

University of New Hampshire

University of New Hampshire Scholars' Repository

Master's Theses and Capstones

Student Scholarship

Spring 2023

Evaluating the Equivalency and Ecological Validity of TPITOS and ITERS-3 Ratings of Quality in New Hampshire's Quality Recognition and Improvement System

Eileen Elizabeth Murphy
University of New Hampshire, Durham

Follow this and additional works at: <https://scholars.unh.edu/thesis>

Recommended Citation

Murphy, Eileen Elizabeth, "Evaluating the Equivalency and Ecological Validity of TPITOS and ITERS-3 Ratings of Quality in New Hampshire's Quality Recognition and Improvement System" (2023). *Master's Theses and Capstones*. 1714.

<https://scholars.unh.edu/thesis/1714>

This Thesis is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Master's Theses and Capstones by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact Scholarly.Communication@unh.edu.

**Evaluating the Equivalency and Ecological Validity of TPITOS and ITERS-3 Ratings of
Quality in New Hampshire's Quality Recognition and Improvement System**

By

Eileen E. Murphy

Bachelor of Architecture, University of Notre Dame, 2009

THESIS

Submitted to the Thesis Committee

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

Human Development and Family Studies

May 2023

This thesis was examined and approved in partial fulfillment of the requirements for the degree of Master of Science in Human Development and Family Studies by:

Thesis Director, Kimberly Nesbitt, Ph.D., Associate Professor,
Human Development and Family Studies

Lisa Ranfos, Ed.D., Clinical Associate Professor,
Human Development and Family Studies

Tyler Jamison, Ph.D., Associate Professor,
Human Development and Family Studies

On April 13, 2023

Approval signatures are on file with the University of New Hampshire Graduate School.

DEDICATION

To my niece, Monroe, who was born four months after I started graduate school, and my older niece and nephew, Rylin and Bradley: Watching you develop and grow keeps me curious and present.

To the children I was lucky enough to care for at Little Sprouts, Growing Places, and my many babysitting engagements: Your journeys, and the joys and challenges your families faced inspire me to do more for young children.

ACKNOWLEDGEMENTS

I'd like to thank Dr. Kimberly Nesbitt and my thesis committee for their support and feedback during the 15 months it took to plan, execute, and report on this project. I'd also like to thank all faculty and staff in the Department of Human Development and Family Studies, the Department of Education, and at the Graduate School at the University of New Hampshire who helped me through my 21 months in the master's program. I could never have imagined how much I would learn and grow over the course of my time here.

To my family: Thank you for always being there.

TABLE OF CONTENTS

COMMITTEE PAGE	ii
DEDICATION	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vii
LIST OF APPENDIX TABLES	vii
LIST OF FIGURES	viii
ABSTRACT	ix
INTRODUCTION	1
Infant and Toddler Development	1
Quality in Early Childhood Care and Education	5
Quality Rating/Recognition and Improvement Systems	8
New Hampshire’s QRIS: <i>Granite Steps for Quality</i>	13
SUMMARY OF THE CURRENT STUDY.....	22
Research Questions	23
METHODOLOGY	23
Participants	23
Procedures	27
Data Analysis	33
RESULTS	37
Equivalency Between TPITOS and ITERS-3 Scores	37
Difference in Scores Related to Exposure	39

Ecological Validity	40
DISCUSSION	42
Granite Steps for Quality	42
Infants and Toddlers in QRISs	43
QRIS Objectives	44
Limitations	45
Implications for Future Research	46
CONCLUSION	47
REFERENCES	48
APPENDICES	58
APPENDIX A: <i>Granite Steps for Quality</i> Pathways, Standard 2	59
APPENDIX B: IRB Approval Letter	61
APPENDIX C: Center and Classroom Surveys	63
APPENDIX D: Detailed Exposure Data	68
APPENDIX E: <i>Granite Steps for Quality</i> Incentives	71

LIST OF TABLES

Table 1: Comparison of <i>Granite Steps for Quality</i> Pathways for Standard 2: Learning Environments	17
Table 2: Comparison of Measurement Tools Used in <i>Granite Steps for Quality</i>	19
Table 3: Characteristics of Participants	26
Table 4: Exposure to the Pyramid Model, TPITOS or TPOT.....	29
Table 5: Exposure to the Environment Rating Scales, ITERS or ECERS	29
Table 6: Comparisons of <i>Granite Steps for Quality</i> Step Ratings According to TPITOS and ITERS-3 Scores as Defined in the Pyramid Model and Environment Rating Scales Pathways	36
Table 7: Comparison of Rankings for TPITOS Score and ITERS-3 Score	38

LIST OF APPENDIX TABLES

Table A1: ERS/ITERS-3 Pathway	59
Table A2: Pyramid Model/TPITOS Pathway	60
Table D1: Teacher’s Exposure to the Pyramid Model, TPITOS or TPOT	68
Table D2: Teacher’s Exposure to the Environment Rating Scales, ITERS or ECERS	69
Table D3: Center’s Exposure to the Pyramid Model	70
Table D4: Center’s Exposure to the Environment Rating Scales	70
Table E1: Quarterly Quality Incentive	71
Table E2: Monthly Child Care and Development Fund Scholarship-based Incentive	71

LIST OF FIGURES

Figure 1: Quality of Care for Infants and Toddlers Research-Based	
Conceptual Model	6
Figure 2: Association Between TPITOS Scores and ITERS-3 Scores	37

ABSTRACT

Quality in early childhood care and education is increasingly being monitored through state-run Quality Rating/Recognition and Improvement Systems (QRISs), however, the design of these systems is under-researched. This study explores the alignment of two observation tools used in a QRIS to assess the quality of infant and toddler care and education – the Teaching Pyramid Infant-Toddler Observation Scale (TPITOS), and the Infant/Toddler Environment Rating Scale, 3rd edition (ITERS-3). These two tools are central to the two pathways for center-based care providers in New Hampshire’s QRIS, which are designed to be equivalent but different methods for measuring quality. In this study, simultaneous TPITOS and ITERS-3 observations were conducted in 10 infant and toddler classrooms in New Hampshire. Quality ratings, according to the QRIS, were determined from scores on each of the tools. A regression analysis indicates a positive association between scores on the two tools, but an analysis of the quality ratings shows that New Hampshire’s QRIS lacks ecological validity. Recommendations for policy and research are provided.

Introduction

From birth to 3 years old, children experience critical growth and development which lay the foundation for lifelong cognitive, mental, and physical health (Center on the Developing Child at Harvard University, 2022; Shonkoff & Phillips, 2000). More than six million (52.9%) infants and toddlers in the United States spend time in child care settings (Keating et al., 2021). Theory and research suggest the quality of the relationships infants and toddlers have with caregivers can significantly impact their development (Morris & Williamson, 2018). Public awareness of the importance of these relationships in early childhood care and education is increasing, but the literature on measuring and improving quality, specifically in infant and toddler classrooms, is scarce. This study aims to contribute to our understanding of measuring and improving the quality of infant and toddler care and education.

This thesis research project describes the quality ratings and program characteristics of a sample of 10 infant and toddler classrooms in licensed child care centers in New Hampshire. The analysis includes an evaluation of the equivalency and ecological validity of the two pathways in New Hampshire's Quality Rating/Recognition and Improvement System and provides considerations for infant and toddler care and education policy.

Infant and Toddler Development

Modern understandings of best practices in early childhood care and education (ECCE) derive from Bronfenbrenner's ecological theory, attachment theory, Vygotsky's sociocultural learning theory, and Piaget's cognitive development theory (Burchinal, 2018; Cohen & Waite-Stupiansky, 2017; Miller, 2016). Ecological systems theory explains that a child's family, community, and broader culture influence their development (Bronfenbrenner, 1986). This highlights the importance of the caregiving environment and the relationships a child forms with

peers and adults in out-of-home care. Attachment theory explains that the behavior of the adults who interact with the child will impact the child's social and emotional development (Ainsworth & Bowlby, 1991). Young children thrive when they trust the adult caregivers around them. Trusting relationships provide a forum for the development of language, and secure bonds with adults give children the confidence to explore their environment and interact with their peers.

The learning theories presented by Jean Piaget and Lev Vygotsky explain cognitive development and skill acquisition (Cohen & Waite-Stupiansky, 2017). Piaget's theory explains that as children age, they are capable of increasingly complex thought (Miller, 2016; Piaget & Inhelder, 1969). When given age-appropriate materials and opportunities, typically developing children are driven to make sense of the world through play and experimentation. Vygotsky's sociocultural learning theory includes the concepts of *scaffolding* and the *zone of proximal development* (Fogel, 1997; Mooney, 2000). Scaffolding refers to the way adults and more skilled children who interact with a less skilled child can help facilitate learning. The zone of proximal development refers to the next skill or developmental level a child is going to acquire. The adult or more skilled child can guide the young child through scaffolding (assistance or modeling) to help them practice the skill until they can independently perform the skill. From these theories, it has been established that children benefit from caregivers who thoughtfully observe their behavior, respond to their needs, provide developmentally appropriate materials, and provide scaffolding to teach new skills.

Theories of child development and early learning situate the young child at the center of a world of opportunity. The physical environment and materials available play an important role in a child's learning. However, the relationships children have with their caregivers – parents, extended family, and non-relative caregivers – are the child's point of entry to the physical

environment (Parlakian, 2003). This is especially true for the youngest infants who are immobile and dependent on adults for interaction with materials (Legerstee et al., 2013). Relationships with important adults in the infant or toddler's life play a vital role in cognitive, social, emotional, and physical development (Morris & Williamson, 2018; National Scientific Council on the Developing Child, 2004; Reis et al., 2000).

As theories of developmental psychology became widely accepted, the latter half of the 20th century saw the emergence of a number of approaches to the care and education of young children. These include the Reggio Emilia approach and Maria Montessori's approach (Follari, 2019). During this period, Magda Gerber and Emmi Pikler began collaborating on an approach specific to infant and toddler caregiving, termed *educaring* (Cohen & Waite-Stupiansky, 2017). Drawing from existing theories of development and learning, with a particular emphasis on respect for the young child, Gerber and Pikler studied the needs of young children in care settings. Later, Gerber worked with others to establish a philosophical approach for professional caregivers of infants and toddlers, called *Resources for Infant Educators* (Cohen & Waite-Stupiansky, 2017).

Gerber's 1998 book describes the main principles of her approach: trust in the child, involve the child in all caregiving activities, allow time for uninterrupted play and free movement, allow for play with peers, create peaceful surroundings, provide varied, simple, open-ended objects, create organized and comfortable spaces, provide primary caregivers and continuity of care, conduct sensitive observation and selective intervention, set clear and consistent limits through authenticity and emotional honesty, provide boundaries and guidance (Gerber & Johnson, 1998). Thoughtful interactions with the child and an intentional design of the learning environment are emphasized in Gerber's work. These qualities further highlight the

importance of the ECCE setting as a context in which the child can be given the opportunity for optimal growth and development.

The Risk of Stress and the Importance of Early Childhood Care and Education

Positive, nurturing relationships lead to learning, growth, and optimal developmental outcomes, while negative or neglectful relationships can have lasting detrimental effects on a child (Center on the Developing Child at Harvard University, 2022). With this in mind, modern infant and child psychologists are exploring how a child's learning and development are promoted or hindered by their social context, including their experiences in early care and education settings.

Research has shown that abuse and neglect put infants and toddlers at an increased risk of mental health issues and developmental delays or disabilities which include emotional, behavioral, and cognitive challenges (Center on the Developing Child at Harvard University, 2022). With more infants and toddlers in the United States spending more time than ever before in ECCE, it is vital that caregivers in these settings do not harm the children with whom they are entrusted (Chaudry & Sandstrom, 2020; Keating et al., 2021).

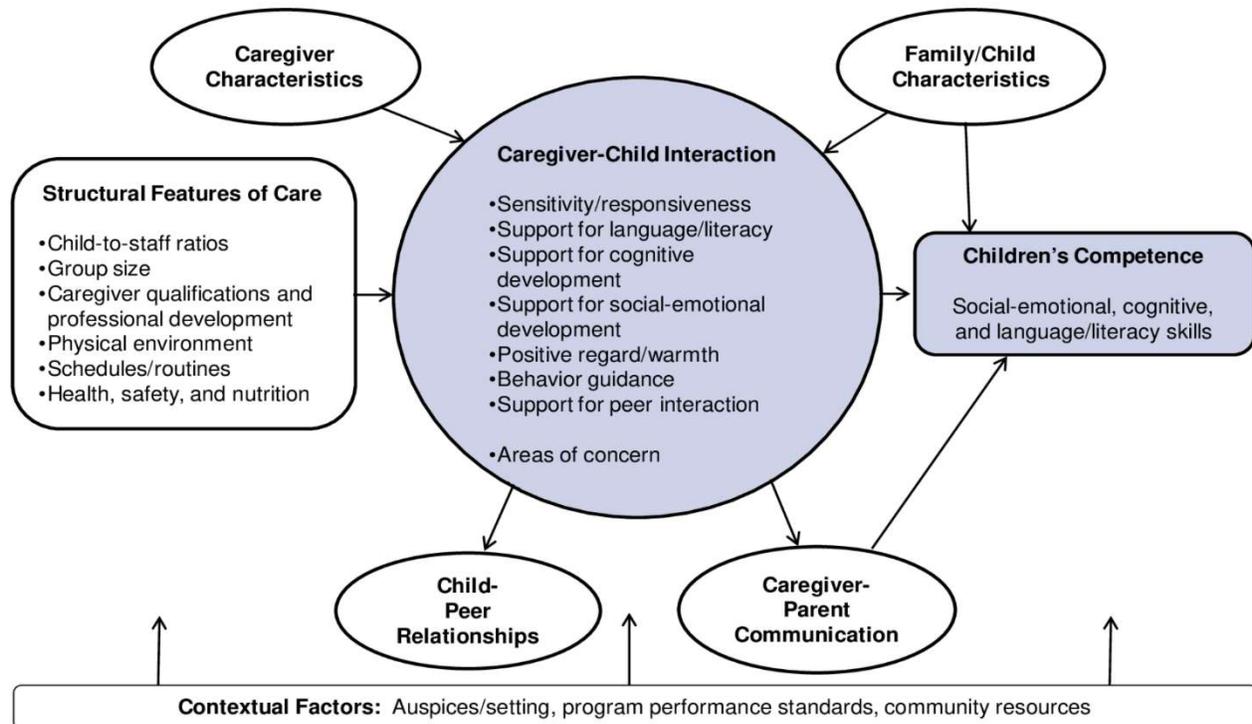
That said, the evidence regarding the benefits of ECCE for infants and toddlers is mixed. Some studies suggest that high-quality ECCE may serve as a protective factor for at-risk infants and toddlers (Berry et al., 2014; Burchinal et al., 1996). Other research – including one study with toddlers – indicates that children who attend high-quality ECCE are significantly better prepared for K-12 education (McCartney, 1984; Parlakian, 2003). More recent research involving infants and toddlers has shown positive associations between high-quality interactions with caregivers and child outcomes, particularly language, social and emotional development (Araujo et al., 2019; Elicker et al., 2021; Mortensen & Barnett, 2015; Pinto et al., 2019).

