 3-5 are considered “Moderate” performance, and scores of 6 or 7 are considered “High” performance.

A total of 96 3- and 4-year old classrooms received a CLASS PreK; mean scores disaggregated by star rating are presented in Table 18. The mean Emotional Support score was 5.5 while the mean Classroom Organization score was 4.8 and the mean Instructional Support score was 2.8.

When assessed by star rating, there is a trend for classrooms in star rated sites to have higher mean scores than classrooms in sites that are not participating in SUTQ. Further, there is a trend for higher rated sites (3-5 stars) to have higher mean ratings, compared to 1- and 2-star sites.

There were statistically significant differences among sites on the Emotional Support subscale ($F_{(5,90)} = 2.733, p < .024$). Specifically, on the Emotional Support subscale there were significant differences between classrooms in non-participating sites and classrooms in 5-star rated sites.

Table 18. Agreement between center star rating and CLASS PreK scores

		Non-rated (n=14)	1-star (n=16)	2-stars (n=14)	3-stars (n=17)	4-stars (n=18)	5-stars (n=17)
Emotional Support	Mean	4.9	5.1	5.3	5.7	5.7	5.9
	SD	1.0	1.2	1.0	.9	.8	.6
Classroom Organization	Mean	4.1	4.5	4.4	5.0	5.2	5.2
	SD	1.5	1.3	1.5	1.4	1.2	1.2
Instructional Support	Mean	2.3	2.5	2.6	2.9	3.1	3.2
	SD	1.4	1.0	1.4	.9	1.0	1.1

Data Source: CLASS PreK

The lower overall scores for Instructional Support also are observed when the percentages of all classrooms scoring in the low, moderate, and high ranges are calculated. As shown in Table 19, classrooms scored relatively well in Emotional Support and Classroom Organization but struggled to achieve higher scores in Instructional Support.

Table 19. Percent of classrooms with low, moderate, and high CLASS PreK scores

	Emotional Support	Classroom Organization	Instructional Support
Low (scores of 1 or 2)	1%	7%	44%
Moderate (scores of 3 to 5)	45%	57%	54%
High (scores of 6 or 7)	54%	35%	2%

Data Source: CLASS PreK

Three other states (Delaware, Pennsylvania, and Minnesota) used the CLASS (Pianta et al., 2008) as a measure of quality in their validation studies . In Delaware, although changes were modest and not always statistically significant, increases in CLASS scores were associated with rising QRIS rating levels (Karoly et al, 2016) and in Pennsylvania, scores were higher for higher rated classrooms on all CLASS subscales (Sirinides, 2010). In Minnesota, however, there were no significant differences across rating tiers (Tout et al., 2011).

In light of findings from previous studies, the results herein are not surprising. A summary of research on the CLASS tool completed by the Center for Advanced Study of Teaching and Learning (2010) reported that students in Early Childhood Education through grade 5 experience moderate to high levels of effective interactions for emotional support and classroom organization, but very low levels of instructional support. Further, a 2007 study conducted by Pianta et al. showed that less than 10% of 1,000 students followed through elementary school had access to classrooms that consistently scored in the mid to upper range for effective interactions. Thus, it is not surprising that participants in the current study also generated lower scores in Instructional Support.

EARLY LANGUAGE AND LITERACY CLASSROOM OBSERVATION (ELLCO)

The ELLCO is validated for use in Early Childhood Education (3- and 4-year old) classrooms and contains three tools [(1) Literacy Environment Checklist; (2) Classroom rating of 14 dimensions of literacy; and (3) Literacy Activities Rating Scale] and two subscales: (1) General Classroom Environment (derived from a combination of items related to classroom structure, organization, contents, management, and climate) and (2) Language and Literacy (derived from items related to language environment and discourse climate, presence of books, approaches to book reading and writing, and curriculum integration). Both are measured on a five-point scale, wherein a “1” represents “deficient” practice and a “5” represents “exemplary” practice.

Ninety-six Early Childhood Education classrooms participating in this study received the ELLCO.

The average General Classroom Environment score was 3.4. The average Language and Literacy score was 2.4.

The mean rating for General Classroom Environment was 3.4 while the mean rating for Language and Literacy was 2.4. ELLCO scores were disaggregated by star rating, as shown in Table 20. As can be seen, classrooms in sites that do not participate in SUTQ tend to have the lowest mean scores. The highest scores in both subscales occurred in classrooms in 4- and 5-star-rated sites.

Analysis of Variance indicates statistically significant differences among sites by star rating on the General Classroom Environment subscale ($F_{(5,90)} = 3.248, p < .10$). Specifically, there were differences between classrooms in non-participating sites and classrooms in 4- and 5-star-rated sites.

Table 20. Agreement between center star rating and ELLCO scores

		Non-rated (n=14)	1-star (n=16)	2-stars (n=14)	3-stars (n=17)	4-stars (n=18)	5-stars (n=17)
General Classroom Environment	Mean	2.8	3.4	3.3	3.5	3.7	3.8
	SD	.8	.8	.7	.8	.8	.6
Language and Literacy	Mean	2.1	2.4	2.2	2.4	2.7	2.7
	SD	.6	.7	.8	.5	.7	.8

Data Source: Early Language & Literacy Classroom Observation

The trends observed in Table 20 are again expressed in Table 21, below. As can be seen, more than half of all classrooms assessed scored in the high range in the General Classroom Environment items. Only 7% of classrooms, in comparison, scored in the same range for Language and Literacy items.

Table 21. Percent of classrooms with low, moderate, and high ELLCO scores

	General Classroom Environment	Language and Literacy
Low (scores of 1 or 2)	18%	51%
Moderate (score of 3)	27%	42%
High (scores of 4 or 5)	55%	7%

Data Source: Early Language & Literacy Classroom Observation

While there is little documentation of the relation of ELLCO scores to QRIS ratings across states, there is evidence that performance on the ELLCO is associated with child outcomes. Jackson et.al.

(2006), for example, linked classroom improvements (as captured on the ELLCO) to child literacy gains. Cunningham (2008) found that *“Students who came from preschool classrooms rated as deficient in their support of language and literacy had the most negative attitudes”*, with regard to reading and writing. Cunningham concluded that *“Although excellent formal reading instruction can influence success in literacy even for high-risk readers, substantial efforts to provide high-quality early literacy environments could provide a major prevention effort for later reading difficulties rather than focusing on remediation after a reading problem has developed.”* Further, there is evidence that a focus on language and literacy supports and a literacy-rich environment can yield results in child appreciation for, attitudes towards, and engagement in literacy-rich behaviors. To the extent that teacher-child interactions are a major predictor of quality and child outcomes (Pianta et.al. 2016), it will be important to focus support on teacher language and literacy practices, in addition to the availability of language and literacy materials and environments.

FAMILY PROVIDER/TEACHER RELATIONSHIP QUALITY SCALE (FPTRQ)

The FPTRQ was developed to assess the strength and quality of parent-teacher engagement and relationships, with a focus on how well teachers facilitate meaningful relationships with families. Three versions of the FPTRQ were used in the study: Director, Teacher, and Parent.

Director Responses

The Director’s version of the FPTRQ contains four subscales: Environment and Policy Checklist, Communication Systems, Information about Resources, and Referrals.

- The Environment and Policy Checklist captures concepts such as the welcoming nature of the site, the availability of culturally-diverse information, and site strategies for providing parenting information. Seventeen items from the assessment are incorporated into this subscale, and the total possible range of scores is 0 to 17.
- The Communication Systems subscale addresses strategies for communicating with families. There are nine items in this subscale and the total possible range of scores is 0 to 9.
- The Information about Resources subscale captures the nature of information made available to families. There are 12 items in this subscale and the total possible range of scores is 0 to 12.
- There are 5 items contained in the Referrals scale, which addresses whether or not programs provide referrals for services such as health screenings or developmental assessments.

Directors reported family engagement practices on par with practices used by the general sample of directors across the country.

The mean scores identified during the instrument’s development provide some guidance for interpreting the scores. To wit, the Environment and Policy Checklist mean score, representing center-based directors, was 13.2, with a range of responses from 6 to 17 (Kim et.al. 2014). (No mean scores were reported for Communications Systems, Information about Resources, or Referrals.) Thus, scores at or above 13.2 in the current study suggest family engagement and communication practices at or above “typical.”

SUTQ Validation Study Results

The mean Environment and Policy checklist score for the current study was 13.2 (the mean score reported by Kim et.al. 2014; see Appendix F for more details), suggesting that, on average, participating directors were performing at a “typical” level, compared to the general sample of directors. The mean Communication Systems score was 7.9—no sample-based mean score was available but it is worth noting that the total range for this subscale is 0 to 9 points. Thus, directors in the current study reported behaviors at the high end of the scale.

The mean score for Information about Resources was 5.5. While no mean sample score was available for comparison, this subscale has a range of 0 to 12 points. Thus, a mean score of 5.5 suggests that directors are, on average, not making a full or comprehensive bank of resources available for parent’s information needs. Finally, the mean score on Referrals was 2.5, at the mid-point of the five-point scale.

Because of low sample sizes on some star rating categories, sites were combined into non-participating, lower rated, and higher rated categories. As can be seen in Table 22, mean scores were relatively high for the Environment and Policy Checklist and Communications Systems subscales, when disaggregated by site rating. Directors from higher rated sites tended to report a higher level of practices or supports for parents, while directors from sites that are not participating in SUTQ tended to report the lowest level of practices or supports (with the exception of Referrals, Table 22).

Sites with higher star ratings reported higher levels of family partnership practices.

Independent samples t-test analyses were conducted to examine differences between lower and higher rated sites. There were statistically significant differences in the Environment and Policy and Referral subscales. As regards the Environment and Policy subscale, higher rated sites had statistically significant higher scores ($t_{(26.259)} = -3.235, p < .003$). The same pattern existed for the Referrals subscale ($t_{(62)} = -2.941, p < .005$).

Table 22. Agreement between lower and higher star rating and Director FPTRQ scores

		Non-rated	Lower Rated (1-to 2-stars)	Higher Rated (3-to 5-stars)
Environment and Policy Checklist; range 0-17	Mean	10.3	11.9	14.4
	SD	2.3	2.95	1.9
	n	6	19	35
Communication Systems; range 0-9	Mean	6.7	7.7	8.0
	SD	2.1	1.1	1.0
	n	6	18	38
Information about Resources; range 0-12	Mean	3.4	4.7	6.4
	SD	2.97	3.3	3.6
	n	8	22	38
Referrals; range 0-5	Mean	2.1	1.8	3.0
	SD	1.8	1.7	1.6
	n	8	23	41

Data Source: FPTRQ-Director Measure

Teacher Responses

In the current study, the mean Knowledge subscale score was 30.4. It is worth noting that the mean score for center-based programs reported in the FPTRQ’s User’s Guide is 33.3. Thus, participating

teachers, on average, reported lower scores for this subscale than the “typical” teacher, suggesting a lower level of knowledge about individual families and their circumstances, compared to the sample used to develop the instrument.

The mean Practices score was 72.8 (while the mean sample-score reported in the User’s Guide is 77.6), again suggesting a lower, on average, level of practice by participating teachers. Finally, the mean

Teacher reports of family partnership practices tended to score below the sample mean.

Attitudes subscale score was 54.9 (the mean sample-score was 54.4). On this subscale, participating teachers appear to be on par with the general, or “typical”, sample used to develop the instrument.

Table 23 presents mean scores disaggregated by star rating. There was a general trend for teachers in higher rated sites to report a higher investment in family outreach and engagement, with a few exceptions. Analysis of Variance indicates statistically significant differences among teachers on the Practices subscale ($F_{(5,224)} = 2.295, p < .046$), although follow-up tests did not reveal distinctive differences among teachers, by star rating.

Table 23. Agreement between star rating and Teacher FPTRQ scores

		Non-Participants	1-Star	2-Stars	3-Stars	4-Stars	5-Stars
Knowledge	Mean	29.1	30.5	29.2	31.5	30.6	31.4
	SD	6.3	6.8	5.4	7.5	6.9	7.4
	n	35	42	40	46	51	39
Practices	Mean	68.5	74.5	69.4	75.5	74.3	73.4
	SD	12.2	11.1	8.89	11.6	11.97	12.8
	n	32	38	38	38	46	38
Attitudes	Mean	55.2	54.8	54.3	55	54.7	55.3
	SD	4.6	4.7	4.9	4.3	4.2	3.97
	n	34	43	33	44	48	40

Data Source: FPTRQ-Teacher Measure

Parent Responses

The Parent version of the FPTRQ contains three subscales: Knowledge, Practices, and Attitudes. The Knowledge subscale addresses a parent’s comfort level with sharing family-specific knowledge with a site. There are 15 items in this subscale and total score ranges from 15 to 60.

The Practices subscale addresses four constructs: Collaboration, Responsiveness, Communication, and Family-Focused Concern. There are 33 items in this subscale and the total range of scores is 33 to 132.

The Attitudes subscale addresses three constructs: Commitment, Understanding Context, and Respect. There are 18 items in this subscale and the total range of scores is 18 to 72.

Parent reports of family partnership practices tended to score below the sample mean.

As with the teacher and director measures, mean scores for the sample that was used to develop the instrument are available to help interpret the findings. In the current study, the mean score for parents on the Knowledge subscale was 51.9 (compared to a mean sample score of 52.6, cited in the User’s Manual). The mean score for the Practices subscale was 100.9 (compared to a mean score of

109.4 cited in the User’s Manual) and the mean score for the Attitudes subscale was 65.2 (compared to a mean score of 67.7 cited in the User’s Manual). Thus, participating parents tended to report lower levels of practice than the sample used to develop the tool.

Disaggregation by participation and star rating (Table 24) indicates that parents from sites that do not participate in SUTQ reported lower scores for the Knowledge and Practices subscales. There were statistically significant differences among responses on the Knowledge ($F_{(5, 63.844)} = 2.684, p < .029$) and Practices ($F_{(5, 51.119)} = 2.646, p < .033$) subscales, with specific differences between parents at non-participating and 5-star rated sites (Knowledge subscale) and parents at non-participating and 3-star rated sites (Practices subscales).

Table 24. Agreement between star rating and Parent FPTRQ scores

		Non-Participants	1-Star	2-Stars	3-Stars	4-Stars	5-Stars
Knowledge	Mean	47.9	52.3	52.4	55.1	50.8	55.1
	SD	12.6	4.8	7.4	7.4	10.4	4.97
	n	33	16	41	15	39	29
Practices	Mean	92.4	102.6	96.4	112.5	103.8	105.2
	SD	19.8	14.3	20.9	20.8	16.3	17.9
	n	28	11	35	15	32	25
Attitudes	Mean	64.9	64.8	65.6	66.9	64.4	65.7
	SD	5.4	6.2	4.2	5.4	6.4	5.6
	n	31	16	38	15	38	29

Data Source: FPTRQ-Parent Measure

As discussed earlier, parent engagement in and support for their child’s preschool learning experience can have positive child outcomes (Ansari & Gershoff, 2016; Arnold et.al. 2008; Hayakawa et.al. 2013; Powell, et.al. 2010). It is important to examine, therefore, not just the willingness to engage parents but also the specific tools and strategies used to work with parents. Results from the FPTRQ Scale—Teacher and Parent Measures suggest that participating teachers function below sample-based “typical” practices in these regards. Further, there appear to be differences in practices that are aligned with quality, wherein teachers and parents at higher rated sites report a higher level of practice and supports. Family partnerships are a popular standard in state QRIS, with 90% of QRIS containing center-based standards and 87% containing standards for home-based care. However, results from the current study suggest that support, coaching, or assistance may be helpful in improving the consistent use of high quality practices.

Importance of Education

Many states have placed a premium on teacher education, encouraging early educators to achieve at least a two-year degree in early childhood education or a related field. It stands to reason that teachers with higher education should be affiliated with higher quality classrooms and educational interactions. However, authors such as Early et.al.(2006) have failed to find significant and consistent relationships between education level, fields of study, or years of experience and child care quality and child academic outcomes. This stated, it is important to note that many studies are exploring the differences between a two-year and four-year degree. There is near universal agreement that a two-

SUTQ Validation Study Results

year degree is critical for providing high quality care and education. This is consistent with staff guidelines and requirements for early educators, and especially lead teachers, to hold a two-year degree in Early Childhood Education (or a related field) and to receive ongoing continuing education and professional development. These requirements also are a necessary benchmark in the “professionalization” of the early learning and development field, as is the development of fields of study and degree options at two- and four-year Institutes of Higher Education.

The study team explored the importance of education by collecting basic education data for teachers from non-participating and SUTQ sites; the data were reported by teachers and site directors. The study team compiled a dataset of 154 teachers, linked to classroom-based observations such as the ERS, CLASS, and ELLCO. Of the 154, 25 (16%) reported having only high school education, 34 (22%) reported having some college, 40 (26%) reported having a two-year degree, and 55 (36%) reported having a four-year degree or higher¹⁶.

Of the teachers who reported having high school education only, 16% were employed in sites that were not participating in SUTQ, while 56% were employed in 1- or 2-star rated sites. Of the teachers who reported having some college education, 15% were employed in sites that were not participating in SUTQ while 53% were employed in 1- or 2-star rated sites. Of the teachers who reported having a two-year degree, 70% were employed in 3-star or higher sites; of the teachers who reported having a four-year degree or higher, 64% were employed in 3-star or higher sites.

Table 25 presents education and experience information, disaggregated by star rating. As can be seen, higher percentages of teachers with two-year or higher degrees are employed in 3-star or higher sites. The distributions were tested with a Chi-square test of independence and were found to be statistically significant (Pearson Chi-square, $p < .000$). This finding is not surprising given the requirement for higher rated sites to have higher proportions of staff at Career Pathways Level 3 or higher—which can be achieved through achievement of a college degree.

Table 25. Education, disaggregated by star rating and SUTQ participation

	Non-Participating Sites (n=21)	1-star (n=28)	2-stars (n=24)	3-stars (n=28)	4-stars (n=28)	5-stars (n=25)
Percent of teachers with less than a college degree	42.9%	67.9%	54.2%	32.1%	25%	8%
Percent of teachers with a two year or four year college degree or higher	57.1%	32.1%	45.8%	67.9%	75%	92%
Mean years of experience	13 (SD=8.3; n=8)	12.8 (SD=8.3; n=18)	8.8 (SD=7.7; n=15)	8.8 (SD=5.97; n=17)	9.0 (SD=8.1; n=13)	12.8 (SD=8.1; n=21)

¹⁶ Note, because these data are self-reported, it will be important to confirm any association between education and classroom quality using verified data.

SUTQ Validation Study Results

The importance of educational attainment, degree, and years of experience were further explored by aligning education with scores on classroom observation. In brief, educational attainment was plotted against Low, Moderate, and High performing scores in classroom observations, displayed in Table 26, Table 27 and Table 28. All distributions were tested using a Chi-square test of independence. None of the findings were statistically significant.

Degree attainment was not significantly associated with observed classroom quality.

Table 26. Alignment of teacher education with low, moderate, and high ERS Scores

	ECERS			ITERS	
	Less than a college degree	Two year degree or higher		Less than a college degree	Two year degree or higher
Low Performing (n=14)	29%	71%	Low Performing (n=15)	53%	47%
Moderate (n=51)	29%	71%	Moderate (n=47)	47%	53%
High Performing (n=12)	33%	67%	High Performing (n=16)	44%	56%

Data Sources: FPTRQ-Teacher Measure; Site Education Profile; ECERS-3; ITERS-R

Table 27. Alignment of teacher education with low, moderate, and high CLASS PreK scores

	Emotional Support			Classroom Organization			Instructional Support	
	Less than a college degree	Two year degree or higher		Less than a college degree	Two year degree or higher		Less than a college degree	Two year degree or higher
Low Performing (n=1)	Sample too small		Low Performing (n=7)	57%	43%	Low Performing (n=34)	35%	65%
Moderate (n=35)	40%	60%	Moderate (n=45)	33%	67%	Moderate (n=44)	70%	70%
High Performing (n=43)	23%	77%	High Performing (n=27)	22%	78%	High Performing (n=1)	Sample too small	

Data Sources: FPTRQ-Teacher Measure; Site Education Profile; CLASS PreK

Table 28. Alignment of teacher education with low, moderate, and high ELLCO Scores

	General Classroom Observation			Language and Literacy	
	Less than a college degree	Two year degree or higher		Less than a college degree	Two year degree or higher
Low Performing (n=14)	43%	57%	Low Performing (n=39)	38%	62%
Moderate (n=19)	37%	63%	Moderate (n=34)	26%	74%
High Performing (n=46)	26%	74%	High Performing (n=6)	17%	83%

Data Sources: FPTRQ-Teacher Measure; Site Education Profile; ELLCO

The current study found no clear alignment between degree attainment and scores on classroom observations. Teachers and providers in high performing classrooms tend to have two-year or higher degrees. However, the same is true in some low and moderately performing classrooms. It is important to note that these findings are based on self-reported data by participating directors and teachers. Moving forward, it will be helpful to re-examine any associations between education and classroom quality using verified data (i.e., educational data that can be cross-checked or cross-referenced with documentation). If the suggested findings can be confirmed, the findings underscore the need to support providers and teachers in their translation of knowledge gained through degree attainment into actual classroom practice, through the ongoing provision of coaching, mentoring, training, and technical assistance.

Administrative Practices

A sub-set of 14 centers volunteered to conduct the PAS assessment. This assessment targets administrative, leadership, and management practices at child care sites.

As with the ERS and the CLASS, possible scores on each item range from 1 to 7, wherein a score of 7 is the highest possible score. Subscales include Human Resources Development, Personnel Cost and Allocation, Center Operations, Child Assessment, Fiscal Management, Program Planning and Evaluation, Family Partnerships, Marketing and Public Relations, Technology, and Staff Qualifications. | The mean PAS score was 3.4

The mean overall score across all 14 sites was 3.4, which is below the mid-point on the seven-point scale. Because of small sample sizes, sites were grouped into lower and higher rated sites, in which lower rated sites include 1- and 2-star rated sites and higher rated sites include 3- to 5-star rated sites. Mean scores for each grouping are presented in Table 29.

Table 29. Agreement between star rating and administrative practices

		Lower Rated (1- to 2-stars; n=6)	Higher Rated (3- to 5-stars; n=7)
Human Resources Development	Mean	2.8	4.5
	SD	2.1	.8
Personnel Cost and Allocation	Mean	1.9	2.5
	SD	1.0	.97
Center Operations	Mean	3.1	3.8
	SD	1.1	1.2
Child Assessment	Mean	4.2	5.8
	SD	2.1	1.1
Fiscal Management	Mean	3.1	2.9
	SD	2.9	1.98
Program Planning and Evaluation	Mean	1.8	4.4
	SD	1.4	.5
Family Partnerships	Mean	5.3	4.9
	SD	1.5	1.7
Marketing and Public Relations	Mean	3.5	4.1
	SD	.8	1.9

SUTQ Validation Study Results

		Lower Rated (1- to 2-stars; n=6)	Higher Rated (3- to 5-stars; n=7)
Technology	Mean	4.6	5.9
	SD	1.4	.9
Staff Qualifications	Mean	2.3	2.5
	SD	1.2	1.2
Overall Score	Mean	3.1	4.0
	SD	.97	.8

Data source: Program Administrative Scale

Program administration and management practices are found in many state QRIS. However, there are relatively few independent studies that quantify the impact of program administration on child outcomes. Heinemeier and Leonard (2013) explored the factors that help child care programs maintain high quality and found multiple administrative and management practices that can contribute to sustainable quality and long-term viability. These factors include the director (a) acting as a champion for quality and translating this philosophy into a working agenda for staff; (b) maintaining a results-based orientation that focused on child development and outcomes; (c) using strategic financing; and (d) remaining open to adapting the program to changing knowledge and expectations for high quality care and education. The findings from the current study suggest that higher rated sites invest in stronger management and leadership techniques than lower rated ones. However, many of the mean scores also suggest that there is much room for improvement.

Findings for Type A and B Homes

It was important for the study team to include family home providers in the current study, to explore the “goodness of fit” between SUTQ criteria and home environments. As documented in the QRIS Compendium, states do not always have the same standards or expectations for home-based care, as for center-based care. Further, there is increasing attention paid to specific supports for family home providers, to facilitate improved quality and positive child outcomes. Forry et.al. (2013), for example, explored the predictors of quality in family home care and found that provider attitudes, beliefs, and practices are not only inter-related but also predictive of measures of global quality. Also important are provider stress, the number of children in care, and engagement with a professional association. The authors concluded that *“the findings of this study related to provider attitudes, beliefs, quality practices, and child outcomes suggest that further research is needed to identify effective delivery methods and content of professional development that addresses both quality practices and providers’ professional attitudes and beliefs.”*

Susman-Stillman et.al. (2013) explored a similar topic and found that there are differences between center-based and home-based provider practices that may be related to provider attitudes and beliefs. More specifically, family home providers tended to provide consistent levels of care over time, even when they had less positive attitudes and beliefs. In contrast, center-based providers exhibited varying levels of quality care, which was sensitive to provider attitudes and beliefs.

Finally, authors such as Raikes et.al. (2013) have documented the challenges of achieving quality in the home-based care environment, especially license-exempt providers, not least of which is shifting standards and expectations regarding quality as well as instruments for assessing quality. The authors also reported that home providers who received subsidies provided, on average, lower quality care than providers who did not receive subsidies; this is a troubling finding.

As discussed earlier, the study team generated a sample of Type A and B homes to include in the study. A total of 17 Type A and B homes agreed to participate; the distribution of sites is shown in Table 30.

