Hansen *et al.*
International Journal of Child Care and Education Policy (2026) 20:17
https://doi.org/10.1186/s40723-026-00184-3

International Journal of Child Care and Education Policy

RESEARCH Open Access

A new approach to predicting early cognitive outcomes from ECEC quality rating scales

Joakim E. Hansen[1*], Erik Eliassen[2], Elisabeth Bjørnestad[3], Edward Melhuish[4,6,7,8,9] and Jacqueline Barnes[4,5]

*Correspondence:
Joakim E. Hansen
joakim.e.hansen@uis.no

Full list of author information is available at the end of the article

Abstract
The quality of ECEC has a significant influence on the development of cognitive skills; however, it is still unclear which dimensions of quality are most relevant. Data from two quality measures (ITERS-R/ ECERS-R) were used to model new latent quality factors designed to describe cognitive stimulation and these factors were used to predict cognitive development. The sample included 1096 children from 93 centers enrolled in 206 toddler groups and 205 preschool groups. Findings indicated that the ITERS-R cognitive stimulation factor was associated with verbal and non-verbal cognitive development at three years, and the ECERS-R cognitive stimulation factor was associated with verbal and non-verbal cognitive development at age five. Although these associations were small in magnitude, they were stronger than for the original scales' total scores. The findings suggest that the new factors represent a useful additional way to document and understand ECEC quality and its relation to children's cognitive development.

Keywords ECEC quality, cognitive development, ITERS-R, ECERS-R

Introduction

Children's cognitive skills develop rapidly during early childhood, and learning experiences in the early years, both in the home and in Early Childhood Education and Care (ECEC), can have a significant influence on the development of cognitive skills and later academic outcomes (Belsky et al., 2007; Keys et al., 2013; Knudsen et al., 2006; Vandell et al., 2010). The benefits of attending ECEC are found to be greater when the quality of care is high, especially for children living in families with low income (Dearing et al., 2009; Melhuish et al., 2015). The quality of pedagogical processes in ECEC is also accepted as a proximal determinant of early learning and child development (Early et al., 2007; Slot et al., 2015). However, the way quality should be defined and measured, and how different dimensions of quality are related to developmental domains, are important issues to explore (Burchinal et al., 2021). A recent meta-analysis of European longitudinal studies found small positive associations of process quality (the nature of interactions experienced) with cognitive development during early childhood and lasting associations with academic outcomes (Ulferts et al., 2019). Lasting effects of ECEC quality on academic outcomes, although small in magnitude, were also identified in a meta-analysis

Springer

by Rademacher et al. (2024). However, although there is agreement on the importance of quality, and particularly process rather than structural factors when related to child development (Howard et al., 2024), there is a need for a refinement of quality measures related to gains in specific developmental outcomes (Burchinal et al., 2021).

In this study, we make use of a new and innovative approach to develop domain-specific quality factors from existing quality measures. The aim is to investigate associations between newly developed quality of care scales, in terms of cognitive stimulation, and cognitive development in a universally accessible ECEC system in toddler and preschool groups. In addition, having data following the same children at two time points makes it possible to look at differences and stability over time since existing individual experiences are accounted for. Such information on the relationship between quality of care and child development may be used to improve policy and practices and in turn provide children with good play and learning experiences in early childhood.

The Environment rating scales (e.g., Infant and Toddler Environment Rating Scale, Revised [ITERS-R], Harms et al., 2006, and Early Childhood Environment Rating Scale, Revised [ECERS-R], Harms et al., 2005) are widely used instruments to assess quality in ECEC centers, both by practitioners and for research (Alpys & Hernández-Torrano, 2026). In addition to its use as a quality assessment tool, a scoping review identified four additional applications of the scale: influence on children's academic and non-academic outcomes, as an intervention assessment tool, as a complementary instrument alongside other methods, and as a means of examining cultural nuances in early childhood settings (Alpys & Hernández-Torrano, 2026). These scales were developed to measure the global quality of the environment from the perspective of the children (Clifford et al., 2020), and they incorporate several subscales related to different developmental domains. Previously, some studies on the scales' psychometric properties have identified problems with criterion and construct validity (Fujimoto et al., 2018; Gordon et al., 2013, 2015; Mayer & Beckh, 2016; Mariano et al., 2019; Perlman et al., 2004). For construct validity, only a few studies have identified a unidimensional quality factor in line with the developers' intention of a global quality measure, and there are somewhat inconsistent recommendations regarding factor structure based on information at the item level (Gordon et al., 2013; Hestenes et al., 2007; Mariano et al., 2019; Mayer & Beckh, 2016). The scales were conceptually designed on three levels: subscales (level 1), consisting of items (level 2), which are scored using dichotomous indicators (level 3). When scoring the indicators (observational response categories scored yes/no), low-quality indicators must be satisfied before high-quality indicators can be assessed (i.e., stop-scoring procedure). However, the indicators do not necessarily follow an ordinal progression with respect to quality, and there may be aspects of different quality dimensions within the indicators for an item (Fujimoto et al., 