Meanwhile, studies that include preschool-age children (3-5 years old) have shown only modestly positive cognitive and social-emotional outcomes for children in ECCE (Burchinal, 2018; Peisner-Feinberg et al., 2001; Ruzek et al., 2014). Other studies involving preschoolers fail to show a correlation between ECCE and long-term cognitive and social-emotional outcomes (Durkin et al., 2022; National Institute of Child Health and Human Development Early Child Care Research, 2000; Phillips et al., 1987).

Some of this evidence is promising, but inconsistent findings point to the need for further research on the effect of infant and toddler care and education programs. At this time, it can be said definitively that stress harms infants and toddlers, and – at a minimum – avoiding undue stress will lead to better outcomes (Center on the Developing Child at Harvard University, 2022). With this in mind, scholars have sought to define high-quality care and education that ensures young children are not harmed and that may encourage optimal development.

Quality in Early Childhood Care and Education

In 2020, Banghart et al. conducted a comprehensive review of the literature on access and quality in infant and toddler care. Due to the complexity of infant and toddler development, which occurs across multiple care contexts and relationships at home and in out-of-home care, child outcomes are very difficult to predict (La Paro et al., 2014; Ragni et al., 2021). Without knowing all the environmental and relational influences a child encounters, using child outcomes to measure quality in ECCE is challenging. However, positive outcomes for children are precisely the objective of parents, ECCE providers, and policymakers. Figure 1 shows just how complex the web of a child's experiences is, and the role of ECCE quality in this system.

Figure 1*Quality of Caregiver-Child Interactions for Infants and Toddlers Conceptual Model*

Note. Reproduced from Halle, T., Anderson, R., Blasberg, A., Chrisler, A., & Simkin, S. (2011).

Quality of Caregiver-Child Interactions for Infants and Toddlers (Q-CCIT): A review of the literature (OPRE 2011-25). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

In an effort to organize the variables in ECCE, a general convention in research and policy is to divide quality in ECCE into two domains: structural features and process features (Burchinal, 2018; Zaslow et al., 2010). The *Quality of Caregiver-Child Interactions for Infants and Toddlers Conceptual Model* lists structural features on the left, and process features are listed in the central grey circle.

Structural features of quality are aspects of the caregiving environment and decisions about the design of an ECCE program that aim to improve the experience for children in the program. Typically, structural features are described in state licensing rules and then implemented and monitored by ECCE program directors and administrators. Structural features include limiting staff-to-child ratios, limiting group size, minimum caregiver qualifications, professional development requirements, creating developmentally appropriate physical spaces, establishing schedules and routines, ensuring health and safety, and promoting nutrition (Halle et al., 2011). Structural features are relatively easy to measure and track for the regulation and oversight of ECCE programs, however, research indicates that structural features are not as highly correlated with child outcomes as process features (Mashburn et al., 2008; Slot et al., 2015). Moreover, it is important to note that ECCE teachers and caregivers – especially in center-based programs – often do not have control over structural features of quality aside from their professional qualifications and ongoing professional development.

In contrast to structural features of quality, process features of quality are related to the day-to-day behavior and actions of teachers and caregivers. ECCE program directors and administrators may design classroom curricula or monitor classroom culture, but process quality is fundamentally what occurs in the interactions between teachers and children (Banghart et al., 2020). Some aspects of high-quality teacher-child interactions are sensitivity and responsiveness, language and cognitive stimulation, positive regard or warmth, behavior guidance, and support for peer interactions (Halle et al., 2011). When defining the quality of interactions, it can be useful to label what they are not. Researchers often call attention to indicators of low-quality care that may be observed (Bigelow et al., 2019; Halle et al., 2011). These negative indicators include detachment, intrusiveness, and negative regard. Again, negative, stressful situations are

associated with negative child outcomes (Center on the Developing Child at Harvard University, 2022).

The nuances of teacher-child relationships and interactions make monitoring and improving process quality challenging, but necessary. Each of the more than six million infants and toddlers in ECCE in the United States is depending on their interactions with a teacher or caregiver to achieve optimal development (Keating et al., 2021; Reis et al., 2000). There are at least one million caregivers of infants and toddlers in the United States who are trusted to develop responsive and nurturing relationships with children. State and federal governments are responding to the need to monitor process quality with the design and implementation of Quality Rating/Recognition and Improvement Systems or QRISs (Administration for Children & Families, n.d.).

Quality Rating/Recognition and Improvement Systems

History

Quality Rating/Recognition and Improvement Systems (QRISs) were originally developed in response to a restructuring of the Child Care and Development Fund in the 1990s (Cannon et al., 2017). In 1998, the Child Care and Development Block Grant was established to manage the use of the Child Care and Development Fund (Child Care Technical Assistance Network, n.d.). As the federal government prepared to pass funding down to states to support and expand child care and school-age care programs for low-income families, it wanted to ensure the programs receiving funding were providing high-quality services. The Child Care and Development Block Grant had two objectives: to support parents who are working toward independence from public assistance and to improve the quality of ECCE and school-age care for all children. The federal government saw that low-quality care was a serious issue across the

United States, and research had proven this was particularly problematic for children from backgrounds that put them at risk of poor developmental outcomes (Center on the Developing Child at Harvard University, 2022; Phillips et al., 1987). Thus, the Child Care and Development Block Grant put pressure on states to develop systems for measuring and improving quality and provided a financial incentive for their efforts (Cannon et al., 2017).

QRISs cover both ECCE, birth to age 5, and school-age care programs (Administration for Children & Families, n.d.). A QRIS is a system that aims to assess the quality of a program, incentivize quality improvement initiatives, and communicate quality ratings to the general public (Cannon et al., 2017; Mitchell, 2005). Essentially, it is the state's way of measuring and monitoring quality in care programs, while encouraging programs to better themselves. QRISs exist in 44 states and the District of Columbia and usually make use of observation-based rating scales to measure a program's level of quality (Administration for Children & Families, n.d.; Build Initiative, 2023). While most states use the phrase, Quality "Rating" and Improvement System, the state of New Hampshire calls their QRIS a Quality "Recognition" and Improvement System (Child Care Aware of NH, 2022). Throughout this paper and in practice, the acronym QRIS is used to refer to both rating and recognition systems.

QRISs aim to ensure quality in ECCE above and beyond the requirements of licensing. The desired outcome of an effective QRIS is improvement in the social, emotional, language, and cognitive development of children in ECCE programs (Cannon et al., 2017). Some studies suggest that process features of quality are associated with child outcomes for children 3 to 5 years old, thus QRISs usually emphasize process quality (Administration for Children & Families, n.d.; Mashburn et al., 2008; Slot et al., 2015). Turning toward infants and toddlers, Gerber's educaring approach and the 2003 ZERO TO THREE publication, *Before the ABCs:*

Promoting School Readiness in Infants and Toddlers show a new focus on social and emotional objectives with this age group (Gerber & Johnson, 1998; Parlakian, 2003). The importance of building confidence and nurturing social-emotional competence has come to the forefront of the scholarly dialogue in recent years, but its origins lie in decades-old theories of developmental psychology. These concepts are rooted in Erik and Joan Erikson's Psychosocial Development Theory (Erikson & Erikson, 1981). This theory explains that infants and toddlers are highly sensitive to their interactions with caregivers, and their confidence and identity evolve through these interactions. It is now widely understood that healthy relationships and interactions are critical to infant and toddler development (Araujo et al., 2019; Cohen & Waite-Stupiansky, 2017; Mortensen & Barnett, 2015; Pinto et al., 2019; Reis et al., 2000)

While there is not enough research to say definitively that promoting healthy social and emotional development in the infant and toddler years leads to long-term benefits such as high school completion or employment, scholars agree that social and emotional learning is a precursor to and prerequisite for language and cognitive skill acquisition (Parlakian, 2003). When measuring the efficacy of a QRIS in infant and toddler programs, the desired outcome should be social and emotional competence. Thus, QRISs and professional development programs would be most effective if they focus on the teacher's *interactions* with infants and toddlers.

Observation-Based Assessments

Observation tools are used in 93% of QRISs to assess teachers' performance as demonstrated by their interactions with children (Mitchell, 2005; National Center on Early Childhood Quality Assurance, 2017). Given the relational nature of infant and toddler care and education, observation-based tools have been thought of as the most effective method for

evaluating teachers of this age group (Thomason & La Paro, 2009). The widespread use of observation tools may also be explained by the relative simplicity of a two- or three-hour observation conducted by a trained third-party assessor. The assessor is presumably objective and an expert in child development, though these two assumptions are rarely verified (Mashburn, 2017).

Recent research suggests that while observations are convenient, they may not be the most valid or reliable evaluation method. In a study published in 2020, researchers noted that it is common for the allocation of funding in ECCE to depend primarily on observational assessments (Thorpe et al., 2020). This places significant weight on the reliability and validity of observations, but as this article highlights, there is minimal reliability and validity data available. This large-scale study analyzed 11,341 observations conducted at different times in 2,306 classrooms for children between 3 and 8 years old using the Classroom Assessment Scoring System, a tool commonly used to assess ECCE process quality. Findings show that the time of day and the activities underway during observations predict a classroom's quality rating. Specifically, quality ratings across all domains in the Classroom Assessment Scoring System steadily decrease as the school day progresses. Quality ratings were higher during large and small group activities, and lower during meal times, physical activity, and transitions (Thorpe et al., 2020). This demonstrates the limitations of observational assessments.

QRIS Efficacy

Observation-based assessments aside, the overall effect of QRISs on the quality of child care in the United States is debatable (Hestenes et al., 2015). In a policy report by Cannon et al. (2017), the authors point out that we have collectively put a significant amount of time, effort, and funding into these systems without having data to support whether they are improving

outcomes for children in ECCE (and school-age care). This report critiques the “first generation” of QRISs designed in response to the Child Care and Development Block Grant and outlines four specific suggestions for the “second generation.” Its recommendations, which shed light on the shortcomings of early QRIS designs, are: QRIS developers should direct more resources to researching and utilizing evidence-based quality improvement methods, for example, professional development such as practice-based coaching; states or the federal government should develop or identify better and more affordable measures for assessing classroom or program quality, possibly video-recorded observations; states should develop simple QRIS rubrics focused on only the most relevant aspects of quality to avoid creating an unnecessary administrative burden for participants; and the state should bear the cost of quality improvements and maintenance dictated by its QRIS (Cannon et al., 2017; Fukkink & Tavecchio, 2010; Helmerhorst et al., 2017).

In support of QRISs, a 2021 empirical study by Elicker et al. determined that a toddler’s attendance in a QRIS-rated high-quality ECCE program was associated with better developmental trajectories. While this is important research, studies that focus on toddler outcomes and testing associations between QRIS ratings and child outcomes are scarce. This can partly be attributed to the complexities of child development described earlier in this review. In addition to associations with QRIS ratings, Elicker et al. (2021) also found correlations between caregiver education and toddlers’ developmental trajectories. Before concluding that QRISs are effective at improving infant and toddler outcomes, more research that supports these results is needed. As we will see in the following sections, QRISs are complex, multi-layered systems designed to measure, recognize, and improve quality. It is difficult to say at this point in history if QRISs have an effect and, if so, which aspects are most impactful.

QRIS Structure

As a state-level system, QRISs currently come in 45 different shapes and sizes (Build Initiative, 2023). Despite their diversity, QRISs generally have some basic elements in common: tiered standards, program accountability, program and teacher supports and incentives, and a plan to inform parents of quality ratings (Mitchell, 2005). *Program accountability* refers to the documentation programs are required to submit to the state as well as observations or other assessments required in classrooms. *Program and teacher supports and incentives* refer to technical assistance for program directors and administrators, and coaching or training for teachers. An important aspect of the QRIS is that it encourages and supports program improvement when a program is rated as low-quality. If a program does not improve, presumably its low ratings would be advertised, and it would be avoided by families. In theory, eventually, low-quality programs would shut down due to low enrollment (Cannon et al., 2017). However, this rarely happens due to the nationwide child care shortage (Haspel, 2019). Another factor that complicates the impact of QRISs is that they are voluntary in most states. Some of the lowest quality programs may choose not to enroll in quality rating systems, thus QRISs would not provide a full picture of the quality of care available across a state.