Table 30. Participation of Type A and B homes

	Type A (n)	Type B (n)
Non-Participating Site	--	1
1-star	2	1
2-stars	1	2
3-stars	2	5
4-stars	--	1
5-stars	--	2

To maintain as much consistency as possible with the study methods used in ODE- and ODJFS-licensed early learning and development centers, Type A and B homes received a similar set of observations as did child care centers and Early Childhood Education classrooms: the FCCERS-R, the CHELLO, and the CIS. Family home providers also completed the Site Questionnaire and the FPTRQ (Director Measure).

FAMILY CHILD CARE ENVIRONMENT RATING SCALE-REVISED (FCCERS-R)

The FCCERS-R has seven subscales: Space and Furnishings, Personal Care Routines, Listening and Talking, Activities, Interaction, Program Structure, and Parents and Provider. The first six subscales were included in analyses, to ensure consistency with the other ERS used in the study.

The mean FCCERS-R rating was 3.2. The mean score in 1- to 2-star rated homes was 2.5 while the mean in 3- to 5-star rated homes was 3.7 (Table 31). While sample sizes are relatively small, there is a trend for 3- to 5-star rated homes to have higher mean scores on all subscales except Personal Care Routines. In fact, on the overall score as well as most subscales, the difference between lower rated and higher rated homes was one point or more.

The mean FCCERS-R score was 3.2.

Table 31. Agreement between star rating and FCCERS scores

		1- to 2-stars (n=5)	3- to 5-stars (n=5)
Space and Furnishings	Mean	2.3	3.8
	SD	.5	1.0
Personal Care Routines	Mean	2.4	2.5
	SD	.8	.4
Listening and Talking	Mean	3.3	4.9
	SD	2.0	1.3
Activities	Mean	1.8	3.4
	SD	.3	.5
Interaction	Mean	4.0	5.4
	SD	.9	1.3
Program Structure	Mean	2.5	3.9
	SD	.6	1.4
Overall Score	Mean	2.5	3.7
	SD	.5	.6

Data Sources: FCCERS-R

CAREGIVER INTERACTION SCALE (CIS)

The CIS was developed to capture the nature and tone of caregiver interactions with children. The CIS has been used in multiple early childhood studies and continues to be assessed (e.g., Colwell et.al. 2013). In the current study, the CIS was used to assess teacher-child interactions in Type A and B homes, as there currently is not a CLASS variant for family child care environments.

The CIS contains four subscales: Teacher Sensitivity, Teacher Harshness, Teacher Detachment, and Teacher Permissiveness. Each is measured on a four-point scale, wherein a score of “1” is the lowest possible rating and a score of “4” is the highest; mean scores closer to “4” are indicative of higher quality, developmentally appropriate, practices.

The mean CIS score was 3.7.

Each of the 17 participating sites in the current study were assessed with the CIS. As can be seen in Table 32, mean scores in the subscales and in the total overall score were relatively high (above a mean of 3, on a four-point scale). This is not surprising, given the findings of Colwell et.al. (2013), who found that the scale may not differentiate between providers who are “moderately” sensitive and those who are “highly” sensitive to children, as providers tend to interact positively with the children in their care. The authors noted that this is similar to items on the CLASS, in which providers tend to score highly on the Emotional Support subscale.

Table 32. Agreement between star rating and CIS scores

		1- to 2-stars (n=6)	3- to 5-stars (n=10)
Sensitivity	Mean	3.5	3.7
	SD	.4	.2
Harshness	Mean	3.6	3.8
	SD	.4	.3
Detachment	Mean	3.6	3.9
	SD	.7	.2
Permissiveness	Mean	3.1	3.0
	SD	.3	.3
Overall Mean Score	Mean	3.5	3.7
	SD	.4	.2

Data Source: CIS

The 1 Type B site that does not participate in SUTQ scored an overall mean score of 3.7.

CHILD HOME EARLY LANGUAGE AND LITERACY OBSERVATION (CHELLO)

The CHELLO was developed to assess the quality of the family home child care environment for supporting and facilitating development of language and literacy skills. The CHELLO contains two subscales: Literacy Environment (total possible score of 26) and Group/Family Observation (total possible score of 65), as well as an overall total score (total possible score of 91). The authors suggest that scores of 21 to 26 on the Literacy Environment Checklist are Exemplary, while scores of 11 to 20 are Fair. Similarly, scores of 55 to 65 on the Group/Family Observation are exemplary, while scores of 44 to 54 are Above Average, scores of 33 to 43 are Basic, and scores of 22 to 32 are Fair.

Type A and B homes that are 1- or 2-star rated received lower mean scores than sites that are more highly rated, as shown in Table 33. However, both groups scored, on average, in the exemplary range for both subscales.

Table 33. Agreement between star rating and CHELLO scores

		1- to 2-stars (n=6)	3- to 5-stars (n=10)
Literacy Environment Checklist	Mean	21.5	23.6
	SD	2.95	2.2
Group/Family Observation	Mean	54.5	62.4
	SD	8.1	2.3
CHELLO	Mean	76	86
	SD	10.9	3.3

Data Source: CHELLO

The 1 Type B site that does not participate in SUTQ scored 11 on the Literacy Environment Checklist and 50 on the Group/ Family Observation.

FAMILY PROVIDER/TEACHER RELATIONSHIP QUALITY SCALE-DIRECTOR MEASURE

Type A and B home providers completed the Director’s version of the FPTRQ scale. To analyze director responses, participating sites were again grouped in lower and higher SUTQ ratings. As shown in Table 34, higher rated sites tended to have higher mean scores in the Environment and Policy, Information about Resources, and Referrals subscales. As with centers and Early Childhood Education classrooms, scores at or above 13.2 for the Environment and Policy subscale suggest provider practices at or above “typical.” Higher rated Type A and B homes exceeded this benchmark, while lower rated homes came very close, with a mean score of 13. Type A and B homes were similar to participating centers in mean scores on the Environment and Policy, Communication Systems, and Information about Resources subscales but, on average, scored lower than centers and Early Childhood Education classrooms on Referrals.

Table 34. Agreement between lower and higher star rating and Director FPTRQ scores

		Lower Rated (1- to 2-stars)	Higher Rated (3- to 5-stars)
Environment and Policy Checklist; range 0-17	Mean	13.0	14.8
	SD	3.5	1.6
	n	3	5
Communication Systems; range 0-9	Mean	8.5	7.7
	SD	.7	.8
	n	2	6
Information about Resources; range 0-12	Mean	5.4	6.9
	SD	4.3	3.1
	n	5	7
Referrals; range 0-5	Mean	1.0	1.3
	SD	1.2	1.4
	n	5	7

Data Source: FPTRQ-Director Measure

The 1 Type B site that does not participate in SUTQ scored 8 on the Environment and Policy Checklist and 1 on Information about Resources.

ADMINISTRATION PRACTICES

Thirteen of the Type A and B participating homes conducted the BAS with trained data collectors. As with the Program Administration Scale, the tool targets management and business

practices, with subscales such as Qualifications and Professional Development, Income and Benefits, Recordkeeping, and Risk Management¹⁷.

The mean BAS score was 4.6.

The mean overall score across the 13 sites was 4.6, which is at the mid-point on the seven-point scale. Because of small sample sizes, sites were grouped into lower and higher rated sites, in which lower rated sites include 1- and 2-star rated sites and higher rated sites include 3- to 5-star rated sites. Mean scores for each grouping are presented in Table 35. Overall, Income and Benefits was the subscale in which lower-rated sites tended to score the lowest. Fiscal Management and Risk Management were the subscales in which higher rated sites tended to score the lowest.

Table 35. Agreement between star rating and administrative practices

		1- to 2-stars (n=5)	3- to 5-stars (n=7)
Qualifications and Professional Development	Mean	4.6	4.1
	SD	2.2	2.6
Income and Benefits	Mean	1.6	3.6
	SD	1.3	1.6
Work Environment	Mean	6.0	6.9
	SD	1.7	.4
Fiscal Management	Mean	2.6	3.3
	SD	2.5	2.6
Recordkeeping	Mean	5.2	5.7
	SD	1.6	1.6
Risk Management	Mean	5.2	3.3
	SD	2.2	1.9
Provider-Parent Communication	Mean	5.4	5.9
	SD	1.9	1.3
Community Resources	Mean	4.4	6.0
	SD	2.5	1.4
Marketing and Public Relations	Mean	4.2	5.7
	SD	2.6	1.3
Overall Score	Mean	4.4	4.9
	SD	1.2	1.1

Data Source: BAS

The 1 Type B site that does not participate in SUTQ scored an overall score of 3.9.

Not all validation studies include an examination of quality in family home providers. Further, not all QRIS are designed with a focus on the quality of care in family home environments. The findings from the current study suggest that the level of observed care and management is similar to center-based and Early Childhood Education classrooms, which is encouraging. However, some of the

¹⁷ The sub-scale Provider as Employer was removed from analysis, to ensure consistency between Type A and B sites.

SUTQ Validation Study Results

challenges and concerns raised by other authors regarding the quality of care in programs accepting publicly funded children may merit additional investigation.

Procedural Validity

In the current study, procedural validity was defined as the strength of the processes used to generate site star rating. In particular, the study team was interested in the following question:

To what extent is star rating sensitive to changes in scoring or measurement?

In short, the study team wished to determine whether or not site star rating would change appreciably if different measurement criteria or techniques were applied. Or, is a robust and unifying construct of quality conveyed by star rating or does star rating depend on measurement tool or technique? If, for example, star rating was relatively insensitive to different measurement techniques (i.e., a highly rated site remained a highly rated site, across techniques or tools), the study team could conclude that star rating was a meaningful and robust procedure as well as construct. On the other hand, if star rating changed markedly with different measurements, then the star rating process may be sensitive to approach and may not represent a robust and unifying concept of quality.

For this phase of the study, the study team used individual assessment and composite scores for sites and classrooms, focusing on child care centers and Early Childhood Education sites, and relying solely on observations of classroom practices. The study-generated scores represent a points system, in which the total possible score for each instrument was calculated and then divided into five rating tiers. Sites and classrooms whose total study-generated score fell into the lowest tier were assigned the lowest study rating, and so on. After assigning a study-generated rating to each site, the study team compared its study-generated rating to the SUTQ rating, to determine what percentage of sites were in agreement. Specifically, the study team believed the two systems were in agreement if the ratings fell within one point of each other. For example, if the SUTQ rating was 3, the study-generated rating constructed by the study team would need to be a 2, 3, or 4 for the two ratings to be in agreement.

Agreement between Star Rating and Individual Assessment Score

The study team first addressed agreement between site mean scores on individual assessment (ERS, CLASS, and ELLCO) and site star rating. Because the ERS and CLASS use seven-point scales, the following distribution was used to calculate a study-generated rating: scores of 6 or 7 were given a study

HIGHLIGHTS

- Independent observations of classroom quality align well with star rating for 2-, 3-, and 4-star rated sites. This suggests that a points-based approach, in which observation scores are used to generate ratings, would support the ratings for these sites.
- There was less alignment with star rating for 1- and 5-star rated sites. For 1-star sites, the relative rating may under-estimate the quality of care being provided at some sites. For 5-star sites, the relative rating may over-estimate the quality of care at some sites.
- Directors and owners differ in their needs for support and assistance in achieving and maintaining quality over time. Grants and financial assistance, however, was a fairly consistent need across sites.

SUTQ Validation Study Results

rating of “highest”, while a score of 5 was given a study rating of “moderate-high”, a score of 4 was given a study rating of “moderate”, a score of 3 was given the study rating of “moderate-low”, and scores of 1-2 were given a study rating of “lowest.” The ELLCO uses a five-point scale; a score of 5 was given the study rating of “highest”, a score of 4 was given the study rating of “high”, a score of 3 was given the study rating of “moderate”, a score of 2 was given a study rating of “fair”, and a score of 1 was given the study rating of “lowest.”

ENVIRONMENT RATING SCALES (ERS)

As shown in Table 36, there was relatively good agreement between mean ERS score and star rating for 2-, 3-, and 4-star sites. Specifically, when the study-generated rating system described above was implemented, there was 80% agreement between study-generated ratings and 2-star sites, 83% agreement between study-generated rating and 3-star sites, and 69% agreement between study-generated ratings and 4-star sites.

There was good agreement between study-generated ratings based on the ERS and 2-, 3-, and 4-star rated sites.

A good degree of agreement, especially such as that shown with 2- and 3-star sites, suggests that the SUTQ ratings are congruent with ERS scores, in that observed practices might correctly predict a site’s star rating (and especially so for 2- and 3-star sites).

Table 36. Agreement between SUTQ star rating and site mean ERS Score

	SUTQ Rating				
	1-star	2-star	3-star	4-star	5-star
Lowest Study Rating	22%	20%	8%	0%	0%
Moderate-Low Study Rating	33%	40%	8%	31%	33%
Moderate Study Rating	11%	20%	58%	31%	33%
Moderate-High Study Rating	22%	0%	17%	15%	17%
Highest Study Rating	11%	20%	8%	23%	17%

Data Sources: ECERS-3; ITERS-R

CLASSROOM ASSESSMENT SCORING SYSTEM (CLASS)

There was good agreement between study-generated ratings based on the CLASS and 3- and 4-star rated sites.

Examining the CLASS, there is good agreement between study-generated ratings and 3- and 4-star sites. This suggests that CLASS scores, distributed as described above, may be good predictors of site star rating for 3- and 4-star sites, but less so for 1-, 2-, and 5-star sites. In 1- and 2-star sites, the study-generated rating based on observed practices using the CLASS might generate a higher star rating than given in SUTQ; for 5-star sites, the converse is true.

SUTQ Validation Study Results

Table 37. Agreement between SUTQ star rating and site mean CLASS score

	SUTQ Rating				
	1-star	2-star	3-star	4-star	5-star
Lowest Study Rating	0%	10%	0%	0%	0%
Moderate-Low Study Rating	22%	10%	0%	8%	0%
Moderate Study Rating	44%	40%	50%	23%	42%
Moderate-High Study Rating	22%	30%	42%	62%	50%
Highest Study Rating	11%	10%	8%	8%	8%

Data Sources: CLASS Infant; CLASS Toddler; CLASS PreK

EARLY LANGUAGE AND LITERACY CLASSROOM OBSERVATION (ELLCO)

Finally, when the ELLCO was examined, agreement between SUTQ star rating and study-generated ratings was strong for 2- and 3-star sites but less so for 4-star sites, and weak for 1- and 5-star sites. In 4- and 5-star sites, ELLCO ratings might generate a lower rating than that earned in SUTQ and in 1-star sites, ELLCO ratings might generate a higher rating.

There was good agreement between study-generated ratings based on the ELLCO and 2- and 3-star rated sites.

Table 38. Agreement between SUTQ star rating and site mean ELLCO Score

	SUTQ Rating				
	1-star	2-star	3-star	4-star	5-star
Lowest Study Rating	0%	0%	0%	0%	0%
Fair Study Rating	22%	30%	25%	23%	17%
Moderate Study Rating	67%	60%	58%	38%	58%
High Study Rating	11%	10%	17%	38%	25%
Highest Study Rating	0%	0%	0%	0%	0%

Data Source: Early Language & Literacy Classroom Observation

This examination of study-generated ratings is highly dependent upon methodology---a change in the distribution of scores will obviously change the level of agreement between the study-generated rating and the SUTQ star rating. Nonetheless, this exercise exhibits possible agreements and discrepancies between SUTQ star rating and the relative quality of observed practices, using well-known instruments such as the ERS, CLASS, and ELLCO. In addition, there is some consistency across these instruments in that 2- to 4-star rated sites tend to have the strongest agreement with the study-generated ratings, while 1- and 5-star rated sites tend to have the weakest agreement.

SUTQ Validation Study Results

Non-Participating Sites

Finally, it may be helpful to examine where sites that are not participating in SUTQ might fall, using the study-generated rating system. As shown in Table 39, sampled sites had observed quality that ranged from the lowest to moderately-high study-generated ratings. In particular, mean CLASS scores appear to be a strength for these sampled sites, while mean ERS scores would assign sampled sites to the lowest two rating categories and mean ELLCO scores would assign these sampled sites to the fair-moderate rating categories.

Table 39. Study-generated ratings for sampled sites that do not participate in SUTQ

	Mean ERS	Mean CLASS		Mean ELLCO
Lowest	38%	0%	Lowest	13%
Moderate-Low	63%	38%	Fair	50%
Moderate	0%	50%	Moderate	38%
Moderate-High	0%	13%	High	0%
Highest	0%	0%	Highest	0%

Data Sources: ERS; CLASS; ELLCO

SITE COMPOSITE RATINGS

After examining the level of agreement between individual assessments and SUTQ star rating, the study team calculated composite ratings, again examining levels of agreement. The first composite score included ERS and CLASS observations, which were averaged across observed classrooms at each site. The total possible range for the ERS-CLASS composite was 2 to 14 points. The composite rating was a “5” if a site averaged 12-14 points, a “4” if the site averaged 9 to 11 points, a “3” if the site averaged 6 to 8 points, a “2” if the site averaged 3 to 5 points, and a “1” if the site averaged 1 or 2 points.

As shown in Table 40, there was no agreement between the composite rating and the SUTQ rating for 1-star sites, with observed quality suggesting a higher rating for these programs (e.g. a composite rating of 3 or 4). There was 70% agreement for 2-star sites with the lack of agreement again representing under-estimated quality. (This is to say, the composite score tended to assign a higher rating to 2-star sites, compared to the SUTQ rating.) While there was strong agreement on 3- and 4-star sites, (92% and 100%, respectively), the composite generated only 50% agreement for 5-star sites. In this latter case, the suggestion is that the composite score would generate a lower rating for some sites.

There was good agreement between an ERS-CLASS composite rating and 2-, 3-, and 4-star sites.

SUTQ Validation Study Results

Table 40. Agreement between ERS and CLASS composite rating and SUTQ star rating

	SUTQ Rating				
	1-star	2-star	3-star	4-star	5-star
Composite level E (lowest)	0%	0%	0%	0%	0%
Composite level D	0%	20%	0%	0%	0%
Composite level C	56%	50%	42%	54%	50%
Composite level B	44%	20%	50%	38%	42%
Composite level A (highest)	0%	10%	8%	8%	8%

Data sources: ERS; CLASS

The second site composite included the ERS, CLASS, and ELLCO observations, again averaged across observed classrooms at each site. The total possible range for the composite was 3 to 19 points. The composite rating was a “5” if a site averaged 16 to 19 points, a “4” if the site averaged 12 to 15 points, a “3” if the site averaged 9 to 11 points, a “2” if the site averaged 6 to 8 points, and a “1” if the site averaged 3 to 5 points.

There was good agreement between an ERS-CLASS-ELLCO composite rating and 3-, and 4-star sites.

Study team observations of classroom practices generated composite scores that ranged from 5 to 16 points. As shown in Table 41, there was 22% agreement between the composite rating and the SUTQ rating for 1-star sites, with observed quality suggesting a higher rating for these programs. There was 60% agreement for 2-star sites with the lack of agreement again representing under-estimated quality. There was again 92% agreement with 3-star sites and 100% agreement with 4-star sites. There was 58% agreement with 5-star sites, with the suggestion, again, that the composite rating would generate a lower rating for some sites.

Table 41. Agreement between ERS, CLASS, and ELLCO composite rating and SUTQ star rating

	SUTQ Rating				
	1-star	2-star	3-star	4-star	5-star
Composite level E (lowest)	0%	0%	0%	0%	0%
Composite level D	22%	20%	0%	0%	0%
Composite level C	33%	40%	50%	54%	42%
Composite level B	44%	30%	42%	38%	50%
Composite level A (highest)	0%	10%	8%	8%	8%

Data sources: ERS; CLASS, ELLCO

The final step was to determine the composite rating for sites that are not yet participating in SUTQ, using the rating procedures described above. As can be seen in Table 42 and Table 43, the composite ratings would generate scores in levels 1 through 3 for non-participating sites, suggesting these sites have low- to moderate quality, based on classroom observations.

SUTQ Validation Study Results

Table 42. ERS and CLASS composite ratings for sampled sites that do not participate in SUTQ

Composite level E (lowest)	0%
Composite level D	38%
Composite level C	63%
Composite level B	0%
Composite level A (highest)	0%

Data sources: ERS; CLASS

Table 43. ERS, CLASS, and ELLCO composite ratings for sampled sites that do not participate in SUTQ

Composite level E (lowest)	13%
Composite level D	38%
Composite level C	50%
Composite level B	0%
Composite level A (highest)	0%

Data sources: ERS; CLASS, ELLCO

It is interesting to note how composite ratings shift when the ELLCO is added. In general, addition of ELLCO observations (which focus on language and literacy supports and practices) tended to lower a site’s overall composite rating, in both SUTQ and non-participating sites. This suggests that language and literacy practices may be an area of weakness for some classrooms; reference to ELLCO scores provided earlier in this report suggests that Language and Literacy practices (as opposed to General Classroom Environment) may be an area for support and assistance.

CLASSROOM COMPOSITE SCORES

A second round of composite ratings were created for Early Childhood Education classrooms (serving 3- and 4-year old children), as opposed to sites. This was to allow classrooms to receive ratings based on their individual observations (whereas site ratings were averaged across observed classrooms, at that site). Classroom composite ratings were assigned using the same methods as site composites, described above.

As can be seen below, for the composite created using the ECERS-3 and the CLASS PreK, there was little agreement between classrooms in 1-star sites and the composite rating, with many classrooms observed at a higher level of quality. There was better agreement with classrooms in 2-star facilities, with 77% agreement between star rating and composite rating. Similarly, there was 100% agreement between star and composite rating for classrooms in 3-star sites and 4-star sites but only 59% agreement between star and composite rating for classrooms in 5-star sites. In the latter case, observed quality was such that the composite rating would give 41% of classrooms a lower score than their current star rating suggests.

There was good agreement between an ERS-CLASS composite classroom rating and 2-, 3-, and 4-star sites.

SUTQ Validation Study Results

Table 44. Agreement between ERS and CLASS composite rating and classroom rating

	SUTQ Rating				
	1-star	2-star	3-star	4-star	5-star
Composite level E (lowest)	0%	0%	0%	0%	0%
Composite level D	20%	23%	0%	0%	0%
Composite level C	40%	54%	41%	50%	41%
Composite level B	40%	15%	59%	39%	47%
Composite level A (highest)	0%	8%	0%	11%	12%

Data sources: ECERS-3; CLASS PreK

Another composite rating was generated for Early Childhood Education classrooms, adding the ELLCO to the composite. The procedures followed to create site composite ratings were again followed for classrooms. There was slightly better agreement between composite and star ratings for classrooms in 1-star sites (27% agreement) and for classrooms in 2-star sites (84% agreement). There was again 100% agreement between composite and star ratings for classrooms in 3-star sites, but 89% agreement between ratings for classrooms in 4-star sites. Finally, there was 47% agreement between ratings for classrooms in 5-star sites. The latter two cases suggest that the addition of observations for language and literacy practices may have decreased the relative placement of some classrooms in 4- and 5-star sites. To the extent that these practices are priority items for SUTQ, support may be necessary to ensure consistent and high quality practices.

There was good agreement between an ERS-CLASS-ELLCO classroom composite rating and 2-, 3-, and 4-star sites.

Table 45. Agreement between ERS, CLASS, and ELLCO composite rating and classroom rating

	SUTQ Rating				
	1-star	2-star	3-star	4-star	5-star
Composite level E (lowest)	0%	0%	0%	0%	0%
Composite level D	27%	38%	6%	11%	0%
Composite level C	47%	46%	65%	44%	53%
Composite level B	27%	15%	29%	39%	41%
Composite level A (highest)	0%	0%	0%	6%	6%

Data sources: ECERS-3, CLASS PreK, ELLCO

Finally, the study team assigned composite ratings to classrooms in non-participating sites to capture the average level of observed quality. For the composite created using the ECERS-3 and the CLASS PreK, 46% of classrooms in non-participating sites were given level 2 composite ratings and 54% were given level 3 composite ratings.

SUTQ Validation Study Results

Table 46. ERS and CLASS composite ratings for classrooms in sampled sites that do not participate in SUTQ

Composite level E (lowest)	0%
Composite level D	46%
Composite level C	54%
Composite level B	0%
Composite level A (highest)	0%

Data sources: ECERS-3; CLASS PreK

Similar findings were generated for the composite created using the ECERS-3, the CLASS PreK, and the ELLCO. As can be seen in Table 47, just over half of classrooms would be given a level 2 composite rating, 38% would be give a level 3 composite rating, and 7% (1 classroom) would be given a level 4 composite rating, based on observed quality of classroom instruction and interactions.

Table 47. ERS, CLASS, and ELLCO composite ratings for classrooms in sampled sites that do not participate in SUTQ

Composite level E (lowest)	0%
Composite level D	54%
Composite level C	38%
Composite level B	7%
Composite level A (highest)	0%

Data sources: ECERS-3; CLASS PreK, ELLCO

As noted earlier in this section, the choice of methodology in creating composite scores and assigning cut values for different ratings will affect the level of agreement between study-generated rating and SUTQ star rating. Therefore, it is perhaps valuable to consider this process as a sensitivity exercise, gauging the extent to which level of agreement shifts as criteria or standards for observed practices are shifted. (How, for example, does shifting cut scores for study-generated ratings affect the level of agreement between the study rating and SUTQ star rating?) The methodology described herein documents some trends for consistency and stability, particularly in 2- to 4-star rated sites and classrooms. In these cases, the approach used in the current study suggests that different levels of observed practice may correlate well with assigned star rating, in 2- to 4- star rated sites. Less strong is the association between observed practices and star rating for 1- and 5-star sites; in the former case, quality may be under-estimated and in the latter, over-estimated. If this is indeed the case, then star rating may not accurately convey the relative level of quality provided at a site.