New Hampshire's QRIS: *Granite Steps for Quality*

History

In 2005, New Hampshire's early childhood experts worked with the New Hampshire Department of Health and Human Services to release the state's first QRIS (Build Initiative, 2023). In 2012, experts, stakeholders, and state officials began working on a revision to the original system. In 2020, with the University of New Hampshire's receipt of the Preschool Development Grant, state, academic, and community partners came together to design a revised

QRIS. The revised system, called *Granite Steps for Quality*, was released in March 2022 and updated in October 2022 (Child Care Aware of NH, 2022).

Structure

Granite Steps for Quality, a voluntary QRIS, has guidelines for four types of programs that provide care to children in New Hampshire: family child care programs, center-based early childhood programs, school-age care programs, and center-based programs providing both early childhood and school-age care services (Child Care Aware of NH, 2022). The focus of the present study is infant and toddler classrooms within center-based early childhood programs (also referred to as center-based ECCE).

Prerequisites for *Granite Steps for Quality* participation are: the program holds an active child care license from the NH Department of Health and Human Services Child Care Licensing Unit for at least 12 months prior to its application; the program administrator completes specific training tasks; and the program is enrolled to receive Child Care and Development Fund (CCDF) scholarship payments on behalf of eligible families (Child Care Aware of NH, 2022). According to the *Granite Steps for Quality* guide, Head Start and Early Head Start center-based programs, which are free to qualified families, do not have to accept CCDF scholarships.

Enrollment in *Granite Steps for Quality* gives a program access to guidance from the state and begins the process of earning a quality designation (Child Care Aware of NH, 2022). To achieve a quality rating, called a *Quality Step* in *Granite Steps for Quality*, programs are evaluated across two standards: Standard 1: Staff Qualifications and Standard 2: Learning Environments. A program selects a Quality Step it hopes to achieve, then submits documentation to the state showing it has met the requirements of both Standard 1 and Standard 2 for that Quality Step. Once the state confirms the program's Quality Step rating, it is posted on the

program's profile in the NH Child Care Search Portal maintained by the state (https://new-hampshire.my.site.com/nhccis/NH_ChildCareSearch).

Each type of care program (family child care, center-based ECCE, and school-age care) has rubrics for Standard 1 and Standard 2. Standard 1 lists minimum staff credentials for each Quality Step. For Standard 2, an Environment Rating Scales (ERS) observation tool and cutoff scores are specified for each program type (i.e., the Family Child Care Environment Rating Scale, the School-Age Care Environment Rating Scale, etc.). Uniquely, for Standard 2 for center-based ECCE, *Granite Steps for Quality* offers two rubrics, called *Pathways*: an Environment Rating Scales Pathway and a Pyramid Model Pathway. A center has the opportunity to choose which Pathway it would like to follow.

Center-Based ECCE Standard 1: Staff Qualifications. The Staff Qualifications outlined in Standard 1 for center-based ECCE are the same regardless of the Pathway chosen. This Standard requires an administrator to hold a current NH Early Childhood Credential and requires increasing percentages of eligible staff to also hold a current NH Early Childhood Credential (from 20% of eligible staff to 40%, 60%, and 80%, respective to Steps 1 to Step 4). At Step 3 there is an additional requirement for 30% of credentialed staff to hold a minimum state teaching credential, which is voluntary in NH. Step 3 allows Head Start programs to meet this criterion by presenting documentation of Teacher Qualifications from their most recent Office of Head Start Monitoring Report. Step 3 also allows National Association for the Education of Young Children (NAEYC) accredited programs to meet this criterion by presenting their current accreditation certificate. Step 4, the highest level in *Granite Steps for Quality*, increases this requirement to 40% of credentialed staff to hold the highest teaching credential available in NH.

Center-Based ECCE Standard 2: Learning Environment. The Learning Environment criteria outlined in Standard 2 for center-based ECCE programs vary greatly between rubrics for the Environment Rating Scales Pathway and Pyramid Model Pathway. Firstly, each Pathway has its own observation tools. The Environment Rating Scales tools are the Infant/Toddler Environment Rating Scale, 3rd Edition (ITERS-3) and the Early Child Environment Rating Scale (ECERS) for preschool classrooms. The Pyramid Model tools are the Teaching Pyramid Infant-Toddler Observation Scale for Infant and Toddler Classrooms (TPITOS) and the Teaching Pyramid Observation Tool for Preschool Classrooms (TPOT). These tools are used to assign a score to the classroom or teacher that is the subject of an observation. A side-by-side comparison of the ITERS and TPITOS Pathways can be seen in Table 1. In the TPITOS Pathway, training and administrative requirements increase with each increasing Step, while in the ITERS Pathway, the same training is required for Steps 1 through 4. The detailed rubrics for Standard 2: Learning Environments can be found in Appendix A, Tables A1, and A2.

Table 1*Comparison of Granite Steps for Quality Pathways for Standard 2: Learning Environments*

Quality Step	Environment Rating Scales Pathway	Pyramid Model Pathway
1	Training requirement	Training requirement
2	Same training requirement as Step 1 ECERS/ITERS-3 self-assessment in all classrooms	Additional training requirements Administrative requirements The program is assigned a coach and establishes a leadership team
3	Same training requirement as Step 1 ECERS/ITERS-3 self-assessment in all classrooms ERS observations are completed in one-third of classrooms <ul style="list-style-type: none"> • Average score of at least 3.5 • All classrooms at least 2.5 on the subscales: Activities and Interactions. <p style="text-align: center;"><i>Head Start and NAEYC Accredited Programs can present documentation to achieve Quality Step 3</i></p>	Additional training requirements Additional administrative requirements TPOT/TPITOS observations and practice-based coaching are occurring in at least one classroom in the program.
4	Same training requirement as Step 1 ECERS/ITERS-3 self-assessment in all classrooms ERS observations are conducted in one-half of classrooms <ul style="list-style-type: none"> • Average score of at least 4.5 • All classrooms at least 3.5 on the subscales: Activities and Interactions 	Additional training requirements beyond Step 3 Leadership team collecting and analyzing TPOT/TPITOS data for continuous improvement to show Pyramid Model practices are being implemented to fidelity TPOT/TPITOS observations occurring with at least three teachers

Note. Adapted from the center-based ECCE rubrics from Child Care Aware of NH. (2022).

Granite Steps for Quality. NH Connections. <https://www.nh-connections.org/providers/quality-care-matters/>

A main difference between the two Pathways in Standard 2 is the amount of time and effort each Quality Step requires. The Pyramid Model Pathway requires more hours of training, more internal data collection, and more administrative effort than the ERS Pathway (Child Care

Aware of NH, 2022). This may lead program directors to gravitate toward the simpler ERS Pathway, but whether the variance in Pathway design influences program participation is unknown. While this would be an interesting area of inquiry, it is too early in the implementation of *Granite Steps for Quality* to evaluate the implications of the design.

Granite Steps for Quality Observation Tools: The ITERS-3 and TPITOS

For all types of care programs, and for both center-based ECCE Pathways, Quality Steps 3 and 4 require one or more observations by a trained, reliable, external observer. A designated number of teachers or classrooms must achieve a minimum score on the observation scale. The quality rating tools chosen for infant and toddler programs participating in *Granite Steps for Quality* are widely used, however, there is minimal research on their validity and reliability. This is partly due to their recent publication, and partly due to the lack of research on indicators of infant and toddler well-being, or assessments designed specifically for this population (Ryberg et al., 2022). That aside, the popularity of other tools and professional development programs created by their developers made them an appropriate choice. As an overview, Table 2 shows a comparison of some characteristics of the ITERS-3 and TPITOS.

Table 2

Comparison of Measurement Tools Used in Granite Steps for Quality

Characteristic	ITERS-3	TPITOS
Subject	One classroom	One teacher
Method	Observation	Observation and interview
Timeframe	No more than 3 hours, except to check environmental items	At least 2 hours for observation, 15 minutes for an interview
Ages of children	75% are 0-35 months	Most are 0-36 months
Teacher-child interaction scales ^a	Subscale 3: Language and Books, Subscale 5: Interactions	Entire scale

Note. Adapted from Bigelow, K. M., Carta, J. J., Irvin, D. D., & Hemmeter, M. L. (2019).

Teaching Pyramid Infant-Toddler Observation Scale (TPITOS) for infant-toddler

classrooms, research edition. Paul H. Brookes Publishing, Co. and Harms, T., Cryer, D.,

Clifford, R. M., & Yazejian, N. (2017). *Infant/Toddler Environment Rating Scale (Third).*

Teachers College Press.

^a *Granite Steps for Quality* has cutoff scores for the ITERS-3 Subscale 4: Activities and Subscale 5: Interactions.

ITERS-3. The Infant/Toddler Environment Rating Scale 3rd Edition (ITERS-3) is the most recent version of the Environment Rating Scale for infant and toddler classrooms released in 2017. It is described as a revision of the Infant/Toddler Environment Rating Scale-Revised and Updated (ITERS-R; Harms et al., 2017). The reliability of the ITERS-3 was tested by developers in a pilot study, and the manual lists reliability at 86.9% in simultaneous observations conducted by two raters. Unfortunately, testing of the validity of the ITERS-3 is less robust. This may be because the scale is just 5 years old and studies were likely impacted by the COVID-19 pandemic. Regardless of the cause, developers of the ITERS-3 defer to the validity of the

ITERS-R and the Early Childhood Environment Rating Scale 3rd Edition (ECERS-3), a measure designed for preschool and pre-K classrooms. They state the validity of these two scales is indicative of the probable validity of the ITERS-3 (Frank Porter Graham Child Development Institute, n.d.; Harms et al., 2017). This is a problematic assumption not often made in research, and it is made more so by some studies that suggest the validity of the ECERS-3 may not be significant (Early et al., 2018).

The lack of validity testing of the ITERS-3 raises questions about using this tool to determine funding for ECCE programs. While it would be ideal to have a more accurate tool for this important purpose, as of 2023, 27 out of the 45 states that have QRISs use Environment Rating Scales (Build Initiative, 2023). The Environment Rating Scales (ERS) is also a popular suite of products already in use by some ECCE programs for internal quality monitoring. In addition, the ERS's broad coverage of infant-toddler programs, family child care programs, and school-age programs as well as preschool and pre-kindergarten programs. This makes it an attractive option for a QRIS that aims to include all forms of Child Care and Development Block Grant funded care (Frank Porter Graham Child Development Institute, n.d.).

TPITOS. The Teaching Pyramid Infant-Toddler Observation Scale (TPITOS) is an observation-based assessment tool for infant and toddler teachers and is associated with the Pyramid Model for Promoting Social-Emotional Competence in Infants and Young Children developed by the Center for the Social and Emotional Foundations of Early Learning (Bigelow et al., 2019). The Teaching Pyramid Observation Tool for Preschool Classrooms (TPOT) was developed in 2014, while the TPITOS was developed in 2019. Similar to the ITERS-3, the TPITOS is relatively new, and testing on validity and reliability is sparse. A comprehensive

report on infant and toddler ECCE assessment tools published by the U.S. Administration for Children and Families in 2020 lists no existing research on the TPITOS (Shah et al., 2020).

As seen with the ITERS-3, the lack of validity testing of the TPITOS raises questions about using this tool to determine funding for ECCE programs. However, its association with the Pyramid Model professional development program and usefulness in individualized coaching make it an attractive option (Bigelow et al., 2019). The Pyramid Model approach focuses on nurturing social and emotional development through thoughtful interactions and provides guidance for teachers of young children with persistent challenging behavior. Research has shown that the Pyramid Model (also called Teaching Pyramid) training program improves the quality of teacher-child interactions, and helps teachers manage challenges when they arise (Hemmeter et al., 2016; Mincic et al., 2009). While not ideal, for the purposes of a QRIS, the TPITOS is a logical choice as it facilitates professional development and produces a numeric score. Despite the lack of validity and reliability testing, it was determined to be a valuable tool for New Hampshire's QRIS purposes and thus it is used in the present study.

Incentives

When the state confirms the program's application, completes observations and verifies the Quality Step rating, the program receives financial incentives. Financial incentives are based on the program's licensed capacity, its Quality Step, and the amount it bills monthly to the CCDF scholarship program. Its monthly billing to the scholarship program aligns with the number of eligible children it serves, their ages, and the amount of time they spend in care (NH Bureau of Child Development and Head Start Collaboration, 2019)

Summary of the Current Study

With a focus on New Hampshire's Quality Rating/Recognition and Improvement System, *Granite Steps for Quality*, two trained and reliable observers conducted simultaneous quality assessment observations in 10 infant and toddler classrooms across the state. These observations used the tools selected by the state of New Hampshire as part of *Granite Steps for Quality*. The tools are the Infant/Toddler Environment Rating Scale, 3rd edition (ITERS-3; Harms et al., 2017), and the Teaching Pyramid Infant-Toddler Observation Scale (TPITOS; Bigelow et al., 2019). The use of these assessments in New Hampshire's QRIS was evaluated using classroom data, teacher data, and Quality Step ratings according to the TPITOS scores and ITERS-3 scores.

Findings from the review of the literature were taken into consideration when designing the current study. Trained observers conducted simultaneous observations to avoid the variability caused by child, teacher, and center characteristics, time of day, and activities observed. This made it possible to compare the two scores from each program to one another. Between programs, there was inevitable variability. Where comparisons are made between programs, this study falls back on the assumptions of validity that all QRISs rely on.

The review of the design of *Granite Steps for Quality* Standard 2: Learning Environments brought to light the differences between the TPITOS and ITERS-3 Pathways. Since the training and administrative requirements are so different, this study bases Quality Step ratings on assessment scores alone. In addition, analyzing staff qualifications described in *Granite Steps for Quality* Standard 1 was beyond the scope of this study. For this study, it was assumed that a program could qualify for the same Quality Step in Standard 1 they are able to achieve in Standard 2.

Research Questions

This project addresses the following research questions:

- 1) Is there an equivalency between the TPITOS and ITERS-3 scores for each program?
 - a. If variability emerges, is there any relationship between teacher or center exposure to an observation tool and differential scores?
- 2) Do the TPITOS and ITERS-3 produce ecologically valid scores? Meaning, do both scores qualify centers for the same amount of funding from *Granite Steps for Quality*?

As reflected in the review of the literature, there is not enough extant research on the TPITOS or the ITERS-3 to form a hypothesis about directional effects. However, as higher scores on both the TPITOS and ITERS-3 are theorized to be indicators of higher quality, it is reasonable to expect that the scores will be positively associated. In addition, since the two Pathways in *Granite Steps for Quality* are designed to provide equivalent measures of quality and Quality Step ratings impact a program's reputation and incentive amounts, the scores from these two tools are expected to translate to the same Quality Step ratings in *Granite Steps for Quality*.

Methodology

Participants

Prior to recruitment and data collection, this project received approval from the University of New Hampshire Institutional Review Board (IRB-FY2022-365; see Appendix B for the approval letter). Teachers indicated informed consent through an online form emailed to them prior to their classroom observations. Data collected included the ECCE center, classroom, and teacher characteristics, teacher demographics, and teacher behaviors.

The researcher identified potential study participants through the University of New Hampshire's Preschool Development Grant partners, the researcher's contacts at the New Hampshire Bureau of Child Development and Head Start Collaboration, and through internet searches. This study included center-based ECCE programs currently licensed by the State of New Hampshire. The terms of the project funding restricted participation to nonprofit centers only. Program directors were initially contacted by email. If they were interested in participating, they were asked to provide the name of an infant or toddler classroom teacher who was willing to be the subject of the 3-hour observation. Some directors copied the participating teacher on email communications. Informed consent, surveys, and scheduling were coordinated by email with the director and included the teacher when an email address was provided. In two instances, the director and the participating teacher were the same individual due to staffing shortages.

Upon completion of the surveys and classroom observations, participating teachers and centers received gift card incentives. Participating teachers received a \$100 gift card for personal use, and center directors received a \$150 gift card intended for developmentally appropriate materials for the center.

Purposive sampling was used to select 10 centers that represent the diversity of center-based infant and toddler ECCE programs in New Hampshire. The intention was to test the alignment of the observation tools in a variety of settings. Participants spanned four counties, were located in both urban and rural municipalities, and included two National Association for the Education of Young Children (NAEYC) accredited centers and one Early Head Start center (US Census Bureau, 2020). Centers ranged in size from 18 to 155 enrolled children ($M = 59$, $SD = 40$). Refer to Table 3 for center and classroom characteristics. Eight out of the 10 participating centers accept Child Care Development Fund (CCDF) scholarships. The Early Head Start center

does not charge tuition so scholarships are not applicable. Of the eight who accept CCDF scholarships, the percentage of currently enrolled students who receive scholarships varied greatly from 0% to 50% ($M = 13.7\%$, $SD = 17.1\%$). Center directors reported weekly tuition for infants ranged from \$265.00 to \$370.00 ($M = \$330.25$, $SD = \$39.48$). Weekly tuition for toddlers ranged from \$220.00 to \$351.00 ($M = \$294.33$, $SD = \$41.84$).

Table 3*Characteristics of Participants*

ID	Center characteristics				Classroom characteristics			
	Classification ^a	Accreditation	Weekly tuition		Enrollment	Class Type ^b	Age range (months)	Child to staff ratio
			Infant	Toddler				
A	Rural		\$370	\$340	49	Infants/Toddlers	2-17	2.3
B	Rural		\$265	\$255	58	Toddlers	11-18	3.0
C	Urban		\$334 ^c	\$290	18	Two-year-olds	22-34	3.3
D	Urban		\$290	\$270	60	Two-year-olds	28-37	5.5
E	Rural		\$335	\$320	20	Toddlers	17-22	4.0
F	Urban	Early Head Start	\$0	\$0	38	Mixed age	7-35	2.5
G	Urban	NAEYC ^d	\$310	\$290	84	Infants/Toddlers	7-19	1.6
H	Urban		n/a	\$220	34	Two-year-olds	24-36	4.0
I	Urban	NAEYC	\$370	\$351	75	Two-year-olds	26-37	5.5
J	Urban		\$368	\$313	155	Infants	3-6	1.5

^a From the U.S. Census Bureau (US Census Bureau, 2020). ^b From the *Infant/Toddler Environment Rating Scale, Third Edition*

(Harms et al., 2017). ^c Program C is open Monday to Thursday. Tuition for four days (\$267) was converted to a five-day equivalent. ^d

NAEYC is the National Association for the Education of Young Children, which has a rigorous accreditation program. *Granite Steps for Quality* provides a shortcut to Step 3 for Early Head Start and NAEYC accredited programs.

Participating classrooms included groups of between three and 12 children. The ages of children spanned from 2 months to 37 months. According to the ITERS classification, this study included groups of infants through 2-year-olds (Harms et al., 2017). Specifically, there were infant groups where the majority of children are 0 to 11 months of age ($n = 1$), toddler groups where the majority are 12 months through 23 months ($n = 2$), infant and toddler mixed groups ($n = 2$), two-year-old groups where the majority are 24 months through 35 months ($n = 4$) and mixed age groups where the children range from infants to 2-year-olds ($n = 1$). Child-to-staff ratios were also diverse, ranging from 1.5 to 5.5 children per adult staff person ($M = 3.3$, $SD = 1.4$).

Participating teachers were all current lead or co-lead teachers in the classroom where the observations took place, according to titles given by the center. Two teachers observed were center directors who had been filling in as teachers for more than two months due to staffing shortages. Four of the 10 teachers observed hold a Director credential from the state, five teachers hold Lead Teacher credentials and one teacher does not hold a credential. Teacher education included “some college” ($n = 3$), completion of an associate degree ($n = 3$), completion of a bachelor’s degree ($n = 1$), and completion of a master’s degree ($n = 3$). Seven participants had degrees in early childhood education, three did not. Years of experience teaching ranged greatly from five to 36 years ($M = 16.8$, $SD = 10.9$). At the time of the study, teachers had been employed by between one and eight centers ($M = 3.0$, $SD = 2.4$).

Procedures

Center and Classroom Surveys

The researcher emailed center directors and participating teachers online surveys to complete prior to the classroom observation. Directors responded to a survey about the center as

a whole and participating teachers responded to a survey about their classroom. The complete surveys can be found in Appendix C. Surveys collected information on demographics as well as familiarity with the TPITOS and ITERS. Respondents indicated their experience with the Pyramid Model approach, the TPOT tool (Teaching Pyramid Observation Tool for Preschool Classrooms), the TPITOS tool, the Environment Rating Scales, the ECERS tool (Early Childhood Environment Rating Scale), and the ITERS tool. Teachers were asked about their personal experiences while directors responded on behalf of the center, (i.e., “What is the center’s experience with the Pyramid Model?”)

Coding Exposure Levels. The researcher used teacher’s and director’s responses to questions about the Pyramid Model, TPOT, or TPITOS and the Environment Rating Scales, ECERS, or ITERS to code each teacher and center as *low* or *high* exposure to each of the observation tools. Exposure is considered the teacher’s or center’s knowledge of or experience with the assessment tool or the teaching approach. For clarity, a respondent’s exposure to any training or assessment related to the TPITOS will be called “TPITOS exposure.” Exposure to any training assessment related to the ITERS will be called “ITERS exposure.” Exposure levels are shown in Tables 4 and 5.

Table 4*Exposure to the Pyramid Model, TPITOS or TPOT*

ID	Teacher exposure	Center exposure	TPITOS score	Standardized TPITOS score
A	Low	High	64.52	-1.32
B	Low	High	88.24	1.94
C	High	Low	75.68	0.22
D	High	High	71.23	-0.39
E	Low	Low	73.85	-0.04
F	High	High	78.38	0.59
G	Low	Low	73.53	-0.08
H	Low	Low	63.41	-1.47
I	High	High	80.26	0.85
J	Low	Low	71.93	-0.29

Note. For detailed exposure data, see Appendix D.**Table 5***Exposure to the Environment Rating Scales, ITERS or ECERS*

ID	Teacher exposure	Center exposure	ITERS-3 score	Standardized ITERS-3 score
A	High	Low	3.35	-1.27
B	Low	High	5.16	0.64
C	Low	Low	5.13	0.61
D	High	High	3.25	-1.37
E	Low	Low	3.81	-0.78
F	Low	Low	5.66	1.17
G	High	High	6.06	1.59
H	Low	Low	4.39	-0.16
I	High	Low	4.59	0.05
J	High	High	4.10	-0.47

Note. For detailed exposure data, see Appendix D.

Teachers who received individual Pyramid Model coaching, or who had previously had the TPITOS conducted on them, or a co-teacher, were coded as high exposure. Those who reported participating in multiple Pyramid Model trainings and who were also familiar with the TPITOS or TPOT observations were also coded as high exposure ($n = 4$). The remainder were coded as low exposure ($n = 6$).

The ITERS tool is older than the TPITOS and is more complex to administer and score (Bigelow et al., 2019; Harms et al., 2017). To differentiate between those who have a deeper understanding of the ITERS and those who have only participated in a general Environment Rating Scales training, the threshold for knowledge of the ITERS was set slightly higher than the TPITOS. Teachers who reported being reliable ITERS or ECERS assessors and those who previously had the ITERS conducted in their classroom were coded as high exposure. The remainder were considered low exposure. Five teachers were coded as having high exposure to the ITERS, five were coded as low.

A center was determined to have high exposure to the TPITOS if the director reported they informally use the Pyramid Model in their program or if they were formally participating in the Pyramid Model pathway of *Granite Steps for Quality*. The remaining centers were considered low exposure. Interestingly, all 10 participating centers reported completing at least “some” Pyramid Model training. Five centers were coded as high exposure to the TPITOS, five were coded as low.

While the TPITOS is tied to the Pyramid Model approach, the ITERS is not tied to any pedagogy. Thus, the ITERS or ERS are not implemented in a program the same way the Pyramid Model is. Because of this, a center was determined to have high exposure to the ITERS if the director reported the center had completed training on the ERS tools. The remaining centers,

whose directors all reported having heard of the ERS, but who had no training, were coded as low exposure. Four centers were coded as high exposure to the ITERS, six were coded as low exposure.

Classroom Observations

Trained observers conducted simultaneous 3-hour classroom observations that were followed by a 15-minute interview with the participating teacher for the TPITOS. The teacher was present for the entire observation. The program director was not required to be present. All observations occurred between the hours of 8:30 a.m. and 12 p.m. on various days of the week between mid-December 2022 and the first week of February 2023. For each observation, a reliable TPITOS observer conducted the TPITOS, and a reliable ITERS assessor conducted the ITERS-3. One TPITOS observer, the researcher for this study, performed all 10 TPITOS observations and two ITERS assessors were utilized.

One ITERS assessor conducted six ITERS-3 observations, the second conducted four. Both ITERS assessors were recruited through the researcher's contacts at the NH Bureau of Child Development and Head Start Collaboration. Both ITERS assessors were contractors of the Bureau of Child Development and Head Start Collaboration during the study. At the time of this study, one assessor was a full-time community college professor in an ECCE teacher training program, the other was a postpartum doula and ECCE professional development consultant. Given the nature of the early childhood community in New Hampshire, the TPITOS observer and both ITERS-3 assessors each had prior relationships with at least one of the participating centers. These relationships ranged from the center being a prior place of employment for the observer or assessor, to the observer or assessor having taught or trained one or more teachers at the center. It was determined that these prior relationships were acceptable knowing this same

circumstance is likely to occur in *Granite Steps for Quality* observations. As in *Granite Steps for Quality* and other QRISs, all trained observers and assessors are entrusted to maintain objectivity.

The TPITOS is intended to measure a teacher's effectiveness in promoting the social and emotional development of infants and toddlers (Bigelow et al., 2019). The ITERS-3 is designed to measure global quality in infant and toddler classrooms (Harms et al., 2017). Both the TPITOS and ITERS-3 observation tools are divided into sections. In the TPITOS, sections are called *Items*; in the ITERS-3, sections are called *Subscales* (Bigelow et al., 2019; Harms et al., 2017).

The 13 TPITOS Items include topics such as Teacher Demonstrates Warmth and Responsivity to Individual Children and Teacher Responds to Children in Distress and Manages Challenging Behaviors. Two of the 13 TPITOS Items are related to the environment: Teacher Conveys Predictability Through Carefully Planned Schedule, Routines, and Transitions, and Environment is Arranged to Foster Social-Emotional Development. The TPITOS has a checklist for each Item, and the observer notes a Yes or No for the teacher's performance on each element in the checklist, called an *Indicator*. The number of Yes Indicators out of the total number of scored Indicators is used to calculate the teacher's overall score (some Indicators can be marked "not applicable"). TPITOS scores are shown as percentages. A 70% or greater is said to indicate that the teacher is implementing the Pyramid Model approach to fidelity. *Granite Steps for Quality* requires a score of at least 70% for Quality Step 4.

The ITERS-3 has six Subscales with two related to the promotion of social and emotional development: Language and Books, and Interactions. Other Subscales include topics such as Space and Furnishings and Program Structure. The ITERS-3 has elements called *Items* within

each Subscale, and Items are further divided into *Indicators* (Harms et al., 2017). From the number of Yes and No Indicators noted for each Item, the ITERS assessor determines where the classroom falls along a scale of 1 to 7 points for the Item. Subscales are made up of three to 10 Items. Average Subscale scores are the sum of Item scores divided by the number of Items scored within that Subscale (some Items can be marked “not applicable”). Total average ITERS-3 scores are computed by adding all Subscale scores and dividing by the total number of Items scored on all Subscales. Both average Subscale scores and total average scores range from 0 to 7. *Granite Steps for Quality* has cutoff scores for both the Activities and Interactions Subscales (2.5 for Quality Step 3, 3.5 for Quality Step 3) and the Total Average Score (3.5 for Quality Step 3 and 4.5 for Quality Step 4).