Recently, some authors (such as Burchinal et.al. 2016; Le et.al. 2015; Hatfield et.al. 2016; Zaslow et.al. 2010) have suggested that quality is not linear in its impact on children. Rather, these authors posit that a threshold level of quality can be achieved, beyond which child-level impacts emerge and escalate. While the current study did not set out to investigate this theory, it may be helpful to consider some of these emerging contributions, especially as they relate to basement and ceiling levels of quality associated with different standards or assessments, which may impact how quality is defined within a state's QRIS.

Moving Forward

Participating directors and owners were given an opportunity, using the Site Questionnaire, to provide feedback and input as to what types of infrastructure would be helpful (if not necessary) to achieve and maintain high quality practices.

Table 48 presents the top-rated items for centers, homes, and Early Childhood Education classrooms. Table 49 presents the percent of respondents who indicated each item was either “very important” or “important” for improving quality. Overall results are presented, along with results for each type of site: child care center, Type A or B home, and Early Childhood Education classrooms.

The highest-rated item, overall, was “grants or financial assistance to buy materials and resources for classrooms”, identified in 94% of responses as either “very important” or “important.” The lowest-rated item, overall, was “assistance or support in becoming accredited”, rated in 61% of responses as either “very important” or “important.” This latter finding is consistent with responses to the value of accreditation in determining quality, in which only 51% of respondents, overall, agreed that accreditation was important.

For child care center directors, the highest-rated items were “grants or financial assistance to buy materials and resources for classrooms”, “financial assistance or support to attract more highly qualified staff”, and “financial assistance or support to retain more highly qualified staff”, rated as “very important” or “important” in 98% of responses. The lowest-rated item was “assistance or support in becoming accredited” (55% of responses).

Participants value grants or financial assistance in support of quality.

For Type A and B homes, the highest-rated item was “grants or financial assistance to buy materials and resources for classrooms” (82% of responses indicated “very important” or “important”), while the lowest-rated items were “regular, on-site, assistance in meeting the requirements for SUTQ ratings” and “support or assistance to understand how to afford and pay for high quality practices” (53% of responses).

Finally, for Early Childhood Education classrooms, the highest-rated item was “more online or computer-based trainings and professional development opportunities”, rated as “very important” or “important” in 100% of responses. The lowest-rated items were “on-site assistance in walking through and understanding the requirements for SUTQ ratings” and “regular, on-site, assistance in meeting the requirements for SUTQ ratings” (33% of responses).

SUTQ Validation Study Results

Table 48. Infrastructure ranked ratings, by respondent type

	Center	Type A and B Homes	Early Childhood Education Classrooms
Highest Rated	<ul style="list-style-type: none"> Grants or financial assistance to buy materials and resources for classrooms Financial assistance or support to attract more highly qualified staff Financial assistance or support to retain more highly qualified staff More trainings and professional development opportunities in my area More online or computer-based trainings and professional development opportunities 	<ul style="list-style-type: none"> Grants or financial assistance to buy materials and resources for classrooms Grants or financial assistance to improve the facility (e.g., landscaping, building repairs, painting) 	<ul style="list-style-type: none"> More online or computer-based trainings and professional development opportunities More trainings and professional development opportunities in my area Financial assistance or support to retain more highly qualified staff Grants or financial assistance to buy materials and resources for classrooms Grants or financial assistance to improve the facility (e.g., landscaping, building repairs, painting) Financial assistance or support to attract more highly qualified staff
Moderate	<ul style="list-style-type: none"> Online or computer-based support for understanding the requirements for SUTQ ratings On-site assistance in walking through and understanding the requirements for SUTQ ratings On-site assistance in walking through and understanding the requirements for SUTQ ratings Support or assistance to understand how to stay at high quality in the future Support or assistance to understand how to afford and pay for high quality practices 	<ul style="list-style-type: none"> On-site assistance in walking through and understanding the requirements for SUTQ ratings Online or computer-based support for meeting the requirements for SUTQ ratings Assistance or support in becoming accredited More trainings and professional development opportunities in my area More online or computer-based trainings and professional development opportunities A mentor or coach I can talk to Support or assistance to understand how to stay at high quality in the future 	<ul style="list-style-type: none"> Online or computer-based support for understanding the requirements for SUTQ ratings Online or computer-based support for meeting the requirements for SUTQ ratings Assistance or support in becoming accredited A mentor or coach I can talk to Support or assistance to understand how to afford and pay for high quality practices
Lowest Rated	<ul style="list-style-type: none"> Grants or financial assistance to improve the facility (e.g., landscaping, building repairs, painting) A mentor or coach I can talk to Online or computer-based support for meeting the requirements for SUTQ ratings Assistance or support in becoming accredited 	<ul style="list-style-type: none"> Online or computer-based support for understanding the requirements for SUTQ ratings Financial assistance or support to attract more highly qualified staff Financial assistance or support to retain more highly qualified staff Regular, on-site, assistance in meeting the requirements for SUTQ ratings Support or assistance to understand how to afford and pay for high quality practices 	<ul style="list-style-type: none"> Support or assistance to understand how to stay at high quality in the future On-site assistance in walking through and understanding the requirements for SUTQ ratings Regular, on-site, assistance in meeting the requirements for SUTQ ratings

SUTQ Validation Study Results

Table 49. Director/Owner feedback on helpful or necessary infrastructure

	Aggregate (n=79)	Centers (n=53)	Home (n=17)	Early Childhood Education Classrooms (n=9)
Grants or financial assistance to buy materials and resources for classrooms	94%	98%	82%	78%
Grants or financial assistance to improve the facility (e.g., landscaping, building repairs, painting)	78%	77%	71%	78%
On-site assistance in walking through and understanding the requirements for SUTQ ratings	74%	80%	65%	33%
Online or computer-based support for understanding the requirements for SUTQ ratings	79%	84%	59%	67%
Regular, on-site, assistance in meeting the requirements for SUTQ ratings	72%	80%	53%	33%
Online or computer-based support for meeting the requirements for SUTQ ratings	74%	73%	65%	67%
Assistance or support in becoming accredited	61%	55%	65%	67%
More trainings and professional development opportunities in my area	91%	92%	65%	89%
More online or computer-based trainings and professional development opportunities	90%	90%	65%	100%
A mentor or coach I can talk to	75%	76%	65%	67%
Financial assistance or support to attract more highly qualified staff	90%	98%	59%	78%
Financial assistance or support to retain more highly qualified staff	92%	98%	59%	89%
Support or assistance to understand how to afford and pay for high quality practices	75%	78%	53%	67%
Support or assistance to understand how to stay at high quality in the future	77%	79%	65%	56%

Data Source: Site Questionnaire

Predictive Validity

Predictive validity is the ability of a score on one measure or assessment to be predictive of a score on another measure or assessment. For the current study, the study team examined whether star rating was predictive of child kindergarten readiness abilities, as measured by Ohio’s Kindergarten Readiness Assessment. Key questions of interest were:

- ***What is the relationship between star rating and child kindergarten readiness skills upon kindergarten entry?***
- ***What is the relationship between independent assessments of quality and child language and literacy skills?***

HIGHLIGHTS

- Participation in higher quality early childhood education environments is aligned with better child scores on KRAs.
- Participation in SUTQ-rated programs is aligned with better child scores on KRAs.

The study team used data extracted from both the Departments of Job and Family Services and Education to respond to the first question. Specifically, the study team received a data extract from the Department of Education containing de-identified KRA data for 2014-2015 and 2015-2016. Students with KRA data were coded for participating in the schools’ Early Childhood Education (ECE) programs (rated 3-stars or higher), Preschool Special Education program, and publicly funded child care (using data provided by the Department of Job and Family Services). Children who were not coded to one of these three categories were assigned to one of the following two groups (using Department of Education indicators): economically disadvantaged students and students who were not considered economically disadvantaged.

There was an association between pre-kindergarten participation in ECE and subsidized private child care and KRA scores.

For the first set of analyses, students were collapsed into cohorts that participated in “lower” (1- to 2-stars) and “higher” (3- to 5-stars) rated Early Childhood Education programs. Students were further disaggregated by their mode of participation; three modes were possible: Early Childhood Education (ECE) classes, Preschool Special Education (PSE) classes, and publicly funded enrollment in private child care (i.e., publicly funded child care). SUTQ star rating was used to group publicly funded children into either lower or higher rated programs. ECE and PSE sites are required to have 3-star or higher ratings, which were considered “higher rated.”

Table 50 presents the findings from the 2014-2015 and 2015-2016 KRA. As can be seen, there are statistically significant differences in mean scores among ECE, PSE, and publicly funded students. Consistently, students that participated in higher rated programming during their pre-kindergarten year

SUTQ Validation Study Results

also had higher KRA scores, followed by students who participated in lower rated programming, and then students that participated in PSE programming¹⁸.

¹⁸ Note, children in PSE classrooms qualify for Individuals with Disabilities in Education Act (IDEA) services and support, due to the presence of special or developmental learning needs. Children who qualify for PSE also have Individualized Education Plans, or IEPs, that target their specific needs. It stands to reason that access to higher-rated early learning and development classrooms will facilitate the development and deployment of the IEP for qualifying children.

SUTQ Validation Study Results

Table 50. KRA mean scores for ECE, PSE, and publicly funded students, disaggregated by lower versus higher star rating: 2014-2015 and 2015-2016 KRA data

		2014-2015 KRA				2015-2016 KRA			
		Publicly funded children in lower-rated programs (1 or 2-stars) (n=2356)	ECE or publicly funded children in higher-rated programs (3-to 5-stars) (n=7141)	PSE children in higher-rated programs (n=8129)	ANOVA	Publicly funded children in lower-rated programs (1 or 2-stars) (n=1093)	ECE or publicly funded children in higher-rated programs (3-to 5-stars) (n=5429)	PSE children in higher-rated programs (n=7552)	ANOVA
Language & Literacy	Mean	264.09	266.31	258.88	p<.000	263.95	266.81	260.12	p<.000
	SD	10.7	11.6	15.8		11.4	11.9	15.0	
Mathematics	Mean	264.84	266.70	259.72	p<.000	262.87	265.68	259.90	p<.000
	SD	12.3	13.0	16.8		12.3	12.7	15.3	
Social Foundations	Mean	264.20	268.25	257.54	p<.000	268.00	272.49	261.69	p<.000
	SD	17.6	18.3	20.3		19.0	19.2	22.1	
Physical Development and Well-Being	Mean	266.39	269.17	257.73	p<.000	268.05	271.41	260.55	p<.000
	SD	16.5	16.5	20.3		17.1	16.6	20.6	
Overall Test Score	Mean	263.51	266.08	257.97	p<.000	263.87	267.06	259.61	p<.000
	SD	10.3	11.2	14.8		11.0	11.6	14.5	

SUTQ Validation Study Results

Student KRA scores also were disaggregated and examined by star rating; only students that had participated in publicly funded child care were included in the 2014-2015 KRA analyses. Sites with 4- and 5-stars were grouped together, because of small sample size in the 5-star category. As is shown in Table 51, in the 2014-2015 KRA data there were significant differences in mean Language & Literacy scores, with specific differences in mean scores for students who participated in 1-star programming versus students who participated in 4- or 5-star programming during their pre-kindergarten year.

Table 51. 2014-2015 KRA mean scores for publicly funded children, disaggregated by star rating

		Publicly funded children in sites that are enrolled in SUTQ sites with a 1-star rating (n=2156)	Publicly funded children in sites that are enrolled in SUTQ sites with a 2-star rating (n=200)	Publicly funded children in sites that are enrolled in SUTQ sites with a 3-star rating (n =1361)	Publicly funded children in sites that are enrolled in SUTQ sites with a 4 or 5-star rating (n =481)	ANOVA
Language & Literacy	Mean	264.10	263.95	265.04	265.53	p<.013
	SD	10.8	10.2	11.4	10.9	
Mathematics	Mean	264.92	263.95	264.98	265.09	ns
	SD	12.3	12.4	12.7	12.2	
Social Foundations	Mean	264.16	264.61	265.68	265.80	ns
	SD	17.4	19.1	18.2	18.1	
Physical Development and Well-Being	Mean	266.33	266.94	267.61	266.60	ns
	SD	16.2	19.0	17.0	16.4	
Overall Test Score	Mean	263.52	263.32	264.36	264.42	ns
	SD	10.3	10.6	10.9	10.6	

ns: Not Significant

The 2015-2016 KRA data, linked to 2014-2015 Early Childhood Education programming, allowed an extended analysis of star rating. For these analyses, star rating was available for the pre-kindergarten year for students in each of the three groups (ECE, PSE, and publicly funded child care). This allowed for a two-way Analysis of Variance to be conducted, wherein the mean KRA scores were examined for each group of students, in each star rating category¹⁹. As shown in Table 52, analyses failed to find significant differences among KRA mean scores, associated with star rating (with one exception, as noted). However, there were significant differences among scores associated with group membership. These findings are largely consistent with those shown in Table 50 and Table 51.

¹⁹ It is important to note the small sample sizes in some star-rated categories. Findings for groups with small sample sizes should be interpreted with caution.

SUTQ Validation Study Results

Table 52. 2015-2016 KRA mean scores for ECE, PSE, and publicly funded children, disaggregated by participation in star rated care

	n	Publicly funded children that were enrolled in sites with a 1-star rating			Publicly funded children that were enrolled in sites with a 2-star rating			ECE, PSE, and publicly funded children that were enrolled in sites with a 3-star rating			ECE, PSE, and publicly funded children that were enrolled in sites with a 4-star rating			ECE, PSE, and publicly funded children that were enrolled in sites with 5-star rating		
		--	--	Subs 737	--	--	Subs 356	ECE 55	PSE 50	Subs 518	ECE 306	PSE 144	Subs 175	ECE 2566	PSE 3081	Subs 147
Language & Literacy	Mean	-	-	263.85	-	-	264.16	266.69	255.28	263.93	269.51	258.90	263.24	267.53	258.72	264.44
	SD	-	-	11.5	-	-	11.3	10.4	16.3	11.5	12.6	15.8	11.6	11.9	15.8	10.9
	Two-way ANOVA	Interaction effects between star rating and ECE, PSE, or Publicly funded status were not statistically significant There were not significant differences among star rating but there were statistically different scores by ECE, PSE, and Publicly funded status														
Mathematics	Mean	-	-	262.96	-	-	262.68	265.04	254.64	262.25	268.99	259.15	262.70	266.09	258.33	262.56
	SD	-	-	12.7	-	-	11.3	9.2	18.4	12.3	13.7	15.3	11.8	12.5	16.0	11.1
	Two-way ANOVA	Interaction effects between star rating and ECE, PSE, or Publicly funded status were not statistically significant There were significant differences among star rating and by ECE, PSE, and Publicly funded status														
Social Foundations	Mean	-	-	268.19	-	-	267.61	272.64	260.58	267.75	274.24	260.78	266.09	274.03	259.75	269.78
	SD	-	-	19.6	-	-	17.6	14.7	23.5	19.1	19.0	20.2	19.7	19.2	22.4	19.4
	Two-way ANOVA	Interaction effects between star rating and ECE, PSE, or Publicly funded status were not statistically significant There were not significant differences among star rating but there were statistically different scores by ECE, PSE, and Publicly funded status														
Physical Development and Well-Being	Mean	-	-	268.14	-	-	267.87	271.56	256.66	268.01	272.91	260.88	265.91	272.59	258.77	268.16
	SD	-	-	17.3	-	-	16.6	15.2	21.1	16.8	16.3	21.4	16.7	16.5	20.8	16.8
	Two-way ANOVA	Interaction effects between star rating and ECE, PSE, or Publicly funded status were not statistically significant There were not significant differences among star rating but there were statistically different scores by ECE, PSE, and Publicly funded status														
Overall Test Score	Mean	-	-	263.92	-	-	263.77	266.76	255.86	263.61	269.5	258.81	262.94	267.95	258.08	264.17
	SD	-	-	11.3	-	-	10.4	8.3	15.8	10.8	12.2	14.7	11.1	11.6	15.0	10.3
	Two-way ANOVA	Interaction effects between star rating and ECE, PSE, or Publicly funded status were not statistically significant There were not significant differences among star rating but there were statistically different scores by ECE, PSE, and Publicly funded status														

SUTQ Validation Study Results

It can be challenging to find significant changes in student performance (a) over a short pre-post window and (b) within the transitional period for a QRIS. Specifically, it is difficult to confirm that changes in student performance can be attributed solely to participation in star rated child care.

Other states have conducted studies that examine the association between child care quality ratings and child outcomes. Of these, five states collected independent child outcome data, including Colorado (Zellman et al., 2008), Indiana (Elicker, 2011), Minnesota (Tout et al., 2011), Missouri (Thornburg et al., 2009), and Washington (Soderberg et al., 2016).

The nature of QRIS often means that experimental designs with random assignment are not possible. Therefore, there is a possibility of selection bias, as some parents (with greater resources) may choose higher quality care. In an effort to mitigate this bias, some studies included controls for family background and/or program and community characteristics, while others relied on a pretest-posttest design controlling for differential levels of development at baseline. Overall, findings are mixed and limited evidence is provided from these studies showing a relationship between program quality ratings and greater developmental gains for children. However, it is also important to note that none of these studies was designed to test a causal link between program quality and child outcomes (Karoly, 2014).

Pre and post data collected from consented students at participating sites indicate gains in receptive and expressive language skills, after adjusting for age expectations. However, variability in student scores prevents conclusions regarding the relationship of SUTQ and short-term gains in language skills.

Brief Summary of QRIS/Child Outcome Studies

• Studies with Independent Assessment of Child Development

Colorado: Children overall and children in poverty in programs at different quality rating levels did not differ systematically.

Indiana: Controlling for parental education and household income, preschool-age developmental assessments were significantly related to quality rating levels only for anxiety/ withdrawal behavior.

Minnesota: Children overall and children in poverty in programs at different quality rating levels did not differ systematically.

Missouri: Children both in poverty and not in poverty in the highest rated programs (4- to 5-stars) did better in social/behavior skills, motivation, self-control, and positive relationships than children in lower-rated programs.

Washington: Children in higher rated programs did better in receptive language, expressive language & fine motor than children in lower-rated programs.

• States Using Teacher Assessment of Child Development

Florida: Positive and significant association between school readiness assessment and quality ratings. Rate of growth of school readiness was higher but not significant for QRIS sites compared with non-QRIS sites.

Pennsylvania: The percentage of “proficient” children was higher for children in 4-star programs than those in 3-star programs, which were the only two levels examined (statistical significance and change scores not reported).

SUTQ Validation Study Results

Virginia: No correlation between quality rating levels and pre-literacy skills. Growth in certain literacy areas (i.e., alphabet knowledge, phonological awareness) was significantly greater in 3- and 4-star as compared to 2-star programs. However, children in higher-rated centers had greater declines in alphabet knowledge after controlling for pre-kindergarten pre-literacy skills, family background, center characteristics, and community characteristics.

Thus, in examining the findings from KRA assessment data (with additional findings presented in Appendix F), it is encouraging to observe differences in mean scores among students who attended higher and lower rated sites. Participation in either ECE, PSE, or publicly funded child care also highlights the variety of program options and needs that students may experience prior to their entry into kindergarten.

Limitations

There are several study limitations to be addressed. First, the study had a narrow time window in which to collect data. All site and child data were collected during a seven-month period, with children granted a three-month window between pre- and post-assessment. This narrow window may have affected the level of change the study was capable of detecting in young children who received Brigance assessments.

Second, the study did not incorporate random assignment of sites, classrooms, or children. The study took advantage of the current distribution of sites in SUTQ and used a stratified random sample to ensure a representative sample of sites was included. However, there may be selection bias in that sites that choose to participate in SUTQ may have inherent differences in attitudes, beliefs, or perceptions, compared to sites that do not choose to participate in SUTQ.

Finally, the study was conducted less than three full years after the transition from a three-tiered rating system to a five-tiered rating system. In fact, as regards child outcomes, some of the data reviewed dated to the transition period, when SUTQ-participating sites were receiving new ratings and the new system was starting its implementation. It is important to interpret findings with the transition period in mind.

Discussion

The study team considers SUTQ a maturing system. It is helpful, therefore, to imagine the processes that are occurring at state, local, and site levels. In this section we present a potential Theory of Change for SUTQ, which posits both shorter- and longer-term outcomes that may be observed. This section also describes, in brief, what actually was observed at study sites.

Theory of Change

What is the Theory of Change for understanding, if not predicting, statewide changes in quality child care and early education practices? First, there are investments—Ohio has allocated significant resources into establishing and maintaining SUTQ, and to broadening its scope and reach. Second are the strategies and activities that the investments support. SUTQ provides guidance for quality improvements, supports and assistance for sites wishing to participate, and financial resources to alleviate the costs of providing and sustaining high quality.

Investments have outputs, in particular the numbers of participants (both sites and children). Over time, SUTQ investments have and will continue to realize increasing participation across the state. As increasing numbers of sites engage in SUTQ, statewide expectations and practices will shift. Finally, there are shorter-term and longer-term outcomes related to SUTQ. Short-term outcomes include the changes in environments, interactions, and practices in sites and classrooms as directors and teachers internalize and implement “high quality.” The time frame for such short-term changes is hard to predict—some sites, already operating at relatively high quality, may need a relatively short amount of time, if any, to achieve these outcomes. Other sites may need a longer timeframe to achieve high quality in status and practice—keeping in mind that staff turnover might lead to pauses or plateaus in a site’s progress towards quality, as the site hires, trains, and coaches new staff.

As a site matures in its quality, it is reasonable to expect consistent and sustained longer-term impacts on children. This is to say that child outcomes in the form of advanced developmental progress or growth may not emerge in the early years of a site’s participation in SUTQ. In the long-term, however, as state infrastructure, expectations, and standard practices change to reflect and accommodate the QRIS, one can expect to see a sustained emphasis on and investment in quality that translates into better child outcomes.

What was expected and what was observed?

Generally speaking, the five-tiered rating system suggests a ranked approach to quality wherein sites with lower ratings (1- or 2-star, for example) provide better quality than sites that have failed to achieve a rating. At the same time, sites with lower ratings still have room for improvement, compared to sites with higher ratings (such as 4- or 5-star). One of the study’s objectives was to test this hypothesis. As discussed below, there was some evidence in support of this hypothesis; findings suggest that quality rankings are not discrete constructs, but rather more gradual shifts in quality along a variety of quality domains and practices. Thus, the differences in quality between 1- and 2-star sites or 2- and 3-star sites may be more a difference in degree rather than a categorical difference. Further, differences in

quality may accrue at different rates in different domains and practices—evidence from the current study suggests that sites achieve advances in some domains while struggling in others²⁰.

1-star sites

Nine center-based 1-star sites (and 3 Type A/B homes), representing 33 observed classrooms, participated in the study. Directors at these sites were in good agreement that Learning and Development and Administrative and Leadership Practices criteria were important for quality, followed by Staff Qualifications, Family and Community Partnerships, and Group Size standards.

Classrooms in 1-star rated centers tended to have observed practices that scored in the lower-to-moderate range of practice on the ERS, CLASS, and ELLCO. That stated, the procedural exercises conducted in this study suggested that observed practices are higher in quality than relative star rating might suggest, with at least some 1-star sites scoring in the moderate range of study-generated ratings for observed practices. Teachers at 1-star rated sites reported family engagement practices that were within range of more highly rated sites, a finding that was supported by parent reports on the same constructs. One-star sites tended to have the lowest proportion of teachers with two-year or higher degrees, compared to 2- through 5-star rated sites. However, mean years of experience were in range of those found at the highest rated (5-star) sites.

2-star sites

Ten 2-star centers (and three Type A/B homes), representing 31 classrooms, participated in the study. The 2-star sites demonstrated interesting patterns in observed and reported findings. For example, when asked what factors are indicative of quality, fewer 2-star directors agreed with questionnaire items than 1-star sites. Further, there was less overall agreement among 2-star directors about the importance of Learning and Development or Administrative and Management Practices, compared to sites with lower and higher star ratings. Aside from sites that were not participating in SUTQ, 2-star sites exhibited the lowest mean scores for observed classroom practices on some scales. This latter finding also was observed in teacher and parent reports of family relationship practices and quality.

One area in which 2-star sites did not exhibit lower overall performance is the percentage of teachers with college degrees—in this, 2-star sites reported a higher percentage of teachers with two-year or higher degrees than 1-star sites, but fewer than 3-star or higher rated sites. This is consistent with the criteria needed to advance in star rating, so this finding is not surprising.

The overall portrait of 2-star sites is confounded by the study's procedural exercises that aligned observation-based ratings of site quality with SUTQ star ratings—in short, there was good or better alignment of study-generated ratings and SUTQ star rating. This was not true for 1-star sites, suggesting that 1-star quality may be under-estimated while a 2-star rating may be appropriate for the observed level of practice.

²⁰ Appendices G and H contain additional information about Ohio's standards for star rated sites, including standards specific to the domains of school readiness.

3-star sites

A 3-star rating can be considered the gateway to high quality practices, as this is the culminating tier of the building blocks approach. Eleven 3-star centers, one 3-star Early Childhood Education classroom, and seven 3-star Type A/B homes participated in the study, representing 35 classrooms.

Directors at 3-star sites tended to align with directors in 1-star sites in their agreement as to the indicators of quality child care and early education. However, in contrast to directors from 1-star sites, directors at 3-star sites rated Learning and Development, Family and Community Partnerships, and Group Size items as most indicative of quality (and reported fewer indicators related to Administration and Management).