Data Analysis

Equivalency Between ITERS-3 and TPITOS Scores

The first research question asked: Is there an equivalency between the ITERS-3 and TPITOS scores for each program? The researcher analyzed equivalency using a regression model and comparisons of rank among the 10 participants according to TPITOS scores and, separately, their rank according to ITERS-3 scores.

Differences in Scores Related to Exposure

The second research question asked whether the difference between a center’s performance on the TPITOS and its performance on the ITERS-3 could be related to its differing levels of knowledge or experience with the observation tools. To explore this question, the researcher first standardized scores on the two tools ($M = 0$, $SD = 1$). This produced a score for each participant based on their performance relative to their peers in this sample (known as a *z*-score; shown in Tables 4 and 5). A participant who had a TPITOS or ITERS-3 score near the

mean for this sample would have a z -score close to zero. The sign of the z -score indicates whether they were above the mean (positive z -score) or below the mean (negative z -score). The value of the z -score indicates how many standard deviations the score is above or below the mean.

The difference between each participant's standardized scores for the TPITOS and ITERS-3 was calculated to determine the variance in scores from the two tools. Differences in standardized scores were then tested for group differences based on the centers' and teachers' exposure to the TPITOS or ITERS-3 tools.

The researcher performed four independent samples t -tests for this analysis. Two tests compared the means of differences in standardized TPITOS and ITERS-3 scores for teachers with low versus high exposure and, separately, for centers with low versus high exposure to the TPITOS. Two more t -tests compared the means of differences in scores for teachers and centers with low versus high exposure to the ITERS. The standardized effect size (Hedges' g) was used to understand the magnitude of the effect of exposure on the difference between scores. Hedges' g was selected due to the small size of this sample (Grissom, 2005; Lakens, 2013).

Ecological Validity

The *Granite Steps for Quality* "Quality Step," or rating, a center would achieve based on the observations in this study was used to answer the third research question: Do the TPITOS and ITERS-3 produce ecologically valid scores? Meaning, do both scores qualify centers for the same amount of incentive funding from *Granite Steps for Quality*?

First, the researcher coded TPITOS and ITERS-3 scores into Quality Steps using a modified version of *Granite Steps for Quality*. For these analyses, it was assumed that all other training, administrative, and classroom participation requirements listed in Standard 2 of *Granite*

Steps for Quality were met. Second, the financial incentives in *Granite Steps for Quality* depend on two factors, Child Care and Development Fund (CCDF) scholarship billing and licensed capacity. The variables associated with CCDF scholarship billing are complex, therefore a case example is used to illustrate the impact of Quality Step ratings.

Determining Quality Steps

This study included just one teacher or classroom at each participating center, and thus the full requirements of either Pathway of *Granite Steps for Quality*: Standard 2 could not be verified. To make these isolated scores meaningful, it was necessary to assume that all other teachers or classrooms in the center, who would be included in *Granite Steps for Quality*, would meet or exceed the performance of study participants. For example, if participant H has an ITERS-3 score that qualifies them for Quality Step 3, it is assumed that all other classrooms at center H meet the minimum requirements for Step 3.

In the Pyramid Model Pathway, centers that have TPITOS observations completed by outside observers, regardless of the score they receive are rated as at least Quality Step 3. A center can achieve Step 4 by demonstrating that the Pyramid Model approach is being implemented to fidelity, which is defined as a score of at least 70% (Bigelow et al., 2019). Thus, a TPITOS score of 69.99% or less was considered a Step 3, and 70% or higher was considered a Step 4. Quality Step determinations are shown in Table 6.

Table 6

Comparisons of Granite Steps for Quality Step Ratings According to TPITOS and ITERS-3

Scores as Defined in the Pyramid Model and Environment Rating Scales Pathways

ID	Quality Step		Step difference
	TPITOS	ITERS-3	
A	3	2	-1
B	4	3	-1
C	4	4	0
D	4	2	-2
E	4	2	-2
F	4	4	0
G	4	4	0
H	3	3	0
I	4	4	0
J	4	3	-1

The ITERS-3 Pathway has designated cutoff scores for Quality Step 3 and Step 4. To be rated as Step 3, a participant must have a total score of 3.5 or higher and their score on the subscales Activities and Interactions must be at least 2.5. For Step 4, total scores must be 4.5 or higher with not less than 3.5 on the Activities and Interactions subscales. Therefore, participants with a total score below 3.49, or Activities *or* Interactions subscale scores below 2.49 were considered a Step 2; those with total scores of 3.5 to 4.49 *and* with Activities and Interactions subscales scores of at least 2.5 were considered a Step 3; and those with total scores above 4.5 *and* Activities and Interactions subscales scores of 3.5 or higher were considered Step 4.

Results

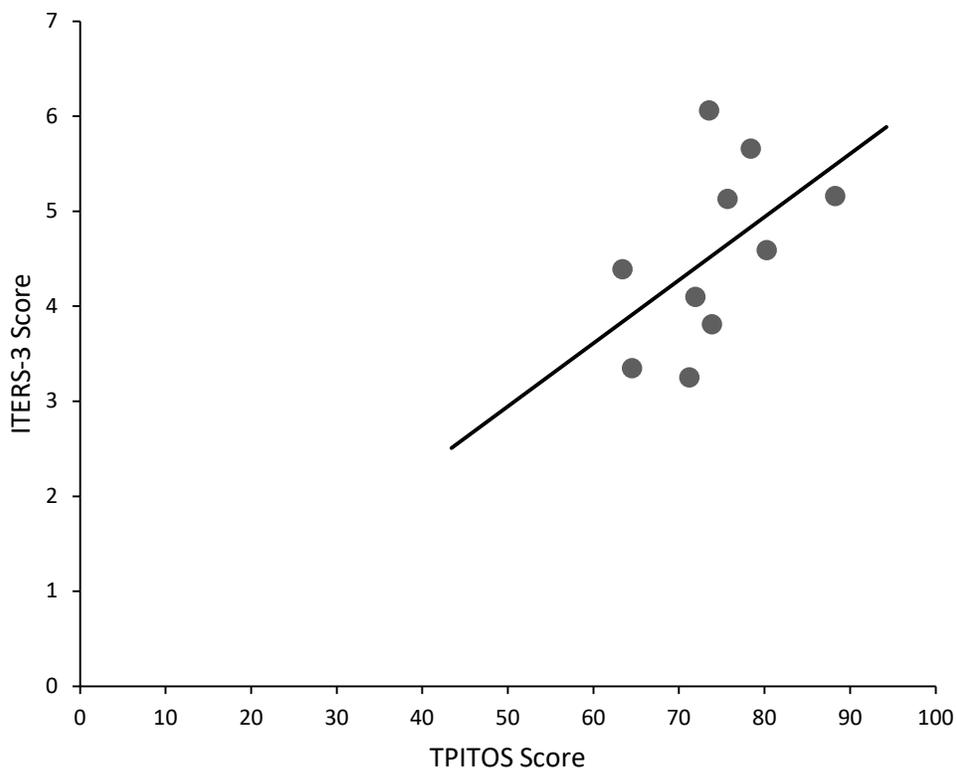
The TPITOS is scored as a percentage, while the ITERS-3 is scored as a numeric value between one and seven. TPITOS scores in this sample ranged from 63.42% to 88.24% ($M = 74.10\%$, $SD = 7.29\%$). ITERS-3 scores ranged from 3.25 to 6.06 ($M = 4.55$, $SD = 0.95$).

Equivalency Between TPITOS and ITERS-3 Scores

As seen in Figure 2, there is a strong positive linear association between TPITOS scores and ITERS-3 scores ($r = .509$, $p = .133$). $R^2 = .259$, thus 25.9% of the variation in ITERS-3 scores is predicted by TPITOS scores, however, this regression model is not significant. While the standardized effect size is robust, the results suggest a lack of clear equivalence between the two tools.

Figure 2

Association Between TPITOS Scores and ITERS-3 Scores



Comparison of Rankings

While a scatterplot demonstrates how TPITOS and ITERS-3 scores relate to one another, a rank comparison allows for a participant's score relative to their peers on one tool to be compared to their score relative to peers on another tool. Rankings are compared in Table 7. As the position change shows, seven out of 10 centers changed rank between the TPITOS and the ITERS-3. If these two observation tools were equivalent measures of quality, the lowest-ranked program on the TPITOS would also be the lowest-ranked program on the ITERS-3. Therefore, in this sample, the TPITOS and ITERS-3 do not provide equivalent measures of quality in terms of program rankings.

Table 7

Comparison of Rankings for TPITOS Score and ITERS-3 Score

ID	Rank		Position change
	TPITOS	ITERS-3	
H	1	5	+4
A	2	2	0
D	3	1	-2
J	4	4	0
G	5	10	+5
E	6	3	-3
C	7	7	0
F	8	9	+1
I	9	6	-3
B	10	8	-2

Note. Participants are listed according to TPITOS score rank. 1 is the lowest-ranked participant, 10 is the highest-ranked participant (with the best score).

Difference in Scores Related to Exposure

Teachers with low exposure to the TPITOS ($n = 6$) tended to have TPITOS scores that were slightly lower than standardized ITERS-3 scores ($M = -0.13$, $SD = 1.15$). Teachers with high exposure to the TPITOS ($n = 4$) tended to have standardized TPITOS scores that were higher than standardized ITERS-3 scores ($M = 0.20$, $SD = 0.80$). However, these findings were not significant ($p = .602$) and have a small effect size (Hedges' $g = -0.29$). Therefore, a teacher's prior knowledge of or experience with the Pyramid Model approach or the TPITOS did not predict differences between their TPITOS and ITERS-3 scores.

Teachers with low exposure to the ITERS ($n = 5$) tended to have ITERS-3 scores that were very slightly higher than standardized TPITOS scores ($M = -0.05$, $SD = 1.05$). While teachers with high exposure to the ITERS ($n = 5$) tended to have standardized ITERS-3 scores that were very slightly lower than TPITOS scores ($M = 0.05$, $SD = 1.05$). These findings were not significant ($p = .896$), and do not have any effect (Hedges' $g = 0.08$). Therefore, a teacher's prior knowledge of or experience with the Environment Rating Scales or ITERS did not predict differences between their TPITOS and ITERS-3 scores.

Centers with low exposure to the TPITOS ($n = 5$) tended to have standardized TPITOS scores that were lower than standardized ITERS-3 scores ($M = -0.50$, $SD = 1.00$). Centers with high exposure to the TPITOS ($n = 5$) tended to have standardized TPITOS scores that were higher than standardized ITERS-3 scores ($M = 0.49$, $SD = 0.78$). These findings were not significant ($p = .125$), but the effect size was large (Hedges' $g = -0.98$). Therefore, a center's experience with the Pyramid Model approach or the TPITOS cannot be said to predict a difference between their TPITOS and ITERS-3 scores. However, a center's prior experience with

the Pyramid Model approach or the TPITOS could explain instances where standardized TPITOS scores were higher than standardized ITERS-3 scores.

Centers with low exposure to the ITERS ($n = 6$) tended to have standardized ITERS-3 scores that were slightly higher than standardized TPITOS scores ($M = -0.13$, $SD = 0.81$). Meanwhile, centers with high exposure to the ITERS ($n = 4$) tended to have standardized ITERS-3 scores that were slightly lower than standardized TPITOS scores ($M = 0.19$, $SD = 1.33$). These findings were not significant ($p = .684$), and the effect size was small (Hedges' $g = -0.28$). Therefore, centers' experience with the Environment Rating Scales or ITERS cannot be said to predict differences between their TPITOS and ITERS-3 scores.

Ecological Validity

The rank comparison in Table 7 demonstrates a lack of equivalency between TPITOS and ITERS-3. Building on this, the real-world consequences of differences in each participant's numeric scores must be explored. The right-most column in Table 6 shows what would happen if, hypothetically, a center that received a Quality Step rating based on the TPITOS switched to a rating based on its ITERS-3 score. In this scenario, five out of 10 study participants drop down one or two Quality Steps. The lower rating would not only impact the center's reputation, it would also impact the financial incentives available to the center from the state.

Illustrative Example

Granite Steps for Quality calculates reimbursement using two factors: Child Care and Development Fund (CCDF) scholarship billing and licensed capacity. In this example, some speculation is required to approximate a real-world scenario. Enrollment is used as a proxy for licensed capacity. The mean enrollment for centers in this sample was 59.1 children with an average of 13.65% of children receiving CCDF scholarships. The hypothetical center has 59

children enrolled, eight of whom (approximately 13.65%) receive CCDF scholarships. The scholarship recipients in this example are two full-time infants, two half-time infants, two full-time toddlers and two half-time toddlers.

If the standard weekly scholarship rate minus a child's cost share of \$20 is billed, the total monthly billing for the center will be: $4.3 \text{ weeks per month} \times (2(\$230.00) + 2(\$173.55) + 2(\$208.70) + 2(\$157.06))$, or $4.3 \times \$1,538.62$ which is \$6,616.07 (NH Bureau of Child Development and Head Start Collaboration, 2019). If this center follows the TPITOS Pathway and is rated at Step 4, they would receive a quality incentive of \$3,600 per quarter and a scholarship-based incentive equal to 10% of their monthly scholarship billing, or \$661.61, each month. This equates to \$14,400 in quality incentives per year and \$7,939.32 in scholarship-based incentives for the year. The overall total of incentives would be \$22,339.32. (Refer to Appendix E for tables showing the incentive structure for non-accredited center-based programs).