Three-star sites tended to exhibit higher scores than 2-star sites on observations of quality practices. This also is true of teacher and parent reports of the quality of family and provider relationships. Three-star sites have higher percentages of two-year (or higher) degreed staff, compared to 1- or 2-star sites but again, this is not surprising given the criteria required to advance in star rating.

Procedurally, there tends to be relatively high agreement between study-generated ratings of quality (using a points approach that relied upon observed practices to generate a rating). This suggests that SUTQ star ratings for 3-star sites are relatively accurate, compared for example to 1-star sites.

4-star sites

There were 15 4-star sites in the study: 10 centers, 1 home, and 3 Early Childhood Education classrooms. These sites encompassed 33 center and Early Childhood Education classrooms.

Directors at 4-star sites aligned with 5-star sites with regard to indicators of high quality, at the 50% and 75% benchmarks for agreement. At the highest benchmark for agreement (agreement with 90% or more of questionnaire items), directors at 4-star sites were more aligned with those at 2-star sites. However, in the types of factors that were highly rated as indicative of quality, directors at 4-star sites aligned with those at 3-star and 5-star sites in valuing Learning and Development, Family and Community Partnerships, and Group Size.

Classrooms in 4-star sites tended to receive higher observed ratings of quality than classrooms in 1- to 3-star sites, with a few exceptions. This is consistent with hypotheses that project gradual improvements in quality, rather than punctuated changes. However, mean FPTRQ scores reported by teachers and parents suggested that observed quality may not extend to the quality of or investments in family relationships, which is surprising. In keeping with criteria for advancement in star rating, 4-star sites have relatively high proportions of two-year (or higher) degreed staff. Finally, as with 3-star sites, there is a relatively high level of agreement with study-generated ratings of quality and SUTQ star ratings.

5-star sites

The 5-star rating is the apex of the SUTQ system; achievement of a 4- or 5-star rating is accomplished by the accrual of points for quality practices. Seven centers, two homes, and five Early Childhood Education classrooms, all at the 5-star level, participated in the study, representing 29 classrooms.

Teachers and directors at 5-star sites were the most likely to report agreement with many indicators of quality presented in the Site Questionnaire. Further, there was high agreement that items in Learning and Development, Administration and Management, Family and Community Partnerships, and Group Size were indicative of quality.

Classrooms in 5-star sites tended to generate the highest scores on observations of classroom quality, a finding that supports the hypothesis that 5-star sites provide the highest level of quality care and education. Teachers and parents who completed the FPTRQ Scale at 5-star sites did not always give these sites the highest ratings for family relationships and investments, however. Teachers at 5-star sites almost universally have two-year or higher degrees, and tend to have more experience than teachers in 2- to 4-star rated sites.

Procedurally, there is a disconnect between study-generated ratings and SUTQ star ratings for 5-star sites. Specifically, the study team's ratings suggested lower ratings for some 5-star sites. The explanation for this phenomenon lies in the fact that although 5-star sites experience the highest relative ratings on observed quality, they don't necessarily achieve the highest possible ratings on these observations. Therefore, 5-star sites both are high achieving (relative to 1- to 4-star sites) but also have room to grow on criterion-referenced scales such as the ERS, CLASS, and ELLCO.

Non-SUTQ participating sites

SUTQ is a voluntary system. Therefore, lack of a star rating does not necessarily mean that a site is providing low quality care for children. The quality of these "non-rated" sites was explored in the current study, which included eight centers and 1 home, representing 29 classrooms.

Generally speaking, non-rated sites generated scores that in some cases were on par with rated sites. Directors from non-rated sites, for example, tended to report higher levels of agreement with Learning and Development and Administrative and Management practices, similar to the responses provided by directors from 4- and 5-star rated sites. Parents and teachers from non-rated sites also reported family relationships that were similar in quality to that of more highly rated sites. In addition, there is a relatively high proportion of two-year or higher degreed staff at non-rated sites, with some sites reporting staff with higher levels of education than staff at 1- or 2-star sites. Staff at non-rated sites also were reported to have relatively high levels of experience, on par with, if not higher than, that reported for 5-star sites.

Overall, non-rated sites experienced the lowest ratings on observations of classroom quality. In fact, there appears to be a benefit to any level of participation in SUTQ, where classroom quality is concerned. Procedurally, the relatively low ratings of observed classroom quality for non-rated sites translated into study-generated ratings in the low to moderate range, with a few exceptions. This was especially true for ratings generated by ERS and ELLCO scores, suggesting that non-rated sites might benefit from SUTQ guidance or assistance in improving structural and process quality.

Conclusions

The current validation study examined the extent to which SUTQ has face, content, construct, procedural, and predictive validity. Findings suggest:

- SUTQ enjoys face validity in multiple components and constructs.
- SUTQ criteria are consistent with evidence-based practices and standards found in other state QRIS, an affirmation of SUTQ's content validity.
- Classrooms in higher rated sites tended to generate higher scores on assessments of classroom quality. However, the highest scores fall, on average, below the high performing benchmarks on these assessments, suggesting that even highly rated sites have room to make improvements in quality. This pattern is not necessarily repeated in teacher or parent reports of family engagement and outreach practices, an area in which all SUTQ sites might benefit from additional assistance and support.
- Study-generated ratings support the procedural validity of star ratings for 2- to 4-star sites. This is to say that a study-generated, points-based, approach to rating sites exhibited relatively strong agreement with the relative ranking of 2- to 4-star sites. This was not the case for 1-star and 5-star sites.
- There is evidence that sites with higher star ratings (3-star or higher) are associated with better child outcomes on the KRA. In addition, trends suggest that participation in any level of star-rated site may provide a benefit for publicly funded children, compared to the status quo. Thus, star ratings (especially when categorized as lower versus higher rated) may be predictive of child performance. It is important to note, however, that the study was not experimental in design—therefore the association of SUTQ rating and child outcomes cannot be interpreted as causal in nature. Further, rating explains a relatively small amount of variance in child performance—other factors not measured influence child performance through direct and indirect avenues.

Recommendations

There are three primary recommendations resulting from the current study:

- (1) Continue to provide, if not expand, support and assistance for child care providers and early educators. Consistent with some research, it is uncertain that higher levels of educational attainment always translate into higher quality practices. Onsite and offsite mentoring, coaching, training, and technical assistance can be important for helping providers and early educators translate knowledge into practice.
- (2) Continue to provide financial support and other resources for sites to offset the costs of providing higher quality care, including the costs of attracting and retaining highly qualified staff. Many authors agree that quality early learning and development is grounded in teacher practices and experience—therefore, sustainable high quality is tied to the ability to

nurture and retain talented staff. Current supports already provided to participating sites and programs include: tiered market rates (wherein higher rates for publicly funded child care are paid for sites with higher star ratings); access to training and professional development opportunities; access to free resources and materials such as curricula and assessments; access to web-based resources and information

(<http://www.earlychildhoodohio.org/index.stm> and <http://education.ohio.gov/Topics/Early-Learning/Step-Up-To-Quality-SUTQ>); collaboration with regional Child Care Resource and Referrals agencies; and others.

- (3) Continue to track SUTQ's shorter- and longer-term outcomes, suggested by a system-level Theory of Change. Shorter-term outcomes are those that were tracked most closely in the current study: improvements in environments and teacher practices. Longer-term outcomes include a positive impact on child outcomes and sustainable high quality practices in financially stable and resilient private and public sites.

Appendices

Appendix A: References

- Ansari, A. & Gershoff, E. (2016). Parent involvement in head start and children's development: Indirect effects through parenting. *Journal of Marriage and Family*, 78(2), pp. 562-579.
- Arnold, D. H., Aeljo, A., Doctoroff, G. L., & Ortiz, C. (2008). Parent involvement in preschool: Predictors and the relation of involvement to preliteracy development. *School Psychology*, 37(1), pp. 74-90.
- Azzi-Lessing, L. (2009). Quality support infrastructure in early childhood: Still (mostly) missing. *Early Childhood Research and Practice*, 11.
- Barnett, W. S. (2008). *Preschool education and its lasting effects: Research and policy implications*. Rutgers NJ: National Institute for Early Education Research.
- Brandes, J. A., Ormsbee, C. K., & Haring, K. A. (2007). From early intervention to early childhood programs: Timeline for Early Successful Transitions (TEST) Intervention in School and Clinic, 42(4), pp. 204-211.
- BUILD Initiative. (2016). QRIS Compendium. Website. As of December 15, 2016: <http://qriscompendium.org/>
- Burchinal, M., Zaslow, M., & Tarullo, L. (2016). Quality thresholds, features, and dosage in early care and education: Secondary data analyses of child outcomes. *Monographs of the Society for Research in Child Development*, 81(2), Serial No. 321.
- Buysse, V., Winton, P. J., & Rous, B. (2009). Reaching consensus on a definition of professional development for the early childhood field. *Topics in Early Childhood Special Education*, 28, pp. 235-243.
- Cassidy, D. J., Lower, J. K., Kinter-Duffy, V. L., Hedge, A. V., & Shim, J. (2011). The day-to-day reality of teacher turnover in preschool classrooms: An analysis of classroom context and teacher, director, and parent perspectives. *Journal of Research in Childhood Education*, 25(1), pp. 1-23.
- Center for Advanced Study of Teaching and Learning. (2010). Measuring and improving teacher-student interactions in pk-12 settings to enhance students' learning. Retrieved from: http://curry.virginia.edu/uploads/resourceLibrary/CLASS-MTP_PK-12_brief.pdf
- Child Care Services Association. (2015). WAGE\$ statewide final report: Fiscal year 2015. Retrieved from: http://www.childcareservices.org/wagesapps/StatewideFinalFY15_Full.pdf
- Clements, D. H., & Samara, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal*, 45, pp. 443-494
- 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), pp. 218-233.

- Cunningham, D. D. (2008). Literacy environment quality in preschool and children’s attitudes toward reading and writing. *Literacy Teaching and Learning*, 12(2), pp. 19-36.
- DeMarco, A., Yzejian, N., & Morgan, J. (2015). Evaluation of Mississippi Child Care Quality Stars program: Final report. Frank Porter Graham Child Development Institute, The University of North Carolina at Chapel Hill.
- Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children’s social and emotional competence: A randomized trial of the preschool “PATHS” curriculum. *Journal of Primary Prevention*, 28, pp. 67-91.
- Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D. Cai, K. Clifford, R. M., Ebanks, C., Griffin J. A., Henry, G. T., Howes, C., Iriondo-Perez, J., Jeon, H., Mashburn, A. J., Peisner-Feinberg, E., Pianta, R. C., Vandergrift, N., & Zill, N. (2007). Teachers’ education, classroom quality, and young children’s academic skills: Results from seven studies of preschool programs. *Child Development*, 78(2), pp. 558-580.
- Early, D. M., Bryant, D., Pianta, R., Clifford, R., Burchinal, M., Ritchie, S., Howes, C., & Oscar, B. (2006). Are teachers’ education, major, and credentials related to classroom quality and children’s academic gains in pre- kindergarten? *Early Childhood Research Quarterly*, 21, pp. 174 – 195.
- Early, D. M., Maxwell, K. L., Ponder, B. D., & Pan, Y. (2017). Improving teacher-child interactions: A randomized controlled trial of Making the Most of Classroom Interactions and My Teaching Partner professional development models. *Early Childhood Research Quarterly*, 38, pp. 57–70.
- Elicker, J., Langill, C. C., Ruprecht, K., Lewsader, J., & T. Anderson, T. (2011). Evaluation of Paths to QUALITY, Indiana’s child care quality rating and improvement system: Final report. West Lafayette, IN: Purdue University. Retrieved from: <https://www.in.gov/fssa/files/PTQFinalReportRev11012.pdf>
- Forry, N. D., Tout, K., Rothenberg, L., Sandstrom, H., Vesely, C. (2013). Child Care Decision- Making Literature Review. OPRE Brief 2013-45. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Frede, E. (2005). Assessment in a continuous improvement cycle: New Jersey’s Abbott Preschool Program. Report for the National Early Childhood Accountability Task Force.
- Gable, S., Laschober, T., Thornburg, K. R., & Mauzy, D. (2007). Cash incentives and turnover in center-based child care staff. *Early Childhood Research Quarterly*, 22, pp. 363–378.
- Gallagher, J., & Clifford, R. (2000). The missing support infrastructure in early childhood. *Early Childhood Research and Practice*, 2.
- Gomez, R. E., Kagan, S. L., & Fox, E. A. (2015). Professional development of the early childhood education teaching workforce in the United States: An overview. *Professional Development in Education*, 41(2), pp. 169-186.

- Grindal, T., Bowne, J. B., Yoshikawa, H., Schindler, H. S., Duncan, G. J., Magnuson, K., & Shonkoff, J. P. (2016). The added impact of parenting education in early childhood education programs: A meta-analysis. *Children and Youth Services Review*, 70, 238-249.
- Hamre, B. K., Le Paro, K. M., Pianta, R. C., LoCasale-Crouch, J. (2014). *Classroom Assessment Scoring System (CLASS) Manual Infant*. Baltimore MD: Paul H. Brookes Publishing Co.
- Han, H. S. (2014). Supporting early childhood teachers to promote children’s social competence: Components for best professional development practices. *Early Childhood Education Journal*, 42, pp. 171–179.
- Harms, T., Clifford, R.M., & Cryer, D. (2015). *Early Childhood Environment Rating Scale*. Third Edition. New York NY: Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R.M. (2007). *Family Child Care Environment Rating Scale*. Revised Edition. New York NY: Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R.M. (2006). *Infant/Toddler Environment Rating Scale*. Revised Edition. New York NY: Teachers College Press.
- Hatfield, B. E., Burchinal, M. R., Pianta, R. C., & Sideris, J. (2016). Thresholds in the association between quality of teacher–child interactions and preschool children’s school readiness skills. *Early Childhood Research Quarterly*, 36, pp. 561–571.
- Hayakawa, M., Giovanelli, A., Englund, M., & Reynolds, A. J. (2016). Not just academics: Paths of longitudinal effects from Parent Involvement to Substance Abuse in Emerging Adulthood. *Journal of Adolescent Health*, 58(4), pp. 433-439.
- Hayakawa, C. M. (2013). *Parent Involvement as a Mechanism of the Effects of Early Childhood Intervention*. Unpublished doctoral dissertation, University of Minnesota, Twin Cities.
- Heinemeier, S. & Leonard, L. (2013). *Sustainability of practice and practices*. Raleigh NC: Wake County Smart Start.
- Hindman, A. H., Skibbe, L. E., & Morrison, F. J. (2013). Teacher outreach to families across the transition to school: An examination of teachers’ practices and their unique contributions to children’s early academic outcomes. *Early Childhood Education Journal*, 41, pp. 391-399.
- Jackson, B., Larzelere, R., St. Clair, L., Corr, M., Fichter, C., & Egertson, H. (2006). The impact of Heads Up! Reading on early childhood educators' literacy practices and preschool children's literacy skills. *Early Childhood Research Quarterly*, 21, pp. 213–226.
- Jamison, K. R., Cabell, S. Q., LoCasale-Crouch, J., Hamre, B. K., & Pianta, R. C. (2014). CLASS-Infant: An observational measure for assessing teacher-infant interactions in center-based child care. *Early Education and Development*, 25(4), pp. 553-572.
- Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, 23, 51–68.






- Karoly, L. A. (2014). Validation studies for early learning and care quality rating and improvement systems: A review of the literature. Working Paper. Santa Monica, CA: RAND Corporation. Retrieved from http://www.rand.org/content/dam/rand/pubs/working_papers/WR1000/WR1051/RAND_WR1051.pdf
- Karoly, L. A., Schwartz, H. L., Setodji, C. M., & Haas, A. C. (2016). Evaluation of Delaware Stars for Early Success. Research Report. Santa Monica, CA: RAND Corporation. Retrieved from: http://www.rand.org/pubs/research_reports/RR1426.html
- Kelley, P., & Camilli, G. (2007). The impact of teacher education on outcomes in center-based early childhood education programs: A meta-analysis. Rutgers NJ: National Institute for Early Education Research
- Kim, K., Porter, T., Atkinson, V., Rui, N., Ramos, M., Brown, E., Guzman, L., Forry, N., & Nord, C. (2014). Family and Provider/Teacher Relationship Quality Measures: User's manual. OPRE Report 2014-65. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Kyzar, K. B., Chiu, C., Kemp, P., Aldersey, H. M., Turnbull, A. P., & Lindeman, D. P. (2014). Feasibility of an online professional development program for early intervention practitioners. *Infants & Young Children*, 27, pp. 174–191.
- Le, V., Schaak, D. D., & Setodji, C. M. (2015). Identifying baseline and ceiling thresholds within the Qualistar early learning quality rating and improvement system. *Early Childhood Research Quarterly*, 30, pp. 215-226.
- LeParo, K. M., Hamre, B. K., and Pianta, R. C. (2012). Classroom Assessment Scoring System (CLASS) Manual Toddler. Baltimore MD: Paul H. Brookes Publishing Co.
- Lee, J. S., Ginsburg, H. P., & Preston, M. D. (2009). Video Interactions for Teaching and Learning (VITAL): Analyzing videos online to learn to teach early childhood mathematics. *Australasian Journal of Early Childhood*, 34, pp. 19-23.
- Lillie, T., & Vakil, S. (2002). Transitions in early childhood for students with disabilities: Law and best practice. *Early Childhood Education Journal*, 30(1), pp. 53-58.
- LoCasale-Crouch, J., Mashburn, A.J., Downer, J. T., & Pianta, R. C. (2008). Pre-kindergarten teachers' use of transition practices and children's adjustment to kindergarten. *Early Childhood Research Quarterly*, 23, pp. 124-139.
- Lonigan, C. J., Phillips, B. M, Clancy, J. L., Landry, S. H., Swank, P. R., Assel, M. Taylor, H. B., Klein, A., Starkey, P., Domitrovich, C. E., Eisenberg, N., de Villiers, J., de Villiers, P., & Barnes, M. (2015). Impacts of a comprehensive school readiness curriculum for preschool children at risk for educational difficulties. *Child Development*, 86, pp. 1773–1793.

- Manning, M., Garvis, S., Fleming, C., & Wong, G. (2015). The relationship between teacher qualification and the quality of the early childhood care and learning environment: A systematic review. The Campbell Collaboration.
- National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education. (2003). Position statement on early childhood curriculum, assessment, and program evaluation. Retrieved from: <https://www.naeyc.org/files/naeyc/file/positions/pscape.pdf>
- National Center on Early Childhood Quality Assurance. (2016). QRIS resource guide. Website: <https://qrisguide.acf.hhs.gov/index.cfm?do=qrisabout>
- Neuman, S. B., Dwyer, J., & Koh, S. (2007). User's guide to the Child/Home Early Language & Literacy Observation (CHELLO) Tool. Baltimore MD: Paul H. Brookes Publishing Co.
- North Carolina Institute for Early Childhood Professional Development (2001). Planning for professional development in child care a guide to best practices and resources.
- Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research*, 21, pp. 381-391.
- Pianta, R., Downer, J., & Hamre, B. (2016). Quality in early education classrooms: Definitions, gaps, and systems, *Futureofchildren.org*, 26(2), pp. 119-137.
- Pianta, R. C., Belsky, J., Houts, R., & Morrison, F. (2007). Opportunities to learn in America's elementary classrooms. *Science*, 30(315), pp. 1795-1796.
- Pianta, R. C., La Paro, K. M., and Hamre, B. K. (2007). Classroom Assessment Scoring System (CLASS) Manual Pre-K. Baltimore MD: Paul H. Brookes Publishing Co.
- Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, 23, pp. 431-451.
- Pianta, R. C. Kraft-Sayre, M., Rimm-Kaufman, S., Gercke, N., & Higgins, T. (2001). Collaboration in building partnerships between families and schools: The National Center for Early Development and Learning's Kindergarten Transition Intervention. *Early Childhood Research Quarterly*, 16, pp. 117-132.
- Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early professional development intervention on head start teachers and children. *Journal of Educational Psychology*, 102(2), pp. 299-312.
- Puccioni, J. (2015). Parents' conceptions of school readiness, transition practices, and children's academic achievement trajectories. *The Journal of Educational Research*, 108(2), pp. 130-147.
- Raikes, H. H., Vogel, C., & Love, J. M. (2013). Family subgroups and impacts at 2, 3, and 5: Variability by race/ethnicity and demographic risk. *Monographs of the Society for Research in Child Development*, 78, pp. 64-92.

- Rous, B., & Hallam, R. (2012). Transition Services for Young Children With Disabilities: Research and Future Directions. *Topics in Early Childhood Special Education, 31*(4), pp. 232–240.
- Rous, B. Hallam, R., McCormick, K., & Cox, M. (2010). Practices that support the transition to public preschool programs: Results from a national survey. *Early Childhood Research Quarterly, 25*(1), pp. 17-32
- Schilder , D., & Carolan, M. (2014). State of the states policy snapshot: State early childhood assessment policies. New Brunswick, NJ: Center on Enhancing Early Learning Outcomes (CEELO).
- Schweinhart, L. J. & Weikart, D. P. (1997). The High/Scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly, 12*, 117-143.
- Shlay, A. B. (2010). African American, White and Hispanic child care preferences: A factorial survey analysis of welfare leavers by race and ethnicity. *Child Care & Early Education Research Connections, 39*(1), pp. 125-141.
- Sirinides, P., Fantuzzo, J., LeBoeuf, W., Barghaus, K., & Fink, R. (2015). An inquiry into Pennsylvania's Keystone STARS. Philadelphia, PA: Consortium for Policy Research in Education.
- Sirinides, P. (2010). Demonstrating quality: Pennsylvania Keystone STARS: 2010 program report. Harrisburg, PA: Office of Child Development and Early Learning. Retrieved from: <http://www.ocdelresearch.org/Reports/Keystone%20STARS/Keystone%20STARS%202010%20Evaluation%20Report.pdf>
- Smith, M. W., Brady, J. P., & Anastasopoulos, L. (2008). User’s Guide to the Early Language & Literacy Classroom Observation Pre-K Tool. Baltimore MD: Paul H. Brookes Publishing Co.
- Soderberg, J., Joseph, G. E., Stull, S., & Hassairi, N. (2016). Early Achievers standards validation study: Final report. Childcare Quality and Early Learning Center for Research & Professional Development, College of Education, University of Washington. Retrieved from: <http://del-public-files.s3-us-west-2.amazonaws.com/EA%20Report%205.31.16.pdf>
- Stone-MacDonald, A., & Douglass, A. (2015). Introducing online training in an early childhood professional development system: Lessons learned in one state. *Early Childhood Education Journal, 43*, 241-248.
- Susman-Stillman, A., Pleuss, J., & Englund, M. M. (2013). Attitudes and beliefs of family- and center-based child care providers predict differences in caregiving behavior over time. *Early Childhood Research Quarterly, 28*(4), pp. 905-917.
- Talan, T. N., & Bloom, P. J. (2009). *Business Administration Scale for Family Child Care*. New York NY: Teachers College Press.
- Talan, T. N., & Bloom, P. J. (2011). *Program Administration Scale*. Second Edition. New York NY: Teachers College Press.
- Teachstone Training. (2014). *Teacher-child interactions in early childhood: Research summary*.