If this same center were to follow the ITERS-3 Pathway and they drop down one step to Step 3, its quality incentive drops to \$3,000 per quarter, but its scholarship-based incentive stays the same. Their yearly incentive would then be \$19,939.32. If the center dropped two steps down to Step 2, both the quality incentive and the scholarship-based incentive decrease. Here the quality bonus is \$2,400 per quarter and the scholarship-based incentive is just 5% of monthly scholarship billing (\$330.80 per month). This equates to \$9,600 in quality incentives and \$3,969.60 in scholarship-based incentives, or \$13,569.60 in incentives for the year. For this hypothetical child care center, the difference between a rating of Step 2 and a rating of Step 4 is \$8,769.72 per year.

Discussion

This paper describes the role of Quality Rating/Recognition and Improvement Systems (QRISs) in measuring and improving the quality of child care. For infants and toddlers in particular, translating caregiver-child relationships and interactions into a numeric rating scale is not an easy task. It's unlikely imperfect observation tools will produce perfect rating systems, thus, as states develop and revise QRISs, there is a need for continuous evaluation. This study and many before it point to the need for federal and state funding for QRIS validation research (Boller & Maxwell, 2015). These systems are as varied and complex as the children receiving care in ECCE and school-age settings. No study includes all 45 QRISs, but each study contributes recommendations for the state or states evaluated, as well as considerations for broader QRIS design and revision.

Granite Steps for Quality

The results of this research suggest a lack of equivalency between the two observation tools used in the two pathways for infant and toddler classrooms in center-based ECCE programs participating in *Granite Steps for Quality*. While this is not a complete evaluation or validation study, these results point to the need for further research on the design of New Hampshire's QRIS. The TPITOS and ITERS-3, and the cutoff scores used in the Environment Rating Scales Pathway do not have enough reliability or validity to be applied in a rating system without thorough testing (Lahti et al., 2015).

One concern with a two-pathway QRIS design where the pathways are not proven to be equivalent is that an objectively lower-quality center could receive the same quality rating as a higher-quality center. The results of this study demonstrate that the TPITOS, as it is used in *Granite Steps for Quality*, consistently leads to higher Quality Step ratings. For families

examining the quality levels of child care centers, this is problematic. As the illustrative example shows, for child care centers seeking financial incentives for their efforts, this is inequitable. A rating system must be reliable for it to have practical utility.

The issue of inter-rater reliability, which was not within the scope of this study, must be addressed in a QRIS which relies heavily on observation scores – as the Environment Rating Scales Pathway does – and which defines cutoff scores, as both pathways do (Kirby et al., 2015; Lahti et al., 2015). Variation between ITERS assessors has significant implications in a system where the difference between a 3.4 and a 3.5 on the ITERS-3 could be several thousands of dollars in annual funding. Inter-rater reliability is all the more important in smaller states like New Hampshire where we see prior relationships between state-contracted observers and the providers participating in the QRIS.

It is recommended that developers of *Granite Steps for Quality* plan periodic evaluations of the overall system, and frequent inter-rater reliability checks for those conducting observations. Based on the results of this study, the cutoff scores for both pathways in *Granite Steps for Quality* should be investigated further. One method that could be used for this investigation is an analysis of initial score distribution (Lahti et al., 2015). This allows cutoff scores to be recalibrated based on early results of the QRIS.

Infants and Toddlers in QRISs

Infancy and toddlerhood are distinctly different from the preschool and prekindergarten periods. It follows that the needs of children from birth to three years old are inherently different, and the standards for infant and toddler teachers should be different. *Granite Steps for Quality* takes a step toward acknowledging this by including observation tools designed specifically for teachers and classrooms with infants and toddlers. However, using this QRIS as an example, a

center's ratings are based on observations conducted in only a sample of the classrooms. There is no requirement that an observation must occur in an infant or toddler classroom. This leads to the possibility that infant and toddler teachers and classrooms will be overlooked.

It is recommended that future research evaluates the validity of QRISs for infant/toddler groups separately from preschool/prekindergarten groups. Even when infant/toddler teachers are rated alongside preschool/prekindergarten teachers, as in *Granite Steps for Quality*, the validity of the system for infants, toddlers, and their care providers must be tested separately. A thorough evaluation study would ask two questions: is this QRIS effectively measuring and improving quality for infant and toddler groups? And is this QRIS effectively measuring and improving quality for preschool and prekindergarten groups? This would contribute to a deeper understanding of the needs of children at all developmental stages and their caregivers.

QRIS Objectives

When designing and evaluating QRISs, it can be beneficial to categorize the elements of a QRIS by the objectives they intend to achieve (Boller & Maxwell, 2015). These could be things such as improved outcomes for children, determining quality ratings for public information, increased professionalization of the workforce, etc. When considering a QRIS with two pathways which are not proven to be equivalent, some theoretical questions arise: Is using more than one strategy to improve outcomes for children logical or equitable? Is it fair to rate centers using two different observation tools without informing the public of the differences? Are all professionals in both pathways of the QRIS benefitting from the professional development plan, and are they improving their practice? Lastly, one might question, who benefits from having two pathways? With *Granite Steps for Quality* entering its second year, it is recommended that its developers and stakeholders consider the efficacy of the two-pathway

design. A QRIS with two pathways is inherently more complex to implement, and more challenging to evaluate and revise, than a QRIS with one pathway.

It should be acknowledged that Granite Steps for Quality is relatively new, and it has been revised every six months since it was released. Given the significant impacts of Quality Step ratings on reputation and funding, ongoing research is needed to determine the effect of each revision. Like ECCE programs, evaluation and revision are critical to a system's continuous quality improvement. All changes must be made thoughtfully.

Limitations

This study was designed to be feasible to complete within 12 months. The small sample size contributed to the lack of significance of some analyses, but, where possible, conservative analytic methods were selected. Due to the nature of this study, the sample may be biased toward higher quality than the general population of *Granite Steps for Quality* participants. Here, directors were asked to select a participating teacher for the observations and the teacher had the option to decline participation. In addition, only one teacher and classroom at each center were evaluated. In *Granite Steps for Quality*, the teachers or classrooms observed are selected at random, and multiple teachers and classrooms at each center are assessed. Thus, it is likely *Granite Steps for Quality* will capture greater variation in TPITOS and ITERS-3 scores. Further, this recruitment procedure limits the generalizability of the study's findings. A robust evaluation study would conduct multiple random observations in each center exactly as *Granite Steps for Quality* dictates.

The small early childhood community in New Hampshire proved to be a limitation in this study. While every effort was made to ensure observer and assessor objectivity, bias is not uncommon in social science research, and it is possible it exists in this study. Between the

TPITOS observer and two ITERS-3 assessors, researchers had prior relationships with directors or teachers from at least five out of the 10 participating centers. Future studies could avoid this potential bias by identifying the researchers' existing relationships and avoiding any opportunities for personal interests to interfere with objectivity.

Another limitation was the restriction to only nonprofit participants – a requirement of the project's funding. As large, for-profit ECCE centers become more prevalent, including them in QRISs and evaluation studies is critical (Gale, 2022). Also, in New Hampshire it was challenging to find Montessori schools that are nonprofit and those that were identified did not elect to participate in this study. Due to the notable number of Montessori schools, the assumption of quality in Montessori programs, and the minimal research to back this assumption, this is a sector of center-based ECCE that should be included in QRISs and QRIS evaluation studies (Marshall, 2017).

Evaluating *Granite Steps for Quality* Standard 1: Staff Qualifications was not within the scope of this study, nor was an evaluation of the administrative and training requirements in Standard 2: Learning Environments. Capturing achievement on all components of *Granite Steps for Quality* and determining whether certain components impact a program's quality ratings more than others would be an interesting direction for future research.

Implications for Future Research

Despite the current study's small size and design limitations, the findings are worth considering as QRIS research and the refinement of *Granite Steps for Quality* continues. The analysis herein suggests the two pathways in *Granite Steps for Quality* are not equivalent, and it is possible a two-pathway design is inherently problematic. Expanded research on choice in QRISs would be helpful. In addition, further research is needed on the reliability and validity of

the TPITOS and the ITERS-3. Large-scale equivalency studies including these two tools could help justify their parallel usage in *Granite Steps for Quality*.

Knowing that part of New Hampshire's aim with the two-pathway design was to have a more accessible QRIS, more research on QRIS accessibility and participation is recommended. As long as most QRISs remain voluntary, it is in the public's best interest for participation rates to be as high as possible. Strategies that improve accessibility, increase participation and lead to real improvement of participating programs should be identified and promoted.

Conclusion

Given the importance of quality in infant and toddler care and education, the systems being implemented to monitor and improve quality require more attention and evaluation than they have received. Further research on New Hampshire's QRIS is needed, as are validity and reliability studies of the measurement tools used in QRISs across the country. Parents should have a dependable method for determining the quality of their children's care settings, and child care providers deserve equitable access to incentives from the state. As this study demonstrates, the current design of New Hampshire's QRIS may lead to inconsistent ratings and incentives for center-based care providers.

References

- Administration for Children & Families. (n.d.). *About QRIS: QRIS resource guide*. Retrieved March 10, 2022, from <https://ecquality.acf.hhs.gov/about-qrisc>
- Ainsworth, M. D. S., & Bowlby, J. (1991). An ethological approach to personality development. *The American Psychologist*, *46*(4), 333–341. <https://doi.org/10.1037/0003-066X.46.4.333>
- Araujo, M. C., Dormal, M., & Schady, N. (2019). Childcare quality and child development. *Journal of Human Resources*, *54*(3), 656–682. <https://doi.org/10.3368/jhr.54.3.0217.8572R1>
- Banghart, P., Halle, T., Bamdad, T., Cook, M., Redd, Z., Cox, A., & Carlson, J. (2020). A review of the literature on access to high-quality care for infants and toddlers. In *Child Trends*. Child Trends.
- Berry, D., Blair, C., Ursache, A., Willoughby, M. T., & Granger, D. A. (2014). Early childcare, executive functioning, and the moderating role of early stress physiology. *Developmental Psychology*, *50*(4), 1250–1261. <https://doi.org/10.1037/a0034700>
- Bigelow, K. M., Carta, J. J., Irvin, D. D., & Hemmeter, M. L. (2019). *Teaching Pyramid Infant-Toddler Observation Scale (TPITOS) for infant-toddler classrooms, research edition*. Paul H. Brookes Publishing, Co.
- Boller, K., & Maxwell, K. (2015). QRIS research: Looking back and looking forward. *Early Childhood Research Quarterly*, *30*, 339–342. <https://doi.org/10.1016/j.ecresq.2014.10.002>
- Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology*, *22*(6), 723–742. <https://doi.org/10.1037/0012-1649.22.6.723>

Build Initiative. (2023). *Quality Improvement Systems Compendium—QIS*.

<https://qualitycompendium.org/>

Burchinal, M. (2018). Measuring early care and education quality. *Child Development Perspectives*, *12*(1), 3–9. <https://doi.org/10.1111/cdep.12260>

Burchinal, M., Roberts, J. E., Nabors, L. A., & Bryant, D. M. (1996). Quality of center child care and infant cognitive and language development. *Child Development*, *67*(2), 606–620.

<https://doi.org/10.1111/j.1467-8624.1996.tb01754.x>

Cannon, J., Zellman, G., Karoly, L., & Schwartz, H. (2017). *Quality rating and improvement systems for early care and education programs: Making the second generation better*.

RAND Corporation. <https://doi.org/10.7249/PE235>

Center on the Developing Child at Harvard University. (2022). *Early childhood mental health*.

Center on the Developing Child at Harvard University.

<https://developingchild.harvard.edu/science/deep-dives/mental-health/>

Chaudry, A., & Sandstrom, H. (2020). Child care and early education for infants and toddlers.

Future of Children, *30*(2), 165–190.

<https://eric.ed.gov/?q=infant+care+quality&pr=on&pg=2&id=EJ1293558>

Child Care Aware of NH. (2022, October). *Granite Steps for Quality*. NH Connections.

<https://www.nh-connections.org/providers/quality-care-matters/>

Child Care Technical Assistance Network. (n.d.). *History and purposes of CCDBG and CCDF*.