- Thornburg, K. R., Mayfield, W. A., Hawks, J.S., & Fuger, K. L. (2009). The Missouri quality rating system school readiness study. Kansas City, M.O.: Center for Family Policy & Research University of Missouri and the Institute for Human Development University of Missouri. Retrieved from: <http://www.elcmdm.org/Knowledge%20Center/reports/MOQRSreport.pdf>
- Torquati, J., Raikes, H., & Huddleston-Casas, C. (2007). Teacher education, motivation, compensation, workplace support, and links to quality of center-based child care and teachers' intention to stay in the early childhood profession. *Early Childhood Research Quarterly*, 2, pp. 261–275.
- Tout, K & Starr, R. (2013). Key elements of a QRIS validation plan: Guidance and planning template. OPRE 2013-11. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Tout, K., Starr, R., Isner, T., Cleveland, J., Albertson-Junkans, L., Soli, M., & Quinn, K. (2011). Evaluation of Parent Aware: Minnesota's quality rating and improvement system pilot, final evaluation report. Prepared for the Minnesota Early Learning Foundation. Minneapolis, MN: Child Trends.
- Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). Issues for the next decade of quality rating and improvement systems. Child Trends and the U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. Issue Brief #3. Retrieved from: http://www.acf.hhs.gov/sites/default/files/opre/improv_systems.pdf
- Whitebook, M., Sakai, L., &Howes, C. (1997). NAEYC accreditation as a strategy for improving child care quality: An assessment. Final report. Washington, DC: Center for the Child Care Workforce.
- Winterbottom, C., & Jones, I. (2014). National accreditation and its role in early education: An analysis of Florida's Gold Seal Quality Child-Care program and licensing standards. *Journal of Early Childhood Research*, 12(1), pp. 64-76.
- Yoshikawa, H., Weiland, C., Brooks-Gunn, J., et.al. (2013). Investing in Our Future: The Evidence Base on Preschool Education. Society for Research in Child Development and the Foundation for Child Development.
- Zaslow, M., Tout, K., Halle, T., Whittaker, J. V., & Lavelle, B. (2010). Toward the identification of features of effective professional development for early childhood educators: Literature review. U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, Washington, D.C.
- 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.
- Zellman, G. L., Perlman, M., Le, V., & Setodji, C. M. (2008). Assessing the validity of the Qualistar Early Learning quality rating and improvement system as a tool for improving child care quality. RAND. Retrieved from: http://www.rand.org/content/dam/rand/pubs/monographs/2008/RAND_MG650.pdf

Appendix B: SUTQ Building Blocks—Centers

Standards	One-Star	Two-Star	Three-Star	Four-Star	Five-Star
					
LEARNING AND DEVELOPMENT					
Curriculum and Planning	Identify curriculum	Obtain curriculum	Implement curriculum	<input checked="" type="checkbox"/> *	<input checked="" type="checkbox"/>
			Demonstrate alignment to the assessment	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Teachers have daily access to curriculum	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	ELDS in classrooms	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Plan of Activities	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Intentional activities for all domains (extra points)	<input checked="" type="checkbox"/> (extra points)
				Learning and exploration opportunities (extra points)	<input checked="" type="checkbox"/> (extra points)
Child Screening and Assessment	Begins process of identifying screening tool to be used	Obtains screening tool	Administers screening tool w/in 60 days of enrollment	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Referrals made w/in 90 days of enrollment	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Results shared with parents	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Identifies staff to be trained on screening tool	Staff trained to administer/score screening tool	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			ELA administered	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Formal and informal assessments conducted	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Results of assessment shared with families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Assessment results used to inform instruction	<input checked="" type="checkbox"/> (extra points)	<input checked="" type="checkbox"/> (extra points)
				Adjust/refine instruction and evaluate child progress (extra points)	<input checked="" type="checkbox"/> (extra points)
				Families involved in using assessment data for child's plan (extra points)	<input checked="" type="checkbox"/> (extra points)

SUTQ Validation Study Results

Standards	One-Star	Two-Star	Three-Star	Four-Star	Five-Star
Interactions & Environment	Program classroom self-assessment selected	Program classroom self-assessment completed	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Teachers set goals/actions based on self-assessment	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			OCOT standards met	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Interactions & Environment				Lead teachers track progress on action steps and readjusts goals as needed (extra points)	<input checked="" type="checkbox"/> (extra points)

ADMINISTRATIVE & LEADERSHIP PRACTICES

Standards	One-Star	Two-Star	Three-Star	Four-Star	Five-Star
Staff Supports	Written wage structure	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		One staff support provided	Two staff supports provided	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Three staff supports provided (extra points)	<input checked="" type="checkbox"/> (extra points)
Program Administration	Program self-assessment completed	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Continuous Improvement Plan (CIP) w/one goal	CIP w/two goals	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			CIP includes input from staff/families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				CIP engages community partners (extra points)	<input checked="" type="checkbox"/> (extra points)
				Conducts annual survey with families and community partners (extra points)	<input checked="" type="checkbox"/> (extra points)
Staff Management	Annual PD plans for administrator/staff	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		One formal observation for teachers annually	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Observation results used for PD plans	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Observations used for Program CIP (extra points)	<input checked="" type="checkbox"/> (extra points)
				Classroom self-assessments used for teacher PD plans (extra points)	<input checked="" type="checkbox"/> (extra points)

Staff Management				Two formal observations for teachers annually (extra points)	<input checked="" type="checkbox"/> (extra points)
------------------	--	--	--	--	--

STAFF QUALIFICATIONS AND PROFESSIONAL DEVELOPMENT					
Standards	One-Star	Two-Star	Three-Star	Four-Star	Five-Star
Staff Education	Administrator has CDA/AA/CPL 2 or higher	Administrator has AA/CPL 3 or higher	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	50% of teachers have CDA, or one teacher has an AA or CPL 3 or higher	25% of teachers have AA/CPL 3 or higher	50% of teachers have AA/CPL 3 or higher	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Lead teacher has at least 3 ½ hours daily of instruction time	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Higher levels of education (extra points)	<input checked="" type="checkbox"/> (extra points)
Professional Development	Administrators and teachers must achieve and maintain a SUTQ PD Certificate every biennium	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Additional levels of training (extra points)	<input checked="" type="checkbox"/> (extra points)

FAMILY AND COMMUNITY PARTNERSHIPS					
Standards	One-Star	Two-Star	Three-Star	Four-Star	Five-Star
Transitions	Written information on transition plans	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Activities to assist children with transitions	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Children's records transferred upon request	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Individualized transition plans	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Written policies for transition plans (extra points)	<input checked="" type="checkbox"/> (extra points)
Communication and Engagement	Family information obtained	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Provides families information on community resources	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Variety of methods of communication used with families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

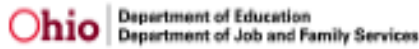
SUTQ Validation Study Results

Standards	One-Star	Two-Star	Three-Star	Four-Star	Five-Star
Communication and Engagement		Health and child development information provided annually	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		One opportunity for family engagement	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			One educational training for families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Written policy re: health screenings	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Child goals developed with families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Documented community partners (extra points)	<input checked="" type="checkbox"/> (extra points)
				Formal model for family engagement (extra points)	<input checked="" type="checkbox"/> (extra points)
				Parent volunteer group (extra points)	<input checked="" type="checkbox"/> (extra points)
Ratios and Accreditation	ADDITIONAL POINTS				
				Improved staff/child ratios and group sizes maintained (extra points)	<input checked="" type="checkbox"/> (extra points)
				Program is accredited by an approved body (extra points)	<input checked="" type="checkbox"/> (extra points)

* The indicates the standard listed previously in the row must be met.

*Please note: This checklist does not fully represent all requirements for rated SUTQ programs. It is meant to be a summary only of the requirements for each rating level. Please see rule 5101:2-17-01 for full program standards for each star-rating level.

Appendix C: SUTQ Building Blocks—Homes



Ohio Department of Job and Family Services
STEP UP TO QUALITY RATINGS PROGRESSIONS
FOR FAMILY CHILD CARE HOMES

Standards	One-Star	Two-Star	Three-Star	Four-Star	Five-Star
LEARNING AND DEVELOPMENT					
Curriculum and Planning	Identify curriculum	Obtain curriculum	Implement curriculum	<input checked="" type="checkbox"/> *	<input checked="" type="checkbox"/>
			Demonstrate alignment to the assessment	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Teachers have daily access to curriculum	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	ELDS in classrooms	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Plan of Activities	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Intentional activities for all domains (extra points)	<input checked="" type="checkbox"/> (extra points)
				Learning and exploration opportunities (extra points)	<input checked="" type="checkbox"/> (extra points)
Child Screening and Assessment	Begins process of identifying screening tool to be used	Obtains screening tool	Administers screening tool w/in 60 days of enrollment	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Referrals made w/in 90 days of enrollment	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Results shared with parents	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Identifies staff to be trained on screening tool	Staff trained to administer/score screening tool	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			EIA administered	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Formal and informal assessments conducted	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Results of assessment shared with families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Assessment results used to inform instruction	<input checked="" type="checkbox"/> (extra points)	<input checked="" type="checkbox"/> (extra points)
				Adjust/refine instruction and evaluate progress (extra points)	<input checked="" type="checkbox"/> (extra points)
				Families involved in using assessment data for child's plan (extra points)	<input checked="" type="checkbox"/> (extra points)

SUTQ Validation Study Results

Standards	One-Star	Two-Star	Three-Star	Four-Star	Five-Star
Interactions & Environment	Program (learning environment) self-assessment selected	Program (learning environment) self-assessment completed	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Owner/teacher sets goals/actions based on self-assessment	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			OCOT standards met	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Teachers track progress on goals (extra points)	<input checked="" type="checkbox"/> (extra points)

ADMINISTRATIVE & LEADERSHIP PRACTICES					
Program Administration	Program (administrative policies) self-assessment completed	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Continuous Improvement Plan (CIP) w/one goal	CIP w/two goals	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			CIP includes input from staff/families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				CIP engages community partners (extra points)	<input checked="" type="checkbox"/> (extra points)
				Conducts annual survey with parents and community partners (extra points)	<input checked="" type="checkbox"/> (extra points)
Staff Management	Annual professional development (PD) plans for staff	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Owner completes the Admin. CKC self-assessment annually	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		One formal observation annually for teachers (except owner)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Observation results used for PD plans	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Observations used for Program CIP (extra points)	<input checked="" type="checkbox"/> (extra points)
				Two formal observations for teachers annually (except owner) (extra points)	<input checked="" type="checkbox"/> (extra points)

STAFF QUALIFICATIONS AND PROFESSIONAL DEVELOPMENT					
Staff Education	Owner has a HS Dip or GED and CDA or a minimum of 25 pts in a CPL 1	Owner has a HS Dip or GED and CDA w/ a minimum of 50 pts in a CPL 2, or a minimum of 75 pts in a CPL 2	Owner has a HS Dip or GED and an AA in ECE, or a minimum of 125 pts in a CPL 2	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			One teacher has AA or a minimum of 100 pts in a CPL 2	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				<input checked="" type="checkbox"/> Higher levels of education (extra points)	<input checked="" type="checkbox"/> (extra points)
	Lead teacher has at least 3 ½ hours daily of instruction time	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Professional Development	Owner and teachers must achieve and maintain a SUTQ PD Certificate every biennium	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Additional levels of PD (extra points)	<input checked="" type="checkbox"/> (extra points)

FAMILY AND COMMUNITY PARTNERSHIPS					
Transitions	Written information to families on transition plans	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Activities to assist children with transitions	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Children's records transferred upon request	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Individualized transition plans	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Written policies for transition plans (extra points)	<input checked="" type="checkbox"/> (extra points)

SUTQ Validation Study Results

Communication and Engagement	Family information obtained	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Provides families information on community resources	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Variety of methods of communication used with families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Health and child development information provided annually	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		One opportunity for family engagement	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			One educational training for families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Written policy re: health screenings	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			Child goals developed with families	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
				Documented community partners (extra points)	<input checked="" type="checkbox"/> (extra points)
				Formal model for family engagement (extra points)	<input checked="" type="checkbox"/> (extra points)
				Parent volunteer group (extra points)	<input checked="" type="checkbox"/> (extra points)

Ratios and Accreditation	ADDITIONAL POINTS				
				Improved staff/child ratios and group sizes maintained (extra points)	<input checked="" type="checkbox"/> (extra points)
				Program is accredited by an approved body (extra points)	<input checked="" type="checkbox"/> (extra points)

* The indicates the standard listed previously in the row must be met.

***Please note:** This checklist does not fully represent all requirements for rated SUTQ programs. It is meant to be a summary only of the requirements for each rating level. Please see rule 3101:2-17-01 for full program standards for each star-rating level.

Appendix D: Methodology

This appendix contains additional details regarding study methodology.

Institutional Review Board Approval

The study team applied for and received Institutional Review Board approval to conduct the study. The consent documents developed to complete this process are provided below.

CONSENT DOCUMENTS

Exhibit 3. Participant Consent and Permission Form

PARTICIPANT CONSENT and PERMISSION FORM

for use with Owner or Directors of Child Care Facilities Sampled for Active Participation in Ohio's SUTQ Validation Study

Study Goals

The goals of the Ohio SUTQ validation study are to assess (a) the extent to which Ohio's 5 child care rating categories capture meaningful differences in child care quality and (b) the extent to which differences in child care quality are meaningful for child development and school readiness.

Study Methods

You are receiving this letter because your site was sampled to participate in the Ohio SUTQ validation study. What this means is that the study team will visit your site and collect data about its operations and classroom quality, including data collected using instruments such as the Environment Rating Scales, the Program or Business Administration Scale, the Classroom Assessment Scoring System (or Caregiver Interaction Scale), the Early Language and Literacy Classroom Observation (or the Child Home Early Language and Literacy Observation), the Family/Provider Teacher Relationship Quality Scale, best practices rubrics, and, if applicable, an assessment of practices for children with disabilities or special learning needs. Your site also will be asked to complete a questionnaire asking for information about the site's organization, staff, and experiences with SUTQ.

For participating child care centers, the study team would like to assess at least one classroom in each age group. For Type A and Type B home providers, the study team will observe the overall educational environment. In addition, the study team would like to interview you and work with you to complete the Program or Business Administration Scale. We anticipate directors will need approximately 15 minutes each, for two surveys (a Site Questionnaire and the Family/Provider Teacher Relationship Quality scale). Further, we request your assistance in checking in upon daily arrival, arranging classroom observations, and introducing our data team to the teachers. We anticipate this will require about 1-2 hours over the duration of the study. Finally, the Program or Business Administration Scale (administered to child care centers or homes, respectively) requires about 6-8 hours of director time, for those directors who agree to complete these scales.

Classroom observations are expected to take 3 days each, but will require minimal assistance or interaction from the teacher. We will ask teachers of 3 and 4 year old children to send parent information and consent and permission forms home with their students and to accept completed, sealed, documents returned by parents. We will ask teachers in all observed classrooms to complete the Family/Provider Teacher Relationship Quality scale—teacher version; we estimate this survey will take about 15 minutes to complete.

The study team also will draw upon data that already are “in house” at the Departments of Education and Job and Family Services. Your data will be treated confidentially and transferred securely.

Child Assessments

The study team will work with your site to identify and recruit 3- and 4-year old children for assessments using the Brigance Inventory of Early Development, Version III (Standardized). We ask your assistance in:

- Sending information, informed consent, and enrollment materials to parents,
- Receiving informed consent and enrollment materials that parents have completed, and
- Providing these documents to the study team for our use in identifying children for participation.

Depending upon the total number of children enrolled, we plan to spend between 1 and 5 days at your site collecting child assessment data.

Confidentiality

All study data will be released to and kept confidential by the study team. Further, any data collected are subject to the confidentiality provisions of appropriate state and federal laws and regulations, which prohibit the disclosure of information without specific written consent, or as otherwise permitted by such regulations. Only authorized members of the study team will have access to your site’s data, including any data collected on individual children, and your site’s identity will not be associated with any specific study findings. Hard copies of the data (as applicable) will be kept in secure locations by the study team and electronic versions of the data will be maintained on password-protected computers. Data will be kept for 3 years and then destroyed.

Assurances

- Your participation in this study is completely voluntary. You have the right not to consent to the collection or use of your site’s education records.
- You have the right to receive a copy of all data collected about your site, upon written request to the study team. This does not include child- or family-level data.
- This consent will remain in effect until or unless it is revoked by you, in writing, and delivered to Compass Evaluation and Research. Any such revocation shall not affect disclosures previously made by Compass Evaluation and Research, prior to the receipt of any such written revocation.
- There is no penalty for withdrawing from the study prior to its completion.
- All data will be collected by trained study team members. All data collected on individual children will be collected on-site, in a quiet area that allows you to view (but not participate in) the child assessment process, to ensure the child’s safety and well-being. The study team will work with you to identify an appropriate spot at your site for conducting child assessments.

Benefits to the Individual

There are no direct benefits to you or your site as a result of this study. That stated, this study is being conducted to validate Ohio’s SUTQ initiative. Your participation may help us better understand the initiative, its policies, procedures, and expectations, and its importance for your site.

Compensation

Your center or Type A or Type B home will receive a \$200 cash payment as a result of participating in this study, allowing site-, classroom-, and child-level data collection at your site, and allowing use of the data collected at your site. You will be asked to complete and sign a Payment Receipt when the payment is issued.

Risks to the Individual

The risks to participating in this research study are minimal and are no more than you (your teachers, or the children who enroll in the study) might experience in day-to-day life. This stated, there may be questions that you do not feel comfortable responding to—in these cases, you can decline to answer a question, you can withdraw yourself from the study, you can withdraw your site from the study, or you can withdraw both yourself and your site from the study. There is no penalty for declining to answer a question or for withdrawing yourself or your site from the study. In the event you decide to withdraw from the study, the study team may request your permission to continue the child assessments expected for children at your site who enroll in the study. In these cases, we would request days and times when we would be able to collect data on enrolled children, at your site, and as needed according to the study team data collection schedule.

Notification of Mandatory Child Report

If during the course of data collection a member of the study team believes that children are being abused or neglected while at your site, she is required to report such incidences to the proper authorities.

Questions:

If, at any time, you have questions or concerns about the study and your participation in it, please contact study directors directly:

Sarah Heinemeier, PhD
Co-Principal Investigator
Compass Evaluation and Research
5720 Fayetteville Road Suite 202
Durham NC 27713
Telephone: 919-544-9004 or 877-652-0189
Email: sarahhei@compasseval.com

Jennifer Hamilton, PhD
Co-Principal Investigator
Westat
1600 Research Blvd
Rockville MD 20850
Telephone: 301-251-1500
Email: JenniferHamilton@westat.com

If you have questions about your rights and welfare as a research participant, please call the Westat Human Subjects Protections office at 1-888-920-7631. Please leave a message with your full name, the name of the research study that you are calling about (Ohio's SUTQ Validation Study), and a phone number beginning with the area code. Someone will return your call as soon as possible.

If you agree to participate, please complete the following page and provide it to a member of the study team. You may retain the remaining pages for your information and files.

Exhibit 4. Agreement to Participate with Informed Consent and Permission

**Ohio SUTQ Validation Study
AGREEMENT TO PARTICIPATE WITH INFORMED CONSENT and PERMISSION**

Name of Site: _____ Address: _____

I, _____ (Director, Owner, or Authorized Representative name), authorize **Compass Evaluation and Research Inc.** (5720 Fayetteville Rd. Suite 202, Durham, NC 27713) to collect and use the following educational records and information:

- Site Questionnaire
- Environment Rating Scales (Infant/Toddler Environment Rating Scale, the Early Childhood Environment Rating Scale, and the Family Child Care Environment Rating Scale)
- Program or Business Administration Scale
- Classroom Assessment Scoring System or Caregiver Interaction Scale
- Early Language and Literacy Classroom Observations or Child Home Early Language and Literacy Observation
- SpecialLink Early Childhood Inclusion Quality Scale
- Family and Provider/Teacher Relationship Quality Scale (director and teacher versions)
- Best Practices Rubric: Early Education for Homeless/CPS/Immigrant Children

for the purpose of:

- Validating Ohio’s SUTQ initiative

I understand that: (1) I have the right not to consent to the collection or use of my site’s educational records; (2) I have the right to receive a copy of such records upon request; (3) and that this consent shall remain in effect until revoked by me, in writing, and delivered to Compass Evaluation and Research, but that any such revocation shall not affect disclosures previously made by Compass Evaluation and Research prior to the receipt of any such written revocation.

I have read the study materials and consent form. I have had the opportunity to ask questions about the study. I agree to participate in the study.

Signature of Director, Owner, or Authorized Representative

Date

Contact telephone number

Contact email address

THIS INFORMATION RELEASE IS SUBJECT TO THE CONFIDENTIALITY PROVISIONS OF APPROPRIATE STATE AND FEDERAL LAWS AND REGULATIONS WHICH PROHIBIT ANY FURTHER DISCLOSURE OF THIS INFORMATION WITHOUT THE SPECIFIC WRITTEN CONSENT OF THE PERSON TO WHOM IT PERTAINS, OR AS OTHERWISE PERMITTED BY SUCH REGULATIONS.

Exhibit 5. Parent and Child Consent and Permission Form

PARENT AND CHILD CONSENT AND PERMISSION FORM

We would like to invite you and your child to be part of a study of Ohio's SUTQ program. The SUTQ program helps track and rate the quality of Ohio child care and pre-kindergarten classrooms. We are inviting you because your child is enrolled at one of the child care centers.

The study is being paid for by Ohio's Department of Job and Family Services, with support from the federal government. Compass Evaluation and Research, Inc. and Westat are external research firms that have been hired to do the study. If you and your child join the study, we would like to:

- **Do some activities with your child** that will help us learn about his/her language and literacy skills. We will use a tool called the *Brigance Inventory of Early Development, Version III*, which involves things like looking at pictures and drawing letters. These activities will take between 15 and 30 minutes to finish. Sometimes, a child is having a bad day and cannot do the activities. If this happens, we will try again another day. Our team will do the activities in a quiet place at the child care center; your child's teacher or the child care director will be able to watch the activities. We would like to do these activities with your child in the spring and then again in the fall.
- **Ask you to complete two forms** which help us learn about you and your family. These two forms are called the *Child and Family Questionnaire* and the *Family and Provider/Teacher Relationship Quality* questionnaire. Each of these forms takes about 15 minutes to complete.

This study is voluntary. What this means is that you can choose whether or not to join the study and there is no penalty if you don't want to join.

There are very low risks to joining the study; the risks are no more than you or your child might experience in day-to-day life. In addition, if there are questions on the two forms that you don't want or don't feel comfortable answering, you don't have to answer them. In fact, you can leave the study at any time, without any penalty. The study does not affect your child's enrollment at the child care site.

There are no direct benefits to you or your child as a result of this study. However, we hope you will agree to join the study because we can learn a lot about the SUTQ program and how well it is working to improve child care quality.

As a thank you for joining the study, your child will receive a copy of the book *Goodnight Gorilla* (or a similar book) after finishing the first set of activities this spring.

All study data are confidential. This means that you and your child will not be identified by name to anyone who is not on the study team. In addition, we will not use you or your child's name in any reports. Your data will only be used by the study team to answer our study questions. The study team will follow state and federal rules for keeping your information confidential.

Notification of Mandatory Child Report. If during the course of data collection a member of the study team believes that children are being abused or neglected while at your site, the team member is required to report such incidences to the proper authorities.

If, at any time, you have questions or want to learn more about the study, please contact the study leaders directly. You can contact Dr. Sarah Heinemeier at 919-544-9004 or Dr. Jennifer Hamilton at 301-251-1500. We are happy to talk with you.

If you have questions about your rights as a part of the study, please call the Westat Human Subjects Protections office at 888-920-7631. Please leave a message with your full name, the name of the study that you are calling about (Ohio's SUTQ Validation Study), and a phone number beginning with the area code. Someone will return your call as soon as possible.

If you agree to join the study, please complete the following page and return it to the director at your child care. Please keep the other pages for your records.

We hope you will join us!

Exhibit 6. Parent and Child Consent and Permission Form

**Ohio SUTQ Validation Study
Parent and Child Consent and Permission Form**

Name of Child: _____ Date of Birth: _____

I, _____ (Parent/Legal Guardian Name), allow **Compass Evaluation and Research Inc.** (5720 Fayetteville Rd. Suite 202, Durham, NC 27713) to collect and use the following information:

- Child and Family Questionnaire (collected in spring 2016)
- Brigance Inventory of Early Development—Version III, Standardized (collected in spring and fall 2016)
- Family and Provider/Teacher Relationship Quality scale—parent version (collected in spring 2016)

for the purpose of:

- studying Ohio’s SUTQ program

I understand that:

- (1) I do not have to join the study; my participation is voluntary.
- (2) I can get a copy of the data that is collected on my child by sending a written note to the study team.
- (3) I can leave the study at any time, without any penalty, by sending a written note to the study team.

I have read the information about the study and this consent form. I have had a chance to ask questions about the study. I agree to participate in the study.

Signature of Parent or Guardian

Date

Contact telephone number

Contact email address

THIS INFORMATION RELEASE IS SUBJECT TO THE CONFIDENTIALITY PROVISIONS OF APPROPRIATE STATE AND FEDERAL LAWS AND REGULATIONS WHICH PROHIBIT ANY FURTHER DISCLOSURE OF THIS INFORMATION WITHOUT THE SPECIFIC WRITTEN CONSENT OF THE PERSON TO WHOM IT PERTAINS, OR AS OTHERWISE PERMITTED BY SUCH REGULATIONS

SAMPLING APPROACH

Data supplied by the Department of Job and Family Services (February 2016) were used to create a stratified random sample of sites for the study. Three types of site were considered: child care centers, Type A/B homes, and Early Childhood Education classrooms in elementary schools. Only sites serving 3- and 4-year old children were considered for the sample. This approach generated a sample of 2,260 child care centers, 100 sites with Early Childhood Education classrooms, and 2,955 homes.

The data sets were stratified by star rating (including sites that were not participating in SUTQ, which were given a rating of “0” for the purposes of creating a sample) and by urban or non-urban status. There were 1,085 urban and 1,275 non-urban child care centers and Early Childhood Education classroom sites and 1,743 urban and 1,212 non-urban Type A/B homes. Approximately 1,440 preschool-aged children were targeted in child care centers and Early Childhood Education classroom sites; children enrolled at Type A/B homes were not included in child assessments.

Table D. 1. Population used for sampling

	Not Participating in SUTQ	SUTQ 1-star	SUTQ 2-stars	SUTQ 3-stars	SUTQ 4-stars	SUTQ 5-stars
Centers and Early Childhood Education Classrooms	1,537	244	230	128	83	138
Type A/B Homes	2,692	166	73	15	2	7

The sample design for the centers and Early Childhood Education classrooms was a stratified two-stage clustered sample, whereas the design approach for the homes was a stratified one-stage design. The goal was to sample 72 child care and Early Childhood Education classroom sites and 25 homes, defined by location and SUTQ rating. An independent sample was selected in each stratum.

To address differences in size across sites, programs were further sorted into size groups within the location by rating strata. For centers/Early Childhood Education classrooms, group boundaries were determined for four sub-strata, with approximately equal numbers of children in each group. For homes, group boundaries were determined for three sub-strata with an approximately equal number of children in each group. With each size sub-strata, center and Early Childhood Education classroom sites were sorted by district. The sampling procedure used a random sort to randomize the selection and select a sample of two programs, with equal probability. For homes, the sampling procedure used a random sort and selected a simple random sample with equal probability. Non-participating and 1-star rated sites were sampled within size sub-strata.

An inflation factor of 30% was applied to compensate for program non-response in centers and Early Childhood Education classroom sites. Rounding to the nearest multiple, the study team sampled 8 programs per location by star rating strata, for a total sample of 96 centers/Early Childhood Education classroom sites. Replacement sites also were selected and assigned to each sampled site, using the

nearest neighbors in the sampling frame. These sites were used when there was non-response from the primary sample.