Retrieved May 2, 2022, from <https://childcareta.acf.hhs.gov/ccdf-fundamentals/history-and-purposes-ccdbg-and-ccdf>

- Cohen, L. E., & Waite-Stupiansky, S. (2017). *Theories of early childhood education: Developmental, behaviorist, and critical* (1st ed.). Routledge.
<https://doi.org/10.4324/9781315641560>
- Durkin, K., Lipsey, M. W., Farran, D. C., & Wiesen, S. E. (2022). Effects of a statewide pre-kindergarten program on children's achievement and behavior through sixth grade. *Developmental Psychology*, *58*(3), 470–484. <https://doi.org/10.1037/dev0001301>
- Early, D. M., Sideris, J., Neitzel, J., LaForett, D. R., & Nehler, C. G. (2018). Factor structure and validity of the Early Childhood Environment Rating Scale – Third Edition (ECERS-3). *Early Childhood Research Quarterly*, *44*, 242–256.
<https://doi.org/10.1016/j.ecresq.2018.04.009>
- Elicker, J., Gold, Z. S., Mishra, A. A., & Lane, S. F. (2021). Toddlers' developmental trajectories as a function of QRIS rated child care quality. *Child & Youth Care Forum*.
<https://doi.org/10.1007/s10566-021-09643-z>
- Erikson, E., & Erikson, J. (1981). On generativity and identity: From a conversation with Erik and Joan Erikson. *Harvard Educational Review*, *51*(2), 249–269.
<https://doi.org/10.17763/haer.51.2.g211757u27732p67>
- Fogel, A. (1997). *Infancy: Infant, family, and society* (Third). West Publishing Company.
- Follari, L. (2019). *Foundations and best practices in early childhood education* (4th ed.). Pearson.
- Frank Porter Graham Child Development Institute. (n.d.). *Reliability and validity of the ITERS-3™: Environment Rating Scales®*. Environment Rating Scale Family of Products. Retrieved May 3, 2022, from <https://ers.fpg.unc.edu/reliability-and-validity-iters-3%E2%84%A2>

- Fukkink, R. G., & Tavecchio, L. W. C. (2010). Effects of Video Interaction Guidance on early childhood teachers. *Teaching and Teacher Education*, 26(8), 1652–1659.
<https://doi.org/10.1016/j.tate.2010.06.016>
- Gale, R. (2022, November 7). The boom of for-profit child care—And where that takes us: A Q and A with Elliot Haspel. *New America*. <http://newamerica.org/better-life-lab/blog/the-boom-of-for-profit-child-care-and-where-that-takes-us/>
- Gerber, M., & Johnson, A. (1998). *Your self-confident baby: How to encourage your child's natural abilities from the very start*. John Wiley and Sons, Inc.
- Grissom, R. J. (2005). *Effect sizes for research: A broad practical approach*. Lawrence Erlbaum Associates. <http://archive.org/details/effectsizesforre0000gris>
- Halle, T., Anderson, R., Blasberg, A., Chrisler, A., & Simkin, S. (2011). *Quality of Caregiver-Child Interactions for Infants and Toddlers (Q CCIIT): A review of the literature* (OPRE 2011-25). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
https://www.acf.hhs.gov/sites/default/files/documents/opre/quality_caregiver.pdf
- Harms, T., Cryer, D., Clifford, R. M., & Yazejian, N. (2017). *Infant/Toddler Environment Rating Scale* (Third). Teachers College Press.
- Haspel, E. (2019). *Crawling behind: America's childcare crisis and how to fix it*. Black Rose Writing.
- Helmerhorst, K. O. W., Riksen-Walraven, J. M. A., Fukkink, R. G., Tavecchio, L. W. C., & Gevers Deynoot-Schaub, M. J. J. M. (2017). Effects of the Caregiver Interaction Profile Training on caregiver-child interactions in Dutch child care centers: A randomized

- controlled trial. *Child & Youth Care Forum*, 46(3), 413–436.
<https://doi.org/10.1007/s10566-016-9383-9>
- Hemmeter, M. L., Snyder, P. A., Fox, L., & Algina, J. (2016). Evaluating the implementation of the pyramid model for promoting social-emotional competence in early childhood classrooms. *Topics in Early Childhood Special Education*, 36(3), 133–146.
<https://doi.org/10.1177/0271121416653386>
- Hestenes, L. L., Kintner-Duffy, V., Wang, Y. C., La Paro, K., Mims, S. U., Crosby, D., Scott-Little, C., & Cassidy, D. J. (2015). Comparisons among quality measures in child care settings: Understanding the use of multiple measures in North Carolina’s QRIS and their links to social-emotional development in preschool children. *Early Childhood Research Quarterly*, 30, 199–214. <https://doi.org/10.1016/j.ecresq.2014.06.003>
- Keating, K., Cole, P., & Schneider, A. (2021). *State of babies yearbook: 2021*. ZERO TO THREE. <https://stateofbabies.org/wp-content/uploads/2021/04/State-of-Babies-2021-Full-Yearbook.pdf>
- Kirby, G., Caronongan, P., Malone, L. M., & Boller, K. (2015). What do quality rating levels mean? Examining the implementation of QRIS ratings to inform validation. *Early Childhood Research Quarterly*, 30, 291–305.
<https://doi.org/10.1016/j.ecresq.2014.08.006>
- La Paro, K. M., Williamson, A. C., & Hatfield, B. (2014). Assessing quality in toddler classrooms using the CLASS-Toddler and the ITERS-R. *Early Education and Development*, 25(6), 875–893. <https://doi.org/10.1080/10409289.2014.883586>
- Lahti, M., Elicker, J., Zellman, G., & Fiene, R. (2015). Approaches to validating child care quality rating and improvement systems (QRIS): Results from two states with similar

- QRIS type designs. *Early Childhood Research Quarterly*, 30, 280–290.
<https://doi.org/10.1016/j.ecresq.2014.04.005>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4.
<https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00863>
- Legerstee, M., Haley, D. W., & Bornstein, M. H. (2013). *The infant mind: Origins of the social brain*. Guilford Press.
- Marshall, C. (2017). Montessori education: A review of the evidence base. *Npj Science of Learning*, 2(1), Article 1. <https://doi.org/10.1038/s41539-017-0012-7>
- Mashburn, A. J. (2017). Evaluating the validity of classroom observations in the Head Start designation renewal system. *Educational Psychologist*, 52(1), 38–49.
<https://doi.org/10.1080/00461520.2016.1207539>
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3), 732–749. <https://doi.org/10.1111/j.1467-8624.2008.01154.x>
- McCartney, K. (1984). Effect of quality of day care environment on children's language development. *Developmental Psychology*, 20(2), 244–260. <https://doi.org/10.1037/0012-1649.20.2.244>
- Miller, P. H. (2016). *Theories of developmental psychology* (6th ed.). Macmillan Higher Education.
- Mincic, M., Smith, B. J., & Strain, P. (2009). Administrator strategies that support high fidelity implementation of the Pyramid Model for promoting social-emotional competence and

- addressing challenging behavior. In *Technical Assistance Center on Social Emotional Intervention for Young Children*. Technical Assistance Center on Social Emotional Intervention for Young Children. <https://eric.ed.gov/?id=ED526388>
- Mitchell, A. (2005). *Stair steps to quality: A guide for states and communities developing quality rating systems for early care and education*. United Way.
http://www.ctearlychildhood.org/uploads/6/3/3/7/6337139/qris_anne_mitchell.pdf
- Mooney, C. G. (2000). *Theories of childhood: An introduction to Dewey, Montessori, Erickson, Piaget & Vygotsky*. Redleaf Press. <https://www.redleafpress.org/Theories-of-Childhood-Second-Edition-An-Introduction-to-Dewey-Montessori-Erikson-Piaget-Vygotsky-P781.aspx>
- Morris, A. S., & Williamson, A. C. (Eds.). (2018). *Building early social and emotional relationships with infants and toddlers: Integrating research and practice*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-03110-7>
- Mortensen, J. A., & Barnett, M. A. (2015). Teacher–child interactions in infant/toddler child care and socioemotional development. *Early Education & Development, 26*(2), 209–229.
<https://doi.org/10.1080/10409289.2015.985878>
- National Center on Early Childhood Quality Assurance. (2017). *QRIS Compendium 2016 fact sheets: Use of observational tools in QRIS*.
- National Institute of Child Health and Human Development Early Child Care Research. (2000). The relation of child care to cognitive and language development. *Child Development, 71*(4), 960–980. <https://doi.org/10.1111/1467-8624.00202>

- National Scientific Council on the Developing Child. (2004). *Young children develop in an environment of relationships* (Working paper no. 1; p. 12).
<http://www.developingchild.net>
- NH Bureau of Child Development and Head Start Collaboration. (2019, November). *Child care scholarship program: Employment related maximum weekly standard rates*. NH Connections. <https://www.nh-connections.org/wp-content/uploads/2021/03/Absentee-Billing-Weekly-Rates.pdf>
- Parlakian, R. (2003). *Before the ABCs: Promoting school readiness in infants and toddlers*. ZERO TO THREE: National Center for Infants, Toddlers and Families.
- Peisner-Feinberg, E. S., Burchinal, M., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade. *Child Development, 72*(5), 1534–1553. <https://doi.org/10.1111/1467-8624.00364>
- Phillips, D., McCartney, K., & Scarr, S. (1987). Child-care quality and children's social development. *Developmental Psychology, 23*(4), 537–543. <https://doi.org/10.1037/0012-1649.23.4.537>
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. Basic Books.
- Pinto, A. I., Cadima, J., Coelho, V., Bryant, D. M., Peixoto, C., Pessanha, M., Burchinal, M., & Barros, S. (2019). Quality of infant child care and early infant development in Portuguese childcare centers. *Early Childhood Research Quarterly, 48*, 246–255.
<https://doi.org/10.1016/j.ecresq.2019.04.003>
- Ragni, B., Boldrini, F., Buonomo, I., Benevene, P., Grimaldi Capitello, T., Berenguer, C., & De Stasio, S. (2021). Intervention programs to promote the quality of caregiver–child

- interactions in childcare: A systematic literature review. *International Journal of Environmental Research and Public Health*, 18(21), Article 21.
<https://doi.org/10.3390/ijerph182111208>
- Reis, H. T., Collins, W. A., & Berscheid, E. (2000). The relationship context of human behavior and development. *Psychological Bulletin*, 126(6), 844–872. <https://doi.org/10.1037/0033-2909.126.6.844>
- Ruzek, E., Burchinal, M., Farkas, G., & Duncan, G. J. (2014). The quality of toddler child care and cognitive skills at 24 months: Propensity score analysis results from the ECLS-B. *Early Childhood Research Quarterly*, 29(1), 12–21.
<https://doi.org/10.1016/j.ecresq.2013.09.002>
- Ryberg, R., Wiggins, L., Moore, K. A., Daily, S., Piña, G., & Klin, A. (2022). Measuring state-level infant and toddler well-being in the United States: Gaps in data lead to gaps in understanding. *Child Indicators Research*. <https://doi.org/10.1007/s12187-021-09902-4>
- Shah, H., Niland, K., Caronongan, P., & Moiduddin, E. (2020). *Competencies of infant and toddler teachers and caregivers: A compendium of measures* (OPRE 2020-21; p. 122). Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Shonkoff, J. P., & Phillips, D. A. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. The National Academies Press.
<https://doi.org/10.17226/9824>
- Slot, P. L., Leseman, P. P. M., Verhagen, J., & Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care

settings. *Early Childhood Research Quarterly*, 33, 64–76.

<https://doi.org/10.1016/j.ecresq.2015.06.001>

Thomason, A. C., & La Paro, K. M. (2009). Measuring the quality of teacher–child interactions in toddler child care. *Early Education and Development*, 20(2), 285–304.

<https://doi.org/10.1080/10409280902773351>

Thorpe, K., Rankin, P., Beaton, T., Houen, S., Sandi, M., Siraj, I., & Staton, S. (2020). The when and what of measuring ECE quality: Analysis of variation in the Classroom Assessment Scoring System (CLASS) across the ECE day. *Early Childhood Research Quarterly*, 53, 274–286. <https://doi.org/10.1016/j.ecresq.2020.05.003>

US Census Bureau. (2020). *Urban and rural*. Census.Gov. <https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html>

Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M. (2010). *Quality dosage, thresholds, and features in early childhood settings: Literature review* (OPRE 2011-5; p. 74). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

APPENDICES

APPENDIX A

*Granite Steps for Quality Pathways, Standard 2: Learning Environments***Table A1***ERS/ITERS-3 Pathway*

STEP	ERS/ITERS-3 Pathway
1	Administrator and 25% of eligible staff have completed an approved ERS training.
2	Administrator and 25% of eligible staff have completed an approved ERS training. AND Program submits ERS self-assessments for all classrooms using age appropriate ERS tools
3	Administrator and 25% of eligible staff have completed an approved ERS training. AND Program submits ERS self-assessments for all classrooms using age appropriate ERS tools. AND ERS observations have been conducted in one third of classrooms. Total Average Scores must be 3.5 or higher with no classroom scoring less than 2.5 on the subscales: Activities and Interactions. <i>Head Start Programs may meet this criteria by documenting 3 out of the 5 following Performance Standards (1302.30, 1302.31, 1302.45, 1304.21 and 1304.53) during their most recent federal monitoring visit.</i> <i>NAEYC Accredited Programs may meet this criteria by presenting documentation of meeting NAEYC Learning Environment standards.</i>
4	Administrator and 25% of eligible staff have completed an approved ERS training. AND Program submits ERS self-assessments for all classrooms using age appropriate ERS tools. AND ERS observations have been conducted in one half of classrooms. Total Average Scores must be 4.5 or higher with no classroom scoring less than 3.5 on the subscales: Activities and Interactions.

Note. Adapted from Child Care Aware of NH. (2022). *Granite Steps for Quality*. NH

Connections. <https://www.nh-connections.org/providers/quality-care-matters/>

Table A2*Pyramid Model/TPITOS Pathway*

STEP	Pyramid Model/TPITOS Pathway
1	<p>Administrator and all eligible staff have completed the approved Pyramid Model introductory training: Building Positive Social Emotional Skills for All Children: Introducing the Pyramid Model Framework</p> <p>AND Administrator and 25% of eligible staff have completed at least two modules (minimum of 12 hours) of Pyramid Model: Infant & Toddler, Preschool or Birth through Five Trainings.</p> <p>Please note: Programs awarded a Quality Step 1 are committing to participation in and completion of the Pyramid Model readiness series within one year of application and will be expected to reapply at a Quality Step 2 upon renewal</p>
2	<p>Administrator and all eligible staff have completed the approved Pyramid Model introductory training.</p> <p>AND a consistent cohort that includes representatives of program administration and 25% of teaching staff have completed the professional learning community: “Program Wide Pyramid Model Implementation: An Introduction and Overview for Child Care Providers” and have confirmed commitment to the Pyramid Model Pathway.</p> <p>AND An administrator and 25% of eligible staff have completed Pyramid - Wellness: Taking Care of Yourself and all lessons included in one of the following trainings: Infant & Toddler, Preschool or Birth through Five Trainings (Modules 1-3).</p> <p>AND Program has been assigned a program coach, established a leadership team, created mission and vision statements, completed their Benchmarks of Quality, and created their action plan based on the Benchmarks of Quality</p>
3	<p>Program has met all requirements of Steps 1 and 2.</p> <p>AND Administrator and 50% of eligible staff have completed Trauma-Informed Care & The Pyramid Model AND Pyramid - Reducing Implicit Bias trainings.</p> <p>AND Leadership team is actively engaged (meets monthly), Practice-Based Coaching and Teaching Pyramid Observation Tool (TPOT) or Teaching Pyramid Infant-Toddler Observation Scale (TPITOS) is occurring in 1 classroom in the program.</p>
4	<p>Program has met all requirements of Steps 1-3.</p> <p>AND Administrator and 50% of all eligible staff have completed the Prevent Teach Reinforce for Young Children (PTR-YC) training.</p> <p>AND Pyramid Model practices are being implemented to fidelity as evidenced by TPOT and TPITOS data and the leadership team is collecting, analyzing and using data for program improvement.</p>

Note. Adapted from Child Care Aware of NH. (2022). *Granite Steps for Quality*. NH

Connections. <https://www.nh-connections.org/providers/quality-care-matters/>

APPENDIX B**IRB Approval Letter****University of New Hampshire**

Research Integrity Services
51 College Road, Durham, NH 03824
research.integrity@unh.edu

Aug 25, 2022 4:22:56 PM EDT

Eileen Murphy

cc: Dean of CHHS (UFSHHS), Human Development & Family Studies (UFS00)

Study Title: Evaluating the Equivalency and Ecological Validity of TPITOS and ITERS-3 Ratings of Quality in New Hampshire's Quality Recognition and Improvement System

IRB #: IRB-FY2022-365

Approval: August 25, 2022

The Institutional Review Board for the Protection of Human Subjects in Research (IRB) has reviewed and approved the protocol for your study as Exempt as described in Title 45, Code of Federal Regulations (CFR), Part 46, Subsection 104(d). Approval is granted to conduct your study as described in your protocol.