An inflation factor of 30% also was applied to the sample of homes, with a total sample of 30 homes. A replacement sample also was generated and was used when there was non-response in the primary sample.

DATA EXTRACTION

The study team entered into a data sharing agreement with the Departments of Job and Family Services and Education and completed four data extracts, two from the Department of Job and Family Services and two from the Department of Education. The data extracts were completed via secure file transfer protocol; the Department of Education extract contained de-identified data that could not be linked to any individual child by name, address, etc.

The data extracts were used to compile and generate information about the scope of SUTQ across the state as well as the connection of publicly funded child enrollment and attendance at SUTQ sites and later KRA scores.

All data were submitted in Excel format and then converted into an SPSS file for analysis. SPSS version 18 was used for all analyses.

SITE AND PARTICIPANT RECRUITMENT EFFORTS

The study team worked with staff at Job and Family Services to send two information letters to child care facilities. The first letter was sent to introduce the study and to inform facilities that there was a chance they could be sampled for inclusion in the study. The second letter was sent after sampling was completed, to inform facilities that were included in the sample of the potential for their participation. The study team contacted these facilities after the second letter was sent.

The study team field coordinator, located outside Cincinnati Ohio, first made contact with sampled sites via telephone. The purpose of the call was to provide additional information on the study and determine if the facility would agree to participate. Facilities that agreed to participate were then scheduled to meet with the lead team member for their area, to receive a study packet containing information, consent documents, instructions, and contact information. After receiving informed written consent from the site director, the lead team member scheduled the site for observations. The objective was to complete a full suite of observations in a classroom in each age grouping, including infants, toddlers, three-year olds, and four-year olds. At some sites, mixed classrooms were available instead of individual classrooms for each age group. Additionally, at some sites more than one classroom in each age grouping was available. In these cases, one classroom in each targeted age group (infants, toddlers, 3-year old, and 4-year old) was randomly selected for observation.

Site directors assisted the onsite data collection team by distributing information and consent forms to parents of 3- and 4-year old children enrolled at the site. Children who would be enrolled at the site in both spring and fall of 2016 were targeted. Once informed written consent was received from parents, the lead team member scheduled the child for assessment. Assessments were conducted within the classroom or office space at each site, in the visual presence of facility staff.

After completion of the study, sites received a \$200 check from the study team as a Thank You for their participation and assistance. Participating children received the book “Goodnight Gorilla”, after completing their first assessment.

DATA COLLECTOR TRAINING

ERS. Lead team members completed Environment Rating Scale (ECERS-3, ITERS-R) trainings provided online by the Environment Rating Scales Institute in March 2016. Team members then participated in double-coding exercises in the field, to ensure reliability with an experienced and highly reliable data collector. Team members that achieved 80% or higher inter-rater reliability with the ERS lead collector were allowed to collect ERS data from sites.

CLASS. Lead team members completed CLASS trainings (Infant, Toddler, and PreK) provided Teachstone in April 2016 in Columbus, Ohio. Team members that successfully completed Teachstone requirements for reliability were allowed to collect CLASS data from sites.

CIS. As there is no formal training for CIS, the study team’s principle investigator provided an overview of the CIS, its written documentation, and scoring instructions, after team members had successfully completed ERS, CLASS, and PAS/BAS trainings. Data were continually reviewed for consistency, outliers, and data entry or coding errors.

ELLCO/CHELLO Team members completed trainings provided by Brookes Publishing in May 2016 in Columbus, Ohio. Only team members who successfully completed the Brookes training were allowed to collect ELLCO or CHELLO data from sites.

PAS/BAS. Lead team members completed PAS/BAS trainings provided by McCormick Center for Early Childhood Leadership in April 2016 in Chicago, Illinois. Team members that successfully achieved 85% or higher reliability were allowed to collect PAS or BAS data from sites.

SpecialLink Inclusion Scale. There is no face-to-face training for Inclusion Scale, although the publisher makes a DVD available. Team members reviewed the DVD and the scale with the principle investigator, after successfully completing ERS, CLASS, and PAS/BAS trainings.

Brigance IED- III. Team members completed trainings provided by Curriculum Associates in March 2016 in Columbus, Ohio. Individual team members completed assessments under the direction of lead team members to ensure consistency. Data were continually reviewed for consistency, outliers, and data entry or coding errors.

INSTRUMENT STATISTICS

Internal consistency (Cronbach’s alpha) was measured for the Site Questionnaire (developed by the study team) and the FPTRQ. Peterson (1994) provides an overview of how to interpret Cronbach’s alpha coefficient. While authors have varied in their recommendations, an alpha coefficient of .7 or higher typically is considered acceptable.

Site Questionnaire. The Site Questionnaire (found in Appendix E) contained sub-items for each question, for questions 4 through 14. A composite variable was constructed by summing the point value of each sub-item, for each question. There were two exceptions to this practice, for questions 11 and 12.

For question 11, a respondent who indicated an Associate’s Degree, but not a Bachelor’s degree, was important for quality, was assigned a value of “1” for the relevant sub-items (11b, 11f, and 11i). Conversely, if the respondent indicated that a Bachelor’s Degree was important for quality, (items 11c, 11g, and 11j), the sub-items were assigned a value of “2”. In these cases, the sub-items 11b, 11f, and 11i did not receive a score, even if the respondent had marked them

For question 12, if the respondent marked item 12a, the question was scored as a “1”. If the respondent marked item 12b, the question was scored as a “2”—in these cases, item 12a did not receive a score, even if the respondent had marked it. If the respondent marked item 12c, the question was scored as a “3”—in these cases, items 12a and 12b did not receive a score, even if the respondent had marked them.

Once each question, from question 4 through 14, was scored, questions were grouped into constructs as follows:

Learning and Development: questions 4-6

Administrative and Leadership Practices: questions 7-10

Staff Education and Professional Development: questions 11-12

Family and Community Partnerships: questions 13-14

Group size and accreditation were not made into a composite variable. Tests for internal consistency were conducted by calculating Cronbach’s alpha for each question (i.e., component) and then each construct. Results were within acceptable limits, with an overall Cronbach’s alpha of .83. However, the alpha coefficient for Staff Education and Professional Development ($\alpha = .50$) was considered too low. Therefore, these items are best interpreted individually rather than grouped into a composite construct.

Table D. 2. Internal Consistency of the Site Questionnaire

Subscale	Number of items	Cronbach’s Alpha
Learning and Development	3	.76
Administrative and Leadership Practices	4	.72
Staff Education and Professional Development	2	.50
Family and Community Partnership	2	.69
Items 4 through 14	11	.83

FPTRQ-Provider/Teacher measure. The provider/teacher measure includes seven subscales. As shown in Table D. 3, most of the subscales show at least acceptable, and mostly good or excellent, reliabilities. Only one subscale (Commitment) shows poor reliability. Lack of variation of responses is the reason for the poor reliability, because nearly all responses within this subscale are a “3” or a “4” (after items PROVQ9b and PROVQ9d were reverse-coded).

Table D. 3. Cronbach's alpha of the provider/teacher measure overall

Provider/teacher measure	Number of items	Cronbach's Alpha
Family-specific Knowledge	12	.90
Collaboration	15	.93
Responsiveness	4	.73
Communication	4	.77
Commitment.....	4	.54
Openness to Change	8	.75
Respect	4	.82

FPTRQ-Parent measure. The parent measure includes eight subscales. As shown in Table D. 4, all of the subscales show good or excellent reliability.

Table D. 4. Cronbach's alpha of the parent measure overall

Provider/teacher measure	Number of items	Cronbach's Alpha
Family-specific Knowledge	15	.97
Collaboration	11	.94
Responsiveness	11	.93
Communication	8	.93
Family-focused Concern	3	.83
Commitment.....	9	.91
Understanding Context	4	.97
Respect	5	.77

DATA COMPILATION AND EXAMINATION

Data files were examined for missing data, out-of-range values, and other data entry errors. All data were verified.

In some instances, there were isolated cases of missing data such as an item that was left blank within the Site Questionnaire, Education Profile, or Brigance IED-III. In these cases, missing data were treated with mean value substitution (in which the mean value for that item or question across similar participants was used in place of the missing value).

After verifying, cleaning, and organizing data (including the construction and testing of composite values), data were examined using histograms, scatterplots, and box plots. Data visualization was conducted to determine whether data fit expectations for a normal distribution and to identify whether or not there appeared to be a continuous range of values or whether the data were forming aggregates or clumps.

Following visual examination of the data, basic descriptive analyses were conducted. For the Site Questionnaire, this involved the generation of frequency distributions for each component and each composite construct. For the remaining instruments, this involved calculation of measures of central tendency, including mean and standard deviation. These data are considered descriptive and are reported in the body of the report as well as in Appendix F: Additional Findings.

DATA ANALYSIS

Quantitative Data

Patterns and trends in data were examined for statistical significance using Chi-squared analysis, Analysis of Variance, Analysis of Covariance, and independent samples t-tests²¹. Analyses were conducted to test for significant relationships in the distribution of values across sites, between mean scores and star rating, and between mean scores and type of site. Follow-up analyses such as Tukey's test were conducted, as appropriate.

Results are reported in the body of the report as well as in Appendix F: Additional Findings.

Qualitative Data

Qualitative data were retrieved from Stakeholder Interviews, selected items from the Site Questionnaire, and selected items from the online survey. All qualitative data were subjected to content analysis. Data coding was completed by two study team members, who established reliability with each other for identifying and coding emergent themes in responses. Team members discussed and reached consensus for items in which there was disagreement about coding.

²¹ All analyses were conducted using Statistical Package for the Social Sciences (SPSS), version 18.

Appendix E: Site Questionnaire

Ohio's SUTQ Site Questionnaire

Thank you for joining the SUTQ validation study. This survey collects some information about you and your facility. All of your information will be kept confidential—this means that we will not identify you or your facility in any published reports. In addition, we will not link what you say to who you are in any reports.

1. Name of site: _____
2. How long has this facility been in operation? _____ years
3. How many total staff (including yourself) work here? _____

In the next sections we present ideas taken from SUTQ guidelines for achieving 1 to 5-star ratings. For each group of items, please check those items that you believe are important for providing high quality early education.

Section 1: Learning and Development

4. Curriculum and Planning

Please check the Curriculum and Planning items that you believe are important for providing high quality early education.

- | | |
|--------------------------|--|
| <input type="checkbox"/> | Using a research-based curriculum that is aligned with Ohio's Early Learning and Development Standards (Birth-K). |
| <input type="checkbox"/> | Ensuring each lead teacher has daily access to a copy of the curriculum and Ohio's Early Learning and Development Standards (Birth-K). |
| <input type="checkbox"/> | Ensuring lead teachers write and use a lesson plan or plan of activities in their classrooms, every day. |
| <input type="checkbox"/> | Ensuring lessons plans or the plan of activities is aligned with all of the developmental domains in Ohio's Birth-Kindergarten Entry Learning and Development Standards. |
| <input type="checkbox"/> | Ensuring lesson plans or the plan of activities is linked to the child care's child assessment process. |

- Ensuring teachers thoughtfully plan class activities and experiences that will meet the needs, interests, and abilities of all of the children in their classroom.

5. Child Screening and Assessment

Please check the Child Screening and Assessment items that you believe are important for providing high quality early education.

- Using a comprehensive, developmentally appropriate, screening instrument(s) for the age groups served.

- Ensuring the appropriate staff are trained to administer, score and use the developmentally appropriate screening instrument correctly.

- Ensuring all Birth-Kindergarten children get a comprehensive developmental screening within 60 business days of entry into the program—and once a year thereafter.

- Talking with families about the results of the developmental screening and, when necessary, making referrals to the appropriate community agencies.

- Using Ohio’s Early Childhood Comprehensive Assessment System to meet state requirements for all enrolled preschool age children.

- Ensuring all children receive ongoing formal and informal assessments and communicating the results with the families.

- Using the results from child assessments to make, adjust and refine instructional decisions and to evaluate child progress

- Ensuring families have multiple opportunities to learn about the assessment process, understand their child’s progress, and make contributions to the child’s education plan.

6. Interactions and Environment

Please check the Interactions and Environment items that you believe are important for providing high quality early education.

- Using a classroom self-assessment tool(s) that addresses the quality of the classroom environment and staff/child interactions.

- Ensuring each lead teacher identifies an area for classroom improvement and creates an action plan to make the improvement(s).

- Ensuring lead teachers document progress on action steps and readjust their improvement goals as needed.

- Ensuring each child experiences a well-structured learning environment and positive interactions with staff and other children.

Section 2: Administrative and Leadership Practices

7. Staff Supports I

Please check the Staff Support items that you believe are important for providing high quality early education.

- A written wage structure for staff at the child care site

- Offering child care staff one or more of the Ohio- approved staff supports.

8. Staff Supports II

Please check the Staff Support items that are **most important for you and your staff.**

- At least five days of paid leave (sick, vacation and/or personal) annually
- Health benefits (medical, dental, vision)
- Retirement benefits
- Discount on child care
- Tuition reimbursement
- Paid professional development
- T.E.A.C.H. Early Childhood--OHIO participation
- Life insurance
- Flexible spending account
- At least one hour of paid planning time weekly
- At least five paid holidays annually

9. Program Administration

Please check the Program Administration items that you believe are important for providing high quality early education.

- Using an annual program self-assessment using a tool that examines practices such as human resource leadership and development, family and community partnerships, program development and evaluation, and business and operations management.

- Creating and using a continuous improvement plan using results from the annual program self-assessment.

- Finding ways to work with community partners to support child and family outcomes.

- Conducting an annual survey with families, staff, and others to get feedback and review accomplishments.

10. Staff Management

Please check the Staff Management items that you

- Creating annual written professional development plans for the site’s administrators, lead teachers, and assistant teachers.

believe are important for providing high quality early education.

- Conducting at least one (but possibly two or more) formal observations of lead and assistant teachers.
- Using formal teacher observations to help create individual professional development plans.
- Using formal teacher observations to help create a program-level continuous improvement plan.
- Using classroom self-assessments to help create individual professional development plans.

Section 3: Staff Qualifications and Professional Development

11. Staff Education

Please check the Staff Education items that you believe are important for providing high quality early education.

- Administrator** has a Child Development Associate (CDA) or Career Pathways Level (CPL) 2
- Administrator** has an AA in ECE (or approved related field for school-age only programs) or a CPL 3
- Administrator** has a Bachelor’s degree or higher in ECE (or approved related field for school-age only programs)
- Administrator** has an administrator credential (ACL3)
- Lead teachers** have a Child Development Associate (CDA) or Career Pathways Level (CPL) 2
- Lead teachers** have a AA in early childhood education (ECE) or an approved related field (or a CPL3)
- Lead teachers** have a Bachelor’s degree or higher in early childhood education (ECE) or an approved related field
- Assistant teachers** have a Child Development Associate (CDA) or Career Pathways Level (CPL) 2
- Assistant teachers** have a AA in early childhood education (ECE) or an approved related field (or a CPL3)
- Assistant teachers** have a Bachelor’s degree or higher in early childhood education (ECE) or an approved related field

12. Professional Development

Please check the Professional Development items that you believe are important for providing high quality early education.

- Ensuring administrators, lead teachers and assistant teachers receive 20 hours of Ohio- approved specialized training every two years.

- Ensuring administrators, lead teachers and assistant teachers receive 25 hours of Ohio- approved specialized training every two years.

- Ensuring administrators, lead teachers and assistant teachers receive 30 hours of Ohio- approved specialized training every two years.

Section 4: Family and Community Partnerships

13. Transitions

Please check the Transition items that you believe are important for providing high quality early education.

- Providing written information to families on transitioning children into, within, and out of the program.

- Providing age-appropriate activities for children to prepare them for the transition to a new classroom or educational setting.

- Transferring any child’s records to the new setting at the family’s request and with the family’s written consent.

- Meeting with families to develop an individualized transition plan that supports a child’s transition to another classroom or educational setting.

- Using written transition policies and procedures that guide the transition process for children and families.

14. Communication and Engagement

Please check the Communication and Engagement items that you believe are important for providing high quality early education.

- Obtaining information about the family structure and routines that are important to the child’s development.

- Providing information regarding resources and community services to families.

- Communicating with families using different modes of communication.

- Each year, providing information on topics addressing health and child development families.

- Creating at least one opportunity for families to engage in activities.

14. Communication and Engagement

- Each year, creating at least one training, workshop or educational event to support families’ engagement in children’s learning and development.

- Using written policies and procedures to ensure that children have received comprehensive health screenings or that families have been provided information on the importance of health screenings and resources to obtain them.

- Working collaboratively with families to create annual written, developmental and/or educational goals for children.

- Using written documentation of formal and/or informal agreements with community partners and other family-serving agencies, programs and entities that may be helpful for the child or family.

- Using a formal model or process to enhance family engagement strategies.

- Creating and supporting an organized and active parent volunteer group.

Section 5: Group Size and Accreditation

15. Group Size and Accreditation

Please check the Group Size and Accreditation items that you believe are important for providing high quality early education.

- Meeting or improving on state-required child:staff ratios.

- Gaining accreditation from an approved accrediting body.

The next few pages ask additional questions about your experience with SUTQ and the supports that may be necessary to improve and maintain your site’s quality.

16. Are you currently participating in SUTQ?

Yes <input type="checkbox"/>	No <input type="checkbox"/>
If you answered yes, please provide some feedback regarding your experiences with SUTQ and achieving your star rating:	If you answered no, why aren’t you currently participating in SUTQ?

17. Which, if any, of the following supports and services might be helpful to you in improving your quality? Please indicate whether each item is *Not Important*, *Somewhat Important*, *Important*, or *Very Important to Improving Quality*.

	Not Important	Somewhat Important	Important	Very Important
Grants or financial assistance to buy materials and resources for classrooms	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Grants or financial assistance to improve the facility (e.g., landscaping, building repairs, painting)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
On-site assistance in walking through and understanding the requirements for SUTQ ratings	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Online or computer-based support for understanding the requirements for SUTQ ratings	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Regular, on-site, assistance in meeting the requirements for SUTQ ratings (e.g., help with curriculum and lesson planning, screening and assessments, learning environments, and developmentally appropriate practices)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Online or computer-based support for meeting the requirements for SUTQ (e.g., help with curriculum and lesson planning, screening and assessments, learning environments, and developmentally appropriate practices) ratings	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Assistance or support in becoming accredited	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
More trainings and professional development opportunities in my area	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
More online or computer-based trainings and professional development opportunities	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A mentor or coach I can talk to	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Financial assistance or support to attract more highly qualified staff	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Financial assistance or support to retain more highly qualified staff	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Support or assistance to understand how to afford and pay for high quality practices	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Support or assistance to understand how to stay at high quality in the future	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

18. How can SUTQ either improve your experience or encourage you to participate?

Appendix F: Additional Findings

This appendix contains additional technical findings generated from the assessments and instruments used in the study. Whereas the body of the report presents findings pertinent to the study's primary questions, the tables found in this appendix provide additional descriptive information that might prove useful.

Early Childhood Environment Rating Scale-3 (ECERS-3)

Several statistics were calculated for the ECERS-3, including overall mean score, mean score in center classrooms and Early Childhood Education classrooms, and mean score in urban and non-urban sites. In addition to overall score, mean scores for each subscale also were calculated and are reported, below, in Table F. 1 and Table F. 2 .

The mean ECERS-3 score was 3.6, on a seven-point scale. Scores for subscales ranged from 2.7 (Learning Activities) to 4.2 (Space and Furnishings and Interactions).

As can be seen in Table F. 1, there were relatively small differences between classrooms in centers and Early Childhood Education classrooms. The largest difference in mean score was found in the subscale Learning Activities, in which center-based classrooms scored slightly higher than Early Childhood Education classrooms. See as well the slight differences between center and Early Childhood Education classrooms in the subscale Program Structure, with Early Childhood Education classrooms scoring slightly higher, on average, than center classrooms.

Table F. 1. Mean ECERS--3 score by type

		Overall (n=95)	Center Classrooms (n=85)	Early Childhood Education Classrooms (n=10) ²²
Space and Furnishings	Mean	4.2	4.2	4.1
	SD	1.4	1.4	1.3
Personal Care Routines	Mean	4.1	4.1	4.0
	SD	1.7	1.7	1.5
Language and Literacy	Mean	3.5	3.5	3.4
	SD	1.6	1.6	.9
Learning Activities	Mean	2.7	2.8	2.0
	SD	1.3	1.4	.5
Interactions	Mean	4.2	4.2	4.2
	SD	1.8	1.8	1.4
Program Structure	Mean	3.9	3.9	4.3
	SD	1.7	1.8	1.4

²² When examining this and related tables, it is important to note the discrepancy in sample size between center and pre-kindergarten classrooms. Tests for statistical significance were not applied.

SUTQ Validation Study Results

		Overall (n=95)	Center Classrooms (n=85)	Early Childhood Education Classrooms (n=10) ²²
Mean ECERS-3 Score	Mean	3.6	3.6	3.4
	SD	1.3	1.4	.7

As regards classrooms in urban versus non-urban settings, for the most part there were slight differences in mean score. The largest differences occurred in the subscales Interactions and Program Structure, in which classrooms in non-urban settings scored slightly higher, on average, than classrooms in urban settings. In one instance, the mean program structure score, there were statistically significant differences between urban and non-urban sites ($p < .037$), with classrooms in non-urban sites exhibiting a significantly higher mean score than classrooms in urban sites.

Table F. 2. Mean ECERS-3 score by urban versus non-urban setting

		Urban Classrooms (n=52)	Non-Urban Classrooms (n=43)
Space and Furnishings	Mean	4.0	4.3
	SD	1.6	1.0
Personal Care Routines	Mean	4.0	4.2
	SD	1.9	1.5
Language and Literacy	Mean	3.4	3.7
	SD	1.7	1.3
Learning Activities	Mean	2.7	2.6
	SD	1.5	1.0
Interactions	Mean	3.9	4.5
	SD	2.0	1.4
Program Structure*	Mean	3.6	4.3
	SD	1.9	1.4
Mean ECERS-3 Score	Mean	3.5	3.7
	SD	1.6	1.0

**statistically significant difference at $p < .037$*

Infant/Toddler Environment Rating Scale-Revised (ITERS-R)

As with the ECERS-3, several statistics were calculated for the ITERS-R: overall mean score and mean score in urban and non-urban sites. Because Early Childhood Education classrooms do not serve infants or toddlers, there were no comparisons to make between center- and Early Childhood Education-based classrooms. Mean scores for each subscale also were calculated and are reported, below, in Table F. 3.

Note that the Parents and Provider subscale was not incorporated, to have consistency with the ECERS-3 scale, which does not include this particular subscale.

The mean ITERS-R score was 3.9, with mean scores ranging from 3.4 to 4.7 on individual subscales, as shown in Table F. 3. The highest subscale score was in the Interactions subscale while the lowest was in the Activities subscale.

Table F. 3. Mean ITERS-R score

		Overall (n=95)
Space and Furnishings	Mean	4.1
	SD	1.5
Personal Care Routines	Mean	3.7
	SD	1.96
Listening and Talking	Mean	3.9
	SD	1.8
Activities	Mean	3.4
	SD	1.5
Interactions	Mean	4.7
	SD	1.9
Program Structure	Mean	3.98
	SD	2.1
Mean ITERS-R Score	Mean	3.9
	SD	1.5

Classrooms in non-urban settings tended to exhibit higher mean scores than classrooms in urban settings, as can be seen in Table F. 4. The smallest differences were in the Activities subscale (a mean difference of .4 between urban and non-urban settings). Several differences in mean score were statistically significant, including Listening and Talking ($p<.028$), Interactions ($p<.011$), Program Structure ($p<.032$), and total mean score ($p<.041$). In each of these instances, classrooms in non-urban sites exhibited a higher mean score, compared to classrooms in urban sites.

Table F. 4. Mean ITERS-R score by urban versus non-urban setting

		Urban (n=46)	Non-Urban (n=49)
Space and Furnishings	Mean	3.8	4.4
	SD	1.6	1.4
Personal Care Routines	Mean	3.4	4.0
	SD	2.0	1.9

SUTQ Validation Study Results

		Urban (n=46)	Non-Urban (n=49)
Listening and Talking*	Mean	3.5	4.3
	SD	2.0	1.6
Activities	Mean	3.2	3.6
	SD	1.7	1.2
Interactions**	Mean	4.2	5.2
	SD	2.0	1.7
Program Structure***	Mean	3.5	4.4
	SD	2.3	1.8
Mean ITERS-R Score****	Mean	3.6	4.2
	SD	1.7	1.3

*statistically significant difference at $p < .028$

**statistically significant difference at $p < .011$

***statistically significant difference at $p < .032$

****statistically significant difference at $p < .041$

Classroom Assessment Scoring System-Infant (CLASS Infant)

There is one overall score for the CLASS Infant: Responsive Caregiving. Forty-three classrooms received the CLASS Infant, with a mean score of 4.6 on a seven-point scale. Only classrooms serving infants were rated using this scale, so no comparisons by type are possible.

Table F. 5. Mean CLASS Infant score

		Aggregate (n=43)
Responsive Caregiving	Mean	4.6
	SD	1.6

Examining Responsive Caregiving by location, classrooms in non-urban settings tended to score higher, on average, on the scale than classrooms in urban settings.

Table F. 6. Mean CLASS Infant score by urban versus non-urban setting

		Urban (n=21)	Non- Urban (n=22)
Responsive Caregiving	Mean	4.2	5.0
	SD	1.6	1.6

Classroom Assessment Scoring System-Toddler (CLASS-Toddler)

The CLASS-Toddler is validated for use in classrooms serving toddlers; no comparisons across type are possible. The scale has two subscales: Emotional and Behavioral Support and Engaged Support for Learning. As can be seen in Table F. 7, classrooms tended to score higher on the Emotional and Behavioral Support subscale than the Engaged Support for Learning subscale, suggesting that teachers in toddler classrooms struggled more with instructional (as compared to emotional) support.

Table F. 7. Mean CLASS-Toddler score

		Aggregate (n=52)
Emotional and Behavioral Support	Mean	5.3
	SD	.99
Engaged Support for Learning	Mean	3.1
	SD	1.2

There were no differences, on average, between classrooms in urban and non-urban settings, for either subscale, as shown in Table F. 8.

Table F. 8. Mean CLASS-Toddler score by urban versus non-urban setting

		Urban (n=26)	Non-Urban (n=26)
Emotional and Behavioral Support	Mean	5.3	5.3
	SD	1.0	.97
Engaged Support for Learning	Mean	3.1	3.1
	SD	1.2	1.3

Classroom Assessment Scoring System-PreK (CLASS PreK)

The CLASS-PreK is validated for use in classrooms serving 3- and 4-year old children. Thus, comparisons between center and Early Childhood Education classrooms are possible, as shown in Table F. 9. There are three subscales for comparison: Classroom Organization, Emotional Support, Classroom Organization, and Instructional Support. As can be seen in Table F. 9, teachers tended to score higher on the first two subscales than on Instructional Support; this trend is exhibited in both center and Early Childhood Education classrooms. Further, there were relatively small differences in mean scores, between centers and Early Childhood Education classrooms.

Table F. 9. Mean CLASS-PreK score

		Aggregate (n=96)	Centers (n=86)	Early Childhood Education Classrooms (n=10)
Emotional Support	Mean	5.5	5.5	5.5
	SD	.96	.98	.8
Classroom Organization	Mean	4.8	4.7	5.2
	SD	1.3	1.4	1.2
Instructional Support	Mean	2.8	2.8	2.7
	SD	1.2	1.2	1.0

Teachers in classrooms in urban-settings tended to score slightly higher in the Classroom Organization subscale, with few differences between urban and non-urban classrooms on the remaining two subscales.

Table F. 10. Mean CLASS-PreK score by urban versus non-urban setting

		Urban (n=44)	Non-Urban (n=52)
Classroom Organization	Mean	4.9	4.6
	SD	1.4	1.3
Emotional Support	Mean	5.4	5.5
	SD	1.0	.9
Instructional Support	Mean	2.8	2.7
	SD	1.2	1.2

Early Language & Literacy Classroom Observation (ELLCO)

The ELLCO is validated for use in classrooms serving 3- and 4-year old classrooms, which makes comparisons by type possible as shown in Table F. 11. On average, teachers exhibited higher scores on the General Classroom Environment subscale, compared to the Language and Literacy subscale—a pattern consistent in both center and Early Childhood Education classrooms. There were relatively small differences between center and Early Childhood Education classrooms on the General Classroom Environment mean subscale score, and no differences on the Language and Literacy mean subscale score.

Table F. 11. Mean ELLCO score by type

		Aggregate (n=96)	Centers (n=86)	Early Childhood Education Classrooms (n=10)
General Classroom Environment	Mean	3.4	3.4	3.6
	SD	.8	.8	.7
Language and Literacy	Mean	2.4	2.4	2.4
	SD	.7	.7	.6

Subscale scores were very similar in classrooms in both urban and non-urban settings, as shown in Table F. 12.

Table F. 12. Mean ELLCO score by urban versus non-urban setting

		Urban (n=52)	Non-Urban (n=44)
General Classroom Environment	Mean	3.5	3.4
	SD	.8	.7
Language and Literacy	Mean	2.5	2.4
	SD	.8	.6

FPTRQ-Director Measure

The FPTRQ (Director Measure) assesses director and site practices for ensuring family outreach, communication, and engagement. The scale was collected from directors and owners in sampled sites, which allows comparison of mean scores by type (Table F. 13). There are four subscales present: Environment and Policy Checklist, Communication Systems, Information about Resources, and Referrals. Directors associated with Early Childhood Education classrooms generated the highest scores in the Environment and Policy Checklist, Information about Resources, and Referrals subscales. Center Directors exhibited the lowest scores on the Environment and Policy Checklist and Information about Resources subscales, but not the Referrals subscale (for which directors from homes generated the lowest mean score).

Table F. 13. Mean FPTRQ (Director Measure) scores by type²³

		Aggregate	Centers	Homes	Pre-k Classrooms
Environment and Policy Checklist	Mean	13.2	12.8	13.4	15.6
	SD	2.7	2.6	3.0	1.4
	n	60	44	9	7
Communication Systems	Mean	7.9	7.8	7.9	7.9
	SD	1.2	1.2	.8	1.5
	n	62	46	8	8
Information about Resources	Mean	5.5	5.1	5.8	7.6
	SD	3.6	3.3	3.7	4.9
	n	68	48	13	7
Referrals	Mean	2.5	2.7	1.1	3.4
	SD	1.7	1.7	1.3	1.5
	n	72	50	13	9

Directors from sites in non-urban settings tended to report higher Environment and Policy Checklist and Communication Systems subscale scores, whereas directors from sites in urban settings tended to report higher Information about Resources subscale scores. There was a small difference in mean scores in the Referrals subscale.

Table F. 14. Mean FPTRQ (Director Measure) scores by urban versus non-urban setting

		Urban	Non-Urban
Environment and Policy Checklist	Mean	12.8	13.6
	SD	2.8	2.6
	n	29	31
Communication Systems	Mean	7.6	8.1
	SD	1.5	.8
	n	29	33
Information about Resources	Mean	5.8	5.2
	SD	3.8	3.3
	n	36	32
Referrals	Mean	2.6	2.5
	SD	1.7	1.8
	n	38	34

²³ It is important to note the relatively small samples sizes for homes and pre-kindergarten classrooms, when reviewing this and related tables.

FPTRQ-Teacher Measure

The FPTRQ (Teacher Measure) assesses teacher knowledge, practices, and attitudes with respect to family outreach, communication, and engagement. The scale was collected from teachers in sampled centers and Early Childhood Education classrooms, which allows comparison of mean scores by type (Table F. 15).

Teachers associated with Early Childhood Education classrooms generated higher scores in the Knowledge and Practices subscales. Teachers from center and Early Childhood Education classrooms exhibited very similar mean scores on the Attitudes subscale, as shown in Table F. 15.

Table F. 15. Mean FPTRQ (Teacher Measure) scores by type

		Aggregate	Centers	Early Childhood Education Classrooms
Knowledge	Mean	30.4	30.3	31.9
	SD	6.8	6.7	7.6
	n	253	238	15
Practices	Mean	72.8	72.4	78.5
	SD	11.7	11.5	12.8
	n	230	217	13
Attitudes	Mean	54.9	54.9	54.5
	SD	4.4	4.4	4.7
	n	242	231	11

Teachers from sites in non-urban settings tended to report higher Knowledge subscale scores, whereas teachers from sites in urban settings tended to report higher Practices and Attitudes subscale scores. In one instance, Attitudes, the difference in mean score was statistically significant ($p < .019$).

Table F. 16. Mean FPTRQ (Teacher Measure) scores by urban versus non-urban setting

		Urban	Non-Urban
Knowledge	Mean	29.9	30.9
	SD	6.9	6.6
	n	126	127
Practices	Mean	73.4	72.1
	SD	10.3	12.8
	n	112	118
Attitudes*	Mean	55.6	54.2
	SD	4.4	4.3
	n	119	123

*statistically significant difference at $p < .019$

FPTRQ-Parent Measure

The FPTRQ (Parent Measure) captures parent feedback on site practices for ensuring family outreach, communication, and engagement. The measure was collected from parents who consented to participate in the study, from sampled centers and Early Childhood Education classrooms.

As with the Teacher Measure, there are three subscales: Knowledge, Practices, and Attitudes. In each subscale, parents whose children attended Early Childhood Education classrooms reported higher scores than parents whose children attended center-based classrooms (Table F. 17).

Table F. 17. Mean FPTRQ (Parent Measure) scores by type

		Aggregate	Centers	Early Childhood Education Classrooms
Knowledge	Mean	51.9	51.6	53.9
	SD	9.1	8.9	10.8
	n	173	152	21
Practices	Mean	100.9	99.6	111.6
	SD	19.5	19.1	19.7
	n	146	130	16
Attitudes	Mean	65.2	64.9	67.9
	SD	5.5	5.6	3.2
	n	167	147	20

There were very small differences between parents whose children attended sites in urban and non-urban settings, on the Knowledge subscale. In comparison, parents whose children attended sites in urban settings tended to report higher scores than parents in non-urban settings on the Practices and Attitudes subscales (Table F. 18). As with teacher FPTRQ scores, there were statistically significant differences in mean score on Attitudes ($p < .048$).

Table F. 18. Mean FPTRQ (Parent Measure) scores by urban versus non-urban setting

		Urban	Non-Urban
Knowledge	Mean	51.7	52.0
	SD	8.7	9.4
	n	65	108
Practices	Mean	101.7	100.5
	SD	20.6	18.9
	n	49	97
Attitudes*	Mean	66.3	64.7
	SD	4.0	6.1
	n	59	108

*statistically significant difference at $p < .048$

FCCERS-R

The FCCERS-R was implemented with participating Type A and B homes. In addition to the overall mean score, the study team calculated six subscale scores, including Space and Furnishings, Personal Care Routines, Listening and Talking, Activities, Interactions, and Program Structure.

Examining results across subscales, participating sites exhibited the highest mean score on the Interactions subscale, and the lowest mean score on the Personal Care Routines subscale. Differences between Type A and B homes were relatively small on most subscales (Table F. 19).

Table F. 19. Mean FCCERS-R score by type

		Overall (n=11)	Type A (n=3)	Type B (n=8)
Space and Furnishings	Mean	3.1	3.3	3.0
	SD	1.1	2.1	.6
Personal Care Routines	Mean	2.5	2.4	2.6
	SD	.7	.4	.7
Listening and Talking	Mean	4.3	3.8	4.5
	SD	1.8	2.4	1.7
Activities	Mean	2.6	2.5	2.7
	SD	.9	1.1	.8
Interactions	Mean	4.9	4.4	5.1
	SD	1.4	1.5	1.4
Program Structure	Mean	3.1	3.0	3.1
	SD	1.3	.3	1.5
Mean FCCERS-R Score	Mean	3.2	3.0	3.2
	SD	.8	1.1	.8

When examining scores by urban or non-urban setting, there is a consistent trend for sites in non-urban settings to exhibit higher mean scores than sites in urban settings (Table F. 20).

Table F. 20. Mean FCCERS-R score by urban versus non-urban setting

		Urban (n=6)	Non-Urban (n=5)
Space and Furnishings	Mean	2.7	3.5
	SD	.7	1.3
Personal Care Routines	Mean	2.4	2.8
	SD	.8	.4
Listening and Talking	Mean	4.1	4.7
	SD	1.8	1.97
Activities	Mean	2.4	2.9
	SD	.7	1.0
Interactions	Mean	4.6	5.4
	SD	1.6	1.1
Program Structure	Mean	2.7	3.5
	SD	.8	1.7
Mean FCCERS-R Score	Mean	2.9	3.5
	SD	.7	.9

CIS

The CIS was implemented with Type A and B homes, as there is not a CLASS variant validated for use in family child care homes. The CIS has four subscales: Sensitivity, Harshness, Detachment, and Permissiveness.

As shown in Table F. 21, providers scored relatively high scores on each subscale. Of the four subscales, the lowest scores were exhibited in the Permissiveness subscale, suggesting that providers struggle the most with permissive behaviors and setting boundaries with or for children.

Table F. 21. Mean CIS score by type

		Overall (n=17)	Type A (n=5)	Type B (n=12)
Sensitivity	Mean	3.6	3.5	3.7
	SD	.3	.4	.2
Harshness	Mean	3.7	3.7	3.7
	SD	.3	.5	.3
Detachment	Mean	3.8	3.7	3.8
	SD	.4	.7	.3
Permissiveness	Mean	3.0	3.0	3.0
	SD	.3	--	.4
Overall Mean Score	Mean	3.7	3.6	3.7
	SD	.3	.5	.2

There was a tendency for providers in non-urban settings to have a slightly higher overall mean score and scores on the subscales, with the exception of the Permissiveness subscale (Table F. 22).

Table F. 22. Mean CIS score by urban versus non-urban setting

		Urban (n=9)	Non-Urban (n=8)
Sensitivity	Mean	3.6	3.7
	SD	.3	.2
Harshness	Mean	3.6	3.8
	SD	.4	.2
Detachment	Mean	3.7	3.9
	SD	.6	.3
Permissiveness	Mean	3.0	3.0
	SD	.3	.3
Overall Mean Score	Mean	3.6	3.8
	SD	.4	.2

CHELLO

The CHELLO was used in Type A and B homes to assess supports for language and literacy development. There are two subscales: Literacy Environment Checklist and Group/Family Observation, which are not comparable to each other.

Type A homes exhibited higher scores on the Literacy Environment Checklist but lower scores on the Group/Family Observation and overall mean score.

Table F. 23. Mean CHELLO score by type

		Overall (n=17)	Type A (n=5)	Type B (n=12)
Literacy Environment Checklist	Mean	22.12	23.0	21.8
	SD	3.8	3.9	3.9
Group/Family Observation	Mean	58.9	57.0	59.7
	SD	6.6	9.6	5.3
CHELLO	Mean	81.0	80.0	81.4
	SD	9.7	13.3	8.4

Homes in non-urban settings exhibited higher scores than homes in urban settings, on average, for both subscales and the overall mean score.

Table F. 24. Mean CHELLO score by urban versus non-urban setting

		Urban (n=9)	Non-Urban (n=8)
Literacy Environment Checklist	Mean	20.2	24.3
	SD	4.2	1.8
Group/Family Observation	Mean	55.2	63.0
	SD	6.98	2.7
CHELLO	Mean	75.4	87.3
	SD	10.2	3.3

PAS

The PAS was conducted with a subset of 14 child care centers. The PAS is a seven-point scale with 10 subscales. The highest scores were generated on the technology, child assessment, and family partnerships subscales. The lowest scores were generated on the personnel cost and allocation and staff qualifications subscales.

Table F. 25. Mean PAS score by type

		Aggregate (n=14)
Human Resources Development	Mean	3.5
	SD	1.8
Personnel Cost and Allocation	Mean	2.2
	SD	1.0
Center Operations	Mean	3.3
	SD	1.3
Child Assessment	Mean	4.8
	SD	2.0
Fiscal Management	Mean	2.9
	SD	2.3
Program Planning and Evaluation	Mean	3.0
	SD	1.7
Family Partnerships	Mean	4.8
	SD	1.9
Marketing and Public Relations	Mean	3.7
	SD	1.5
Technology	Mean	5.1
	SD	1.5
Staff Qualifications	Mean	2.3
	SD	1.2
Overall Score	Mean	3.4
	SD	1.1

Data source: Program Administrative Scale

SUTQ Validation Study Results

PAS scores also were examined by location. In overall mean score and each subscale (except center operations), centers in non-urban settings exhibited higher mean scores than centers in urban settings.

Table F. 26. Mean PAS score by urban versus non-urban setting

		Urban (n=7)	Non-Urban (n=7)
Human Resources Development	Mean	2.5	4.5
	SD	1.4	1.6
Personnel Cost and Allocation	Mean	1.9	2.5
	SD	1.2	.9
Center Operations	Mean	3.3	3.3
	SD	1.4	1.4
Child Assessment	Mean	4.1	5.4
	SD	2.4	1.3
Fiscal Management	Mean	2.6	3.1
	SD	2.1	2.6
Program Planning and Evaluation	Mean	2.5	3.6
	SD	1.5	1.8
Family Partnerships	Mean	4.6	4.9
	SD	2.1	1.7
Marketing and Public Relations	Mean	3.2	4.1
	SD	1.5	1.5
Technology	Mean	4.4	5.7
	SD	1.6	1.2
Staff Qualifications	Mean	1.8	2.8
	SD	1.1	1.1
Overall Score	Mean	3.0	3.9
	SD	1.1	.99

Data source: Program Administrative Scale

BAS

The BAS was conducted with a subset of 13 Type A and B homes. The BAS is a seven-point scale with 9 subscales calculated for the current study. The highest overall scores were generated on the work environment and provider-parent communication subscales. The lowest overall scores were generated on the income and benefits and fiscal management subscales. In overall mean score and most subscales, Type B homes generated higher mean scores than Type A homes (with the exception of the risk management subscale).

Table F. 27. Mean BAS score by type

		Overall (n=13)	Type A (n=4)	Type B (n=9)
Qualifications and Professional Development	Mean	4.3	3.8	4.6
	SD	2.3	2.1	2.4
Income and Benefits	Mean	2.6	2.0	2.9
	SD	1.8	1.4	1.9
Work Environment	Mean	6.5	6.0	6.7
	SD	1.1	2.0	.5
Fiscal Management	Mean	2.9	1.5	3.6
	SD	2.4	.6	2.6
Recordkeeping	Mean	5.5	4.8	5.9
	SD	1.5	1.5	1.5
Risk Management	Mean	3.9	4.8	3.6
	SD	2.1	2.2	2.1
Provider-Parent Communication	Mean	5.7	5.3	5.9
	SD	1.5	2.2	1.2
Community Resources	Mean	5.4	4.5	5.8
	SD	1.9	2.9	1.4
Marketing and Public Relations	Mean	4.8	4.0	5.2
	SD	2.1	2.9	1.6
Overall Score	Mean	4.6	4.1	4.9
	SD	1.1	.9	1.2

Data Source: Business Administrative Scale

SUTQ Validation Study Results

BAS scores also were examined by location. In addition to overall mean score, sites in non-urban settings generated higher scores on some subscales (income and benefits, work environment, fiscal management, community resources, and marketing and public relations), while in other subscales, sites in urban settings were higher, on average (qualifications and professional development and provider-parent communication).

Table F. 28. Mean BAS score by urban versus non-urban setting

		Urban (n=7)	Non-Urban (n=6)
Qualifications and Professional Development	Mean	4.6	4.0
	SD	2.6	1.9
Income and Benefits	Mean	2.3	3.0
	SD	1.98	1.5
Work Environment	Mean	6.0	7.0
	SD	1.4	.0
Fiscal Management	Mean	2.7	3.2
	SD	2.3	2.6
Recordkeeping	Mean	5.6	5.5
	SD	1.5	1.6
Risk Management	Mean	4.0	3.8
	SD	2.5	1.8
Provider-Parent Communication	Mean	5.9	5.5
	SD	.9	2.1
Community Resources	Mean	5.1	5.7
	SD	1.9	2.2
Marketing and Public Relations	Mean	3.6	6.3
	SD	2.1	.5
Overall Score	Mean	4.4	4.9
	SD	1.2	1.0

Data Source: Business Administrative Scale

Teacher Education and Experience

Teacher education was retrieved from participating teachers and directors and coded to reflect the attainment of a two-year or higher degree (versus some college, High School completion or GED only, or less than High School completion). As shown in Table F. 29, Early Childhood Education classrooms were staffed with teachers who reported a two-year or higher degree, while approximately 60% of centers were staffed with teachers with the equivalent levels of education. Teachers in Early Childhood Education classrooms, on average, had more experience than teachers in child care centers.

Table F. 29. Overall teacher education by type

	Mean (n=154)	Centers (n=146)	Early Childhood Education Classrooms (n=8)
Percent of teachers with less than a college degree	38.3%	40.4%	0%
Percent of teachers with more a two year or four year college degree	61.7%	59.6%	100%
Mean years of experience	10.9 (SD=7.8; n=92)	10.6 (SD=7.7; n= 85)	13.7 (SD=9.3; n=7)

When examined by location, there were relatively small differences in educational attainment by teachers in urban versus non-urban settings. Similarly, there was almost no difference in mean years of experience in teachers in these different settings.

Table F. 30. Teacher education by urban versus non-urban setting

	Urban (n=79)	Non-Urban (n=75)
Percent of teachers with less than a college degree	40.5%	36%
Percent of teachers with more a two year or four year college degree	59.5%	64%
Mean years of experience	10.9 (SD=7.5; n=41)	10.8 (SD=8.1; n=51)

Comparison of Kindergarten Readiness Assessment Scores for Children Who Did and Did Not Participate in Star-Rated Programs during their Pre-kindergarten Year

Working with both Department of Education and Job and Family Services datasets, the study team isolated children who (a) participated in school-supported Early Childhood Education (ECE) programs; (2) participated in Preschool Special Education (PSE) programs; and (3) participated in publicly funded child care. In addition, the study team used Department of Education records to isolate students who were considered economically disadvantaged (but for whom there wasn't a record of participation in ECE, PSE, or publicly funded child care²⁴) and students who were not considered economically disadvantaged.

As shown in Tables Table F. 31 and Table F. 32, children who were not considered economically disadvantaged received the highest scores, on average, on the KRA. They were followed by children who participated in ECE programs and children who participated in publicly funded child care. Children who were considered economically disadvantaged, but did not have a record of participating in ECE, PSE, or publicly funded child care, and children participating in PSE programs consistently exhibited the lowest mean scores. The differences across groups were statistically significant, as explained in Table F. 33.

Table F. 31. KRA mean scores, disaggregated by pre-kindergarten experience: 2014-2015

		Children who participated in ECE in their pre-kindergarten year (n=5299)	Children who participated in PSE in their pre-kindergarten year (n=7451)	Children who were coded as both ECE/PSE in their pre-kindergarten year (n=678)	Children who participated 1 week or more in subsidy program (n=10686)	Children who were considered economically disadvantaged but did not participate in the ECE, PSE, or subsidy programs (n=42090)	Children who were not considered economically disadvantaged (n=47859)	ANOVA
Language & Literacy	Mean	266.71	258.87	259.01	263.38	261.57	271.22	p<.000
	SD	11.6	15.9	14.1	11.5	12.6	11.9	
Mathematics	Mean	267.29	259.71	259.84	263.86	262.82	272.88	p<.000
	SD	13.0	17.0	14.1	12.5	13.8	13.4	
Social Foundations	Mean	269.13	257.42	258.95	264.40	264.45	273.88	p<.000
	SD	18.3	20.4	19.0	18.3	18.7	17.8	
Physical Development and Well-Being	Mean	269.81	257.57	259.47	266.23	265.37	273.13	p<.000
	SD	16.3	20.4	18.8	17.0	17.4	15.7	
Overall Test Score	Mean	266.67	257.92	258.52	263.00	262.08	271.35	p<.000
	SD	11.3	14.9	12.5	10.8	11.8	11.5	

²⁴ This group of children may have participated in child care during their pre-kindergarten year, including child care that was star rated. Information about their pre-kindergarten experiences was not available.

Table F. 32. KRA mean scores, disaggregated by pre-kindergarten experience: 2015-2016

		Children who participated in ECE in their pre-kindergarten year (n=4589)	Children who participated in PSE in their pre-kindergarten year (n=6853)	Children who were coded as both ECE/PSE in their pre-kindergarten year (n=699)	Children who participated 1 week or more in subsidy program (n=4228)	Children who were considered economically disadvantaged but did not participate in the ECE, PSE, or subsidy programs (n=47062)	Children who were not considered economically disadvantaged (n=49043)	ANOVA
Language & Literacy	Mean	267.35	259.96	261.73	263.39	261.90	271.29	p<.000
	SD	11.9	15.1	13.8	11.8	13.0	12.2	
Mathematics	Mean	266.28	259.81	260.85	262.12	261.55	271.03	p<.000
	SD	12.7	15.4	14.2	12.0	13.5	13.3	
Social Foundations	Mean	273.36	261.20	266.48	267.56	267.51	277.67	p<.000
	SD	19.1	22.0	22.4	19.2	20.0	18.3	
Physical Development and Well-Being	Mean	272.11	260.16	264.43	267.62	267.00	275.26	p<.000
	SD	16.5	20.6	20.4	17.1	17.7	15.7	
Overall Test Score	Mean	267.70	259.39	261.76	263.28	262.60	271.98	p<.000
	SD	11.6	14.5	14.0	11.0	12.2	11.9	

Table F. 33. Differences in KRA scores, by participation in ECE, PSE, or publicly funded child care

	2014-2015	2015-2016
Language & Literacy	Children in ECE programs versus all other groups; children in PSE programs versus children in ECE programs, children in publicly funded child care, economically disadvantaged children who did not participate in ECE or publicly funded child care, and children who were not economically disadvantaged; children who participated in publicly funded child care versus all other groups; children coded to both ECE and PSE versus children in ECE programs, children in publicly funded child care, children considered economically disadvantaged, and children not considered economically disadvantaged; children considered economically disadvantaged versus all other groups; and children not considered economically disadvantaged versus all other groups.	Children in all groups except children coded to both PSE and ECE and children considered economically disadvantaged
Mathematics	Children in all groups except children in PSE or PSE/ECE, children in PSE/ECE and children in publicly funded child care, and children in PSE/ECE and economically disadvantaged children.	Children in all groups except children in PSE/ECE and children in publicly funded child care; children in PSE/ECE and economically disadvantaged children; and economically disadvantaged children versus children not considered economically disadvantaged
Social Foundations	Children in all groups except children in PSE/ECE and children in publicly funded child care; children in PSE/ECE and economically disadvantaged children; and economically disadvantaged children versus children not considered economically disadvantaged	Children in all groups except economically disadvantaged children versus children not considered economically disadvantaged
Physical Development and Well-Being	Children in all groups except children in PSE/ECE and children in publicly funded child care; and children in PSE/ECE and children not considered economically disadvantaged	Children in all groups except children in PSE/ECE and children in publicly funded child care; and children in PSE/ECE and children not considered economically disadvantaged
Overall Test Score		Children in all groups except children in PSE/ECE and children in publicly funded child care; and children in PSE/ECE and children not considered economically disadvantaged

Further Examination of Kindergarten Readiness Assessment Scores for Children in Publicly Funded Child Care

When subsidy participants were examined by race/ethnicity, African-American and Hispanic students did not perform as well as Caucasian students or students categorized as “Other.” There were statistically significant differences across tests and racial/ethnic groups, as shown in Tables Table F. 34

Table F. 35.