Researchers who conduct studies involving human subjects have responsibilities as outlined in the document, [Responsibilities of Directors of Research Studies Involving Human Subjects](#). Please read this document carefully before commencing your work involving human subjects.

Note: IRB approval is separate from UNH Purchasing approval of any proposed methods of paying study participants. Before making any payments to study participants, researchers should review the Payment of Incentives/ Compensation to Research Participants [guidance](#) to ensure they are complying with institutional requirements. If such institutional requirements are not consistent with the confidentiality or anonymity assurances in the IRB-approved protocol and consent documents, you may need to request a modification from the IRB.

Upon completion of your study, please submit a study closure form through [Cayuse IRB/Human Ethics](#) along with a report of your findings.

If you have questions or concerns about your study or this approval, please feel free to contact Melissa McGee at 603-862-2005 or melissa.mcgee@unh.edu. Please refer to the IRB # above in all correspondence related to this study. The IRB wishes you success with your research.

For the IRB,

A handwritten signature in blue ink that reads "Julie F. Simpson". The signature is written in a cursive style with a large initial 'J' and 'S'.

Julie F. Simpson
Director

APPENDIX C

Center and Classroom Surveys

Center Survey (completed by center director)

1. Program Name
Required to answer. Single line text.
2. Name of person completing this form
Required to answer. Single line text.
3. Title of person completing this form
Required to answer. Single line text.
4. Town where the center is located.
Required to answer. Single line text.
If your program has multiple locations, please indicate the town where the observation for this study will take place.
5. Do you provide a sliding scale for tuition?
Required to answer. Single choice.
If yes, please provide the average or full cost of tuition in Questions 6 and 7.
 - Yes
 - No
6. Average cost of tuition for infants
Required to answer. Single line text.
Please indicate whether this is the weekly or monthly rate.
7. Average cost of tuition for toddlers
Required to answer. Single line text.
Please indicate whether this is the weekly or monthly rate.
8. Number of children enrolled in your program
Required to answer. Single line text.
If there are multiple locations, only include children at the location where the observation will take place.
9. Age range of children enrolled in your program at this location
Required to answer. Single line text.
10. Number of full time classroom staff in all classrooms at this location
Required to answer. Single line text.
11. Number of part time classroom staff

Required to answer. Single line text.

Include teaching assistants and non-credentialed staff.

12. If your program accepts State Scholarships, number of children receiving NH Child Care Scholarship Program funding.

Required to answer. Single line text.

If you do not accept State Scholarships, write "N/A."

13. Center accreditation or licensing level

Required to answer. Multiple choice.

Check all that apply.

- NAEYC
- Early Head Start
- Licensed Plus
- Licensed

14. What is the center's experience with the Pyramid Model?

Required to answer. Multiple choice.

Select all that apply.

- We have never heard of it.
- We have heard of it.
- We have completed some Pyramid Model training.
- We informally use the Pyramid Model approach in our program.
- We are participating in the Pyramid Model Pathway (TPITOS/TPOT) in *Granite Steps for Quality*.
- Observations have shown that we are implementing the Pyramid Model approach to fidelity.

15. What is the center's experience with the Environment Rating Scales (ITERS, ECERS, etc.)?

Required to answer. Multiple choice.

Select all that apply.

- We have never heard of it.
- We have heard of it.
- We have completed an Environment Rating Scales training.
- We self-assess our classrooms with the Environment Rating Scales (ITERS, ECERS, etc.)
- We are participating in the Environment Rating Scales (ITERS/ECERS) Pathway in *Granite Steps for Quality*.
- We have had outside professionals complete Environment Rating Scales (ITERS/ECERS) observations in our classrooms.

16. What is the center's teaching philosophy/pedagogy

Required to answer. Multiple choice.

Select all that apply.

- Waldorf

- Montessori
- Reggio Emilia
- Nature-based
- Resources for Infant Educators (RIE)
- Other

Classroom Survey (completed by participating teacher)

1. Your name (first, last)
Required to answer. Single line text.
2. Total number of children on your roster
Required to answer. Single line text.
3. Average number of children in your classroom each day
Required to answer. Single line text.
Or provide a range such as 4-7 children/day.
4. Age range of children on your roster
Required to answer. Single line text.
Please indicate unit: "months" or "years and months."
5. Number of children on your roster with identified special needs
Required to answer. Single line text.
Emotional, behavioral, physical, etc.
6. Number of children on your roster who you feel require specialized support but who do not have a diagnosis or treatment plan.
Required to answer. Single line text.
7. Number of children on your roster who are dual-language learners *and* who are given individualized support to learn English
Required to answer. Single line text.
8. Number of full time teachers dedicated to your classroom
Required to answer. Single line text.
9. Number of part time, assistant, or float teachers who spend time in your classroom
Required to answer. Single line text.
Please specify "4 staff each week" or "1 part time staff person each day"
10. Average total number of teachers and staff in your classroom each day
Required to answer. Single line text.
Response can be a range or number.

11. Your job title at this center (example: Lead Teacher, Infant Teacher, etc.)
Required to answer. Single line text.
12. Your NH Teaching Credential
Required to answer. Single choice.
- Director Qualified
 - Lead Teacher
 - Associate Teacher
 - Assistant Teacher
 - No official credential
13. Number of years you have worked in early childhood care and education
Required to answer. Single line text.
14. Number of different centers you have worked for
Required to answer. Single line text.
15. Education level
Required to answer. Single choice.
- High school
 - Some college-level courses
 - Associate degree
 - Bachelor's degree
 - Master's degree
16. Did you receive a degree in Early Childhood Care and Education or a related field?
Required to answer. Single choice.
- Yes
 - No
17. What is your gender?
Required to answer. Single line text.
18. What is your experience with the Pyramid Model? (Select all that apply.)
Required to answer. Multiple choice.
TPITOS = Teaching Pyramid Infant-Toddler Observation Scale, TPOT = Teaching Pyramid Observation Tool (used in preschool and pre-K classrooms)
- I have completed an online training.
 - I have completed an in-person training.
 - I learned about the Pyramid Model in a high school or college-level course.
 - I have heard of the TPITOS or TPOT observations.
 - I am a reliable TPITOS or TPOT observer.
 - The TPITOS has been conducted in my classroom.
 - The TPITOS has been conducted on me before.
 - I have received individualized coaching in Pyramid Model strategies.

19. What is your experience with the Environment Rating Scales (ITERS, ECERS, etc.)?
(Select all that apply.) Required to answer. Multiple choice.

ITERS = Infant/Toddler Environment Rating Scale, ECERS = Early Childhood Environment Rating Scale (used in preschool and pre-K classrooms)

- I have completed an online training.
- I have completed an in-person training.
- I learned about the ITERS or ECERS in a high school or college-level course.
- I have heard of the ITERS or ECERS observations.
- I am a reliable ITERS or ECERS observer.
- The ITERS has been conducted in my classroom.
- I have received individualized coaching using the ITERS criteria.

APPENDIX D

Detailed Exposure Data

Table D1

Teacher's Exposure to the Pyramid Model, TPITOS or TPOT

ID	Completed online PM training	Completed in-person PM training	Learned about PM in a high school or college course	Has heard of the TPITOS or TPOT observation	Is a reliable TPITOS or TPOT observer	TPITOS has been conducted in their classroom	TPITOS has been conducted on them	Has received individualized coaching on PM strategies	Exposure rating
A	Yes	-	-	-	-	-	-	-	Low
B	Yes	-	-	-	-	-	-	-	Low
C	Yes	-	Yes	Yes	-	-	-	-	High
D	Yes	Yes	-	Yes	-	Yes	Yes	-	High
E	-	-	Yes	-	-	-	-	-	Low
F	Yes	Yes	-	Yes	-	-	-	Yes	High
G	Yes	-	-	-	-	-	-	-	Low
H	-	-	Yes	-	-	-	-	-	Low
I	Yes	Yes	Yes	Yes	-	-	-	-	High
J	-	-	-	Yes	-	-	-	-	Low

Note. PM is the Pyramid Model approach, TPITOS is the Teaching Pyramid Infant-Toddler Observation Scale, and the TPOT is the

Teaching Pyramid Observation Tool for Preschool Classrooms.

Table D2*Teacher's Exposure to the Environment Rating Scales, ITERS or ECERS*

ID	Completed online ERS training	Completed in-person ERS training	Learned about ERS in a high school or college course	Has heard of the ITERS or ECERS observation	Is a reliable ITERS or ECERS assessor	ITERS has been conducted in their classroom	Has received individualized coaching on the ITERS criteria	Exposure rating
A	-	-	-	Yes	Yes	-	-	High
B	Yes	-	-	Yes	-	-	-	Low
C	Yes	-	Yes	Yes	-	-	-	Low
D	-	-	Yes	Yes	-	-	Yes	High
E	-	-	Yes	-	-	-	-	Low
F	-	-	Yes	Yes	-	-	-	Low
G	Yes	Yes	-	Yes	Yes	-	-	High
H	-	-	-	Yes	-	-	-	Low
I	Yes	Yes	Yes	Yes	-	Yes	-	High
J	Yes	-	Yes	Yes	-	Yes	-	High

Note. ERS refers to the family of Environment Rating Scales products, ITERS is the Infant/Toddler Environment Rating Scales, and

ECERS is the Early Childhood Environment Rating Scale used in preschool and pre-kindergarten classrooms.

Table D3*Center's Exposure to the Pyramid Model*

ID	Have heard of the Pyramid Model	Completed some PM training	Informally use the PM approach	Participating in the PM Pathway in GSQ	Observations have shown that PM is being implemented to fidelity	Exposure rating
A	Yes	Yes	Yes	-	-	High
B	Yes	-	Yes	-	-	High
C	Yes	Yes	-	-	-	Low
D	Yes	Yes	-	Yes	-	High
E	Yes	-	-	-	-	Low
F	Yes	Yes	Yes	Yes	Yes	High
G	Yes	-	-	-	-	Low
H	Yes	-	-	-	-	Low
I	Yes	Yes	-	Yes	-	High
J	Yes	Yes	-	-	-	Low

Note. PM is the Pyramid Model and GSQ is *Granite Steps for Quality*.

Table D4*Center's Exposure to the Environment Rating Scales*

ID	Have heard of the ERS	Completed ERS training	Self-assess using ERS tools	Participating in the ERS Pathway in GSQ	Outside observers have completed ERS assessments	Exposure rating
A	Yes	-	-	-	-	Low
B	Yes	Yes	-	Yes	-	High
C	Yes	-	-	-	-	Low
D	Yes	Yes	-	-	-	High
E	Yes	-	-	-	-	Low
F	Yes	-	-	-	-	Low
G	Yes	Yes	-	-	Yes	High
H	Yes	-	-	-	-	Low
I	Yes	-	-	-	-	Low
J	Yes	Yes	Yes	Yes	Yes	High

Note. ERS refers to the family of Environment Rating Scales products. GSQ is *Granite Steps for Quality*.

APPENDIX E

Granite Steps for Quality Incentives for Non-accredited Centers

Table E1

Quarterly Quality Incentive

GSQ Quality Step	License capacity as of date of application				
	1-25	26-50	51-100	101-200	201+
1	\$600	\$1200	\$1800	\$2400	\$3000
2	\$1200	\$1800	\$2400	\$3000	\$3600
3	\$1800	\$2400	\$3000	\$3600	\$4200
4	\$2400	\$3000	\$3600	\$4200	\$4800

Note. Adapted from Child Care Aware of NH. (2022). *Granite Steps for Quality*. NH

Connections. <https://www.nh-connections.org/providers/quality-care-matters/>

Table E2

Monthly Child Care and Development Fund Scholarship-based Incentive

GSQ Quality Step	Percent of child care scholarship billing
1	5%
2	5%
3	10%
4	10%

Note. The Child Care and Development Fund scholarships are referred to as NH Child Care Scholarships in New Hampshire. “All programs enrolled in and receiving the NH Child Care Scholarship Program funds will automatically receive this payment.” Adapted from Child Care Aware of NH. (2022). *Granite Steps for Quality*. NH Connections. <https://www.nh-connections.org/providers/quality-care-matters/>