Table F. 34. KRA mean scores for publicly funded children, disaggregated by race/ethnicity: 2014-2015

		African-American (n=4888)	Caucasian (n=3972)	Hispanic (n=787)	Other (n=1003)	ANOVA
Language & Literacy	Mean	262.15	265.21	260.85	264.61	p<.000
	SD	11.6	10.8	11.7	11.6	
Mathematics	Mean	262.43	266.15	261.32	264.25	p<.000
	SD	12.2	12.4	12.2	12.0	
Social Foundations	Mean	263.18	265.60	264.65	265.86	p<.000
	SD	18.7	17.7	17.7	17.6	
Physical Development and Well-Being	Mean	265.18	267.21	266.42	267.81	p<.000
	SD	17.2	16.5	16.4	16.9	
Overall Test Score	Mean	261.81	264.62	261.49	264.04	p<.000
	SD	10.9	10.4	10.4	10.6	

**some students were missing demographic data and are not included in this table*

Table F. 35. KRA mean scores for publicly funded children, disaggregated by race/ethnicity: 2015-2016

		African-American (n=1785)	Caucasian (n=1728)	Hispanic (n=289)	Other (n=404)	ANOVA
Language & Literacy	Mean	262.14	265.26	259.21	263.83	p<.000
	SD	11.9	11.2	12.9	11.0	
Mathematics	Mean	260.39	264.35	258.82	262.53	p<.000
	SD	11.6	12.0	11.8	11.9	
Social Foundations	Mean	266.04	268.85	266.83	269.13	p<.000
	SD	19.4	19.0	18.9	19.3	
Physical Development and Well-Being	Mean	266.45	268.30	267.73	269.50	p<.000
	SD	17.1	16.8	17.4	17.7	
Overall Test Score	Mean	261.93	264.87	260.91	264.03	p<.000
	SD	10.9	10.7	11.6	10.8	

Table F. 36. Differences in KRA scores, among publicly funded children, by race

	2014-2015	2015-2016
Language & Literacy	African-Americans versus all groups; Caucasians versus African-American; and Hispanic versus all groups	Children in all groups except Caucasians and "Other"
Mathematics	African-Americans versus Caucasians and "Other"; Caucasian versus all groups; and "Other" versus all groups	African-Americans versus Caucasians and "Other"; Caucasian versus Hispanic and "Other"; and Hispanic versus "Other"
Social Foundations	African-Americans versus Caucasians and "Other"	African-Americans versus Caucasians and "Other"
Physical Development and Well-Being	African-Americans versus Caucasians and "Other"	African-Americans versus Caucasians and "Other"
Overall Test Score	African-Americans versus Caucasians and "Other"; Caucasian versus Hispanic; and Hispanic versus "Other"	African-Americans versus Caucasians and "Other"; Caucasian versus Hispanic; and Hispanic versus "Other"

As with the general dataset, publicly funded students with disabilities did not perform as well on the KRA as publicly funded students who were not identified as having disabilities (Table F. 37). Independent samples t-tests confirmed that the differences in mean scores between publicly funded children with and without disabilities were statistically significant.

Table F. 37. KRA mean scores for publicly funded children, disaggregated by disability status

		2014-2015		2015-2016	
		Publicly funded children with disabilities (n=368)	Publicly funded children without disabilities (n=10282)	Publicly funded children with disabilities (n=178)	Publicly funded children without disabilities (n=4029)
Language & Literacy	Mean	253.78	263.77	254.89	263.76
	SD	13.2	11.2	13.3	11.5
Mathematics	Mean	255.26	264.22	255.16	262.42
	SD	14.6	12.2	13.4	11.8
Social Foundations	Mean	250.66	264.94	253.30	268.18
	SD	19.5	18.0	21.0	18.9
Physical Development and Well-Being	Mean	251.73	266.79	254.88	268.16
	SD	19.1	16.6	18.8	16.8
Overall Test Score	Mean	252.85	263.41	254.13	263.67
	SD	12.5	10.5	12.3	10.7

Appendix G

Ohio's Early Learning and Development Standards in All Essential Domains of School Readiness (Birth – Age 5)

INTRODUCTION

In December 2011, Ohio was awarded the Race to the Top Early Learning Challenge Grant. To be awarded the funding, Ohio was required to have *Early Learning and Development Standards in all Essential Domains of School Readiness, Birth to Age 5*. These five domains included:

- Social and Emotional Development
- Physical Well-being and Motor Development
- Approaches Toward Learning
- Language and Literacy Development
- Cognition and General Knowledge

Ohio's Early Learning and Development Standards describe key concepts and skills that young children develop during the birth-to-five-year period. Their purpose is to support the development and well-being of young children and to foster their learning. The standards promote the understanding of early learning and development, provide a comprehensive and coherent set of expectations for children's development and learning, and guide the design and implementation of curriculum, assessment and instructional practices with young children.

The standards present a continuum of learning and development from birth to age five in each of the domains. Because the infant/toddler years are marked by rapid developmental change, the standards are divided into three meaningful transitional periods: Infants (birth to around 8 months), Young Toddlers (6 to around 18 months), and Older Toddlers (16 to around 36 months). The standards during the preschool years describe those developmental skills and concepts children should know and be able to do at the end of their preschool experience.

The Ohio Early Learning and Development Standards were created as part of a collaborative effort of state agencies serving young children including Ohio Department of Education, Ohio Department of Job and Family Services, Ohio Department of Health, Ohio Department of Mental Health, Ohio Department of Developmental Disabilities, and the Governor's Office of Health Transformation. The state agencies worked with national experts and writing teams made up of Ohio-based content experts and stakeholders to revise and expand the standards in the five developmental domains.

Ohio's revision of standards builds upon the strong set of existing standards in Ohio's Infant and Toddler Guidelines (for children birth to 36 months of age) and the Pre-Kindergarten Standards (for children ages 3 to 5). Ohio's *Infant and Toddler Guidelines* was the major source for the development of the infants' and toddlers' standards. Similarly, Ohio's *Pre-Kindergarten Content Standards* were revised and expanded in the Language and Literacy and Cognitive Development domains. The Cognition and

General Knowledge standards were aligned with the kindergarten Common Core State Standards in English-Language Arts and Mathematics and Ohio’s Revised Academic Content Standards in Science and Social Studies. Finally, the standards were reviewed and revised with particular attention to being appropriate for children with disabilities and for children with diverse cultural and linguistic backgrounds. Knowledge of the strengths and needs of each child is pertinent in order to implement differentiation strategies and culturally responsive pedagogy in a manner to help each child meet the standards.

ORGANIZATION OF THE STANDARDS

The standards within each domain are organized according to **strands**, the developmental or conceptual components within each domain. Each strand contains one or more **topics**, the area of focus within each strand, and the **standard statements**, those concepts and skills children should know and be able to do for the different age-groups. Some topics reflect learning and development across the birth-to-five continuum, with standards for all age levels: infants, young toddlers, older toddler, and Pre-K, while other topics pertain only to a specific age- period. For example, some knowledge and skills such as *the ability to identify and describe shapes* or skills related to social studies and science emerge in preschool. Topics that address those competencies include standards only at the Pre-K level. Other topics such as *Self- Comforting* and *Social Identity* have standards only at the infant-toddler levels, because these foundational skills developed during the early years lead to more specific competencies at the preschool level.

AN OVERVIEW OF THE DOMAINS

Social and Emotional Development. The standards for Social and Emotional development involve behaviors that reflect children’s emotional growth and their growing ability to successfully navigate their social worlds through interactions with teachers and peers. These standards include a focus on children’s developing abilities to regulate attention, emotions, and behavior, and to establish positive relationships with familiar adults and with peers. Research indicates that early skills of social competence and self-regulation are foundational to children’s long-term academic and social success (National Research Council, 2008). Strands in the social and emotional domain are *Self* and *Relationships*.

Physical Well-Being and Motor Development Physical Well-Being and Motor Development standards address motor skills and health practices that are essential for children’s overall development. These skills include the ability to use large and small muscles to produce movements, to touch, grasp and manipulate objects, and to engage in physical activity. These standards also describe the development of health practices that become part of children’s daily routines and healthy habits such as nutrition and self-help. These skills and behaviors play an important role in children’s physical well-being and set children on a path leading toward a healthy lifestyle. Healthy children are more likely to attend school, to be physically active, and to learn more effectively (Bluemenshine and others, 2008). The two strands in this domain are *Motor Development* and *Physical Well-Being*.

Approaches Toward Learning. Approaches Toward Learning centers on the foundational behaviors, dispositions, and attitudes that children bring to social interactions and learning experiences. It includes children’s initiative and curiosity, and their motivation to participate in new and varied experiences and challenges. These behaviors are fundamental to children’s ability to take advantage of

learning opportunities, and to set, plan, and achieve goals for themselves. This domain also includes children’s level of attention, engagement, and persistence as they do a variety of tasks. These factors are consistent predictors of academic success (Duncan et al., 2007). Finally, children’s creativity, innovative thinking and flexibility of thought allow them to think about or use materials in unconventional ways, and to express thoughts, ideas and feelings in a variety of media. The standards in the domain Approaches Toward Learning are organized in the following strands: *Initiative; Engagement and Persistence; and Creativity*.

Language and Literacy. The standards for language and literacy reflect knowledge and skills fundamental to children’s learning of language, reading and writing. Young children’s language competencies pertain to their growing abilities to communicate effectively with adults and peers, to express themselves through language, and to use growing vocabularies and increasingly sophisticated language structures. Early literacy skills include children’s developing concepts of print, comprehension of age-appropriate text, phonological awareness, and letter recognition. Research has identified early skills of language and literacy as important predictors for children’s school readiness, and their later capacity to learn academic knowledge (National Early Literacy Panel, 2008). The Language and Literacy domain consists of the following strands: *Listening and Speaking, Reading and Writing*.

Cognition and General Knowledge. This domain includes those cognitive processes that enable all other learning to take place, as well as children’s knowledge of the social and physical world. This domain is organized into the strand, *Cognitive Skills* and those concepts and skills in **sub- domains**, *Mathematics, Social Studies and Science*.

Cognitive Skills. This strand refers to the underlying cognitive mechanisms, skills and processes that support learning and reasoning across domains, including the development of memory, symbolic thought, reasoning and problem-solving.

- **Mathematics.** The sub-domain of mathematics encompasses the mathematical concepts and skills that children develop during the birth-to-five-year period, including children’s developing understanding of number and quantity, number relationships, and basic algebraic concepts. A meta-analysis conducted by Duncan and colleagues (2007) suggests that specific early math skills such as knowledge of numbers and ordinality are important predictors of later achievement in math and reading. The Mathematics sub- domain also addresses children’s developing knowledge of key attributes of objects, including size and shape, and the way objects fit, are positioned, and move in space. The standards in the domain of mathematics are organized in four strands: *Number Sense, Number Relationships and Operations; Algebra; Measurement and Data; and Geometry*.
- **Social Studies.** The sub-domain of social studies includes basic skills and competencies that set the foundation for learning about concepts of social science. At a young age, children begin to develop their social identity and to think about their place in the social world. As they grow, they develop an increased awareness of their personal histories and heritage, and a sense of time and place. Through everyday interactions with children and adults, they develop an appreciation for rights and responsibility within a group, and how social rules help people in promoting safety and fairness (Mindes, 2005). Such competencies are described in the domain

of Social Studies under the following strands: *History; Geography; Government; and Economics*.

- **Science.** This sub-domain focuses on children's curiosity to explore and learn about their environment. It includes behaviors of exploration and discovery, and fundamental conceptual development such as problem solving and cause and effect. These early behaviors develop into increasingly systematic inquiry skills, and the ability to observe, investigate and communicate about the natural environment, living things, and objects and materials (Gelman and Brenneman, 2004). Early competencies in science are organized in four key strands: *Science Inquiry and Application; Earth and Space Science; Physical Science; and Life Science*.

Ohio's early learning and development standards illuminate the breadth of learning and development from birth to kindergarten entry that strengthens school readiness. An understanding of learning and development in each domain guides programs and teachers as they plan developmentally appropriate learning opportunities and environments for young children. In particular, teachers can use an understanding of standards to focus on the kinds of interactions and environments that support, for example, language development or approaches toward learning. While the standards facilitate a focused look at young children's learning in each domain, teachers and others responsible for the care and education of young children need to keep in mind that infants, toddlers, and preschool-age children learn holistically.

Moreover, social and emotional development stands at the center of their learning. For example, as an infant or toddler builds security in a relationship with a caring adult, that child is also learning to communicate with language and to use the relationship as a secure base for practicing new movement skills and building knowledge about the world through exploration. Likewise, as preschool-age children tell stories about family experiences they are expanding their self-awareness, using their growing cognitive capacity to remember the past, and practicing narrative skills. Such examples of integrated learning are endless. In addition to providing focused looks in each domain, the standards can help us see how learning occurs in different domains at the same time.

Teachers and others can use the standards as starting points for observing and understanding young children's learning and development. With each learning encounter teachers observe, they can refer to the standards and ask what knowledge and skills are the children gaining in the areas of language and literacy, cognition and general knowledge, social and emotional development, physical well-being and motor development, and approaches toward learning. Teachers can use their observations of integrated learning to plan new learning encounters for young children and support the building of knowledge in all essential domains of school readiness.

References

- Bluemenshine, S. L. and others (2008). "Children's School Performance: Impact of General and Oral Health," *Journal of Public Health Dentistry*, Vol. 68, No. 2, 82-87.
- Duncan, G.J. et al. (2007). School readiness and later achievement. *Developmental Psychology*, 43 (6), pp. 1428-1466.
- Gelman, R., and K. Brenneman. 2004. "Science Learning Pathways for Young Children." *Early Childhood Quarterly Review* 19:150–58.
- Mindes, G. (2005). "Social Studies in Today's Early Childhood Curricula," *Beyond the Journal: Young Children on the Web*, Vol. 60, No. 5, 12-18.
- National Early Literacy Panel (2008). *Developing Early Literacy: A Scientific Synthesis of Early Literacy Development and Implications for Intervention*.
- National Education Goals Panel (1995). *Reconsidering children's early development and learning: Toward common views and vocabulary*. Washington DC: Author.
- National Research Council (2008). *Early Childhood Assessment: Why, What, and How*. Committee on Developmental Outcomes and Assessment for Young Children, C. E. Snow and S.B. Van Hemel, Editors. Board on Children, Youth, and Families, Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.






Appendix H

Standards for Programs Department of Education Department of Job and Family Services



STAR RATING SYSTEM

One-, two- and three-star programs must meet all standards at the lower levels before advancing to the next. Beyond three stars, programs will be awarded points based on their adherence to three-star requirements and their ability to earn additional points. Programs must achieve a minimum of at least one point in each of the four domains and can earn extra points for accreditation, lower staff/child ratios and group size.

Domains	ONE STAR 	TWO STARS 	THREE STARS 	 FOUR AND FIVE STARS 
LEARNING & DEVELOPMENT (MAX 21 POINTS)	Curriculum & Planning Program engages in a process to identify a written, research-based, comprehensive curriculum aligned with the Early Learning and Development Standards and/or Ohio's K-12 Standards (appropriate to the age groups served) that is developmentally appropriate. Each classroom has a copy of the Early Learning and Development Standards and/or Ohio's K-12 Standards (appropriate to the age groups served). Teachers utilize a written, dated plan of activities.	Program obtains a written, research-based, comprehensive curriculum aligned with the Early Learning and Development Standards and/or Ohio's K-12 Standards (appropriate to the age groups served) that is developmentally appropriate. Each teacher has daily access to a copy of the curriculum. Teachers use a written, dated plan of activities that is aligned to all developmental domains in the Early Learning and Development Standards and/or Ohio's K-12 Standards (appropriate to the age groups served).	Program implements a written, research-based, comprehensive curriculum aligned with the Early Learning and Development Standards and/or Ohio's K-12 Standards (appropriate to the age groups served) and demonstrates its alignment to assessment. Teachers use a written, dated plan of activities that is aligned to all developmental domains in Early Learning and Development Standards and/or Ohio's K-12 Standards (appropriate to the age groups served).	Teachers plan intentional and purposeful activities and experiences that meet the needs/interests/abilities of children across all developmental domains (5 points). Teacher supports children's active engagement through opportunities for exploration and learning (1 point).
	Child Screening & Assessment Program engages in a process to select a comprehensive, developmentally appropriate screening instrument(s) for the age groups served. (Does not apply to school-age children.) Program identifies staff to be trained to administer, score and use the instrument(s) appropriately.	Program obtains a comprehensive, developmentally appropriate screening instrument(s) for the age groups served. (Does not apply to school-age children.) Staff are trained to administer, score and use the instrument(s) appropriately.	Program ensures that all children (except school-age children) receive a comprehensive, developmental screening that is valid and reliable within 60 business days of entry into the program and annually thereafter. Necessary referrals are completed within 90 days of identification of need, and the results are formally communicated with families. Program administers assessments that meet state requirements for all enrolled preschool-aged children. Program assesses all children enrolled using formal and informal methods on an ongoing basis to inform instruction. Results are shared with families.	Ongoing child assessment results are used to make, adjust and refine instructional decisions and to evaluate child progress (5 points). Families are provided multiple opportunities to understand the assessment process used and data collected, review and contribute to their child's education plan, and share information about their child's progress toward learning goals (4 points).
	Instructional Environment Program selects a self-assessment tool(s) that addresses the quality of the classroom environment and staff/child interactions that is developmentally appropriate for each age group served.	Staff complete an annual self-assessment that addresses the quality of the classroom environment and staff/child interactions for each group of children.	Each lead teacher identifies an area for improvement and creates an action plan, which includes a goal and action steps toward completing the goal, based on the results of the annual self-assessment. Program supports each child's development by providing well-structured learning environments and positive interactions among staff and children.	Lead teacher documents progress on action steps and readjusts goals as needed (4 points).

Step 4 and Step 5 programs must score at least one point in each of the four domains. They can score additional points by meeting the standards listed below.
 To achieve a Step 4 designation, programs must score 35-55 points.
 To achieve a Step 5 designation, programs must score 56-79 points.

ADDITIONAL POINTS

Accreditation Accredited from an approved accrediting body (5 points).	Preschool • 36 to < 48 months 1:10 / 2:20 and 48 months to < school age 1:12 / 2:24 (2 points); or • 36 months to < 48 months 1:11 / 2:22 and 48 months to < school age 1:13/2:26 (1 point)
Staff/Child Ratio and Group Size Infants • Birth to < 18 months 1:4 / 2-8 / 3:10 (3 points) or 1:5 / 2:10 (2 points)	Mixed Age Groups follow the ratio for the youngest child in the group School age • K to age 14 1:15 (2 points); or 1:16 (1 point)
Toddlers • 18 to 30 months 1:5 / 2:10 • 30 to 36 months 1:6 / 2:12 (3 points); or • 18 to 30 months 1:6 / 2:12 • 30 to 36 months 1:7 / 2:14 (2 points)	Mixed Age Groups follow the ratio for the youngest child in the group

SUTQ Validation Study Results

Domains	ONE STAR	TWO STARS	THREE STARS	FOUR AND FIVE STARS
ADMINISTRATIVE & LEADERSHIP PRACTICES (MAX 18 POINTS)	Staff Support The program has a written wage structure.	Program offers one of the approved staff supports to employees.	Program offers two of the approved staff supports to employees.	Program offers three or more of the approved staff supports to employees (3 points).
	Program Administration Program completes an annual program self-assessment.	Program completes a continuous improvement plan, which includes a minimum of one goal and action steps toward completing the goal, using results from self-assessment.	Program completes an annual continuous improvement plan, which includes a minimum of two goals and action steps toward completing the goals, using results from the self-assessment.	Program's annual continuous improvement plan has strategies to engage community partners to support child and family outcomes (3 points). The program conducts an annual survey with families. Identified stakeholders and/or community partners to review accomplishment of program goals (3 points).
	Staff Management Administrators, lead teachers and assistant teachers have annual professional development plans.	Lead and assistant teachers have at least one formal observation annually.	Results of the formal observations are used to inform individual professional development plans.	Results of the formal observations are used to inform the program's annual continuous improvement plan (4 points). Results of annual classroom self-assessment are used to inform individual professional development plans (3 points). Two formal observations are completed annually for all lead and assistant teachers (2 points).
STAFF QUALIFICATIONS & PROFESSIONAL DEVELOPMENT (MAX 17 POINTS)	Staff Education Administrator has Child Development Associate (CDA) or Career Pathways Level (CPL) 2. 50 percent of lead teachers have a CDA (at least two lead teachers must have a CDA or higher to meet this standard), or one lead teacher has an associate's degree (AA) appropriate to the age groups noted below or a CPL 3. • Early Childhood Teachers—An AA in early childhood education (ECE) or an approved related field. • School-Age Teachers—An AA in an approved field.	Administrator has an AA in ECE (or approved related field for school-age-only programs) or a CPL 3. 25 percent of lead teachers have an AA appropriate to the age groups noted below or a CPL 3. • Early Childhood Teachers—An AA in ECE or an approved related field. • School-Age Teachers—An AA in an approved field.	Administrator has an AA in ECE (or approved related field for school-age-only programs) or CPL 3. 50 percent of lead teachers have an AA appropriate to the age groups noted below or a CPL 3. • Early Childhood Teachers—An AA in ECE or an approved related field. • School-Age Teachers—An AA in an approved field.	Administrators • Master's degree in ECE or approved related field, or bachelor's degree (BA) with CPL 5, or BA in ECE with an Administrator Credential Level (ACL) 3 (5 points); or • BA in ECE or approved related field with an administrator license, or AA in ECE or approved related field with an ACL 3 (4 points); or • BA in ECE or approved related field, or CPL 4, or AA in ECE or approved related field with a CPL 3 (3 points); or • Associate's degree in ECE or approved related field with an ACL 3 (2 points). Lead Teachers • 100 percent have a BA in ECE or approved related field or a CPL 4 (5 points); or • 50 percent have a BA in ECE or approved related field or a CPL 4, and the other 50 percent have a minimum of an AA appropriate to the age groups noted in Step 3 or a CPL 3 (4 points); or • 100 percent have an AA appropriate to the age groups noted in Step 3 or a CPL 3 (3 points). Assistant Teacher (applies only to groups that need an assistant teacher in order to maintain ratio) • 100 percent of classrooms have an assistant teacher with a CDA or CPL 2 (4 points); or • 50 percent of classrooms have an assistant teacher with an AA in ECE (or related field for school-age teachers) (3 points); or • 75 percent of classrooms have an assistant teacher with a CDA or CPL 2 (2 points); or • 50 percent of classrooms have an assistant teacher with a CDA or CPL 2 (1 point).
	Professional Development Administrators, lead teachers and assistant teachers receive a minimum of 20 clock hours of approved specialized training every two years. Individuals who currently are enrolled in a degree-granting program in early childhood or related field can use coursework to fulfill the annual training requirement.	Administrators, lead teachers and assistant teachers receive a minimum of 20 clock hours of approved specialized training every two years. Individuals who currently are enrolled in a degree-granting program in early childhood or related field can use coursework to fulfill the annual training requirement.	Administrators, lead teachers and assistant teachers receive a minimum of 20 clock hours of approved specialized training every two years. Individuals who currently are enrolled in a degree-granting program in early childhood or related field can use coursework to fulfill the annual training requirement.	100 percent of administrators and lead and assistant teachers have obtained 30 or more hours of approved professional development over a two-year period (3 points); or 100 percent of administrators and lead and assistant teachers have obtained 25 or more hours of approved professional development over a two-year period (2 points); or 50 percent of administrators and lead and assistant teachers have obtained 30 or more hours of approved professional development over a two-year period (2 points); or 50 percent of administrators and lead and assistant teachers have obtained 25 or more hours of approved professional development over a two-year period (1 point).
	Transitions Program provides written information to families on transitioning children into, within and out of the program.	Program provides age-appropriate activities for children to prepare them for the transition to a new classroom or educational setting. Program transfers any child's records to the new setting at the family's request and with the family's written consent.	Program meets with families to develop an individualized transition plan that supports a child's transition to another classroom or educational setting.	The program has written transition policies and practices that include strategies for supporting transitions into, within and out of the program for both children and families (2 points).
FAMILY & COMMUNITY PARTNERSHIPS (MAX 8 POINTS)	Communication & Engagement Program obtains information about the family structure and routines that is important to the child's development. Program provides information regarding resources and community services to families.	Program communicates with families using different modes of communication, at least one of which promotes two-way communication. At least once per year, information on topics addressing health and child development is provided to families. Program offers at least one opportunity for families to engage in activities.	Program organizes at least one educational training, workshop or event to support families' engagement in children's learning and development. Program has written policies and procedures to ensure that children have received comprehensive health screenings and/or that families have been provided information on the importance of health screenings and resources to obtain them. Programs and parents collaborate to create annual written, developmental and/or educational goals for children.	Program has written documentation of formal and/or informal agreements with community partners and other family-serving agencies, programs and entities (3 points). Program uses a formal model or process to enhance family engagement strategies (2 points). Program has an organized and active parent volunteer group (1 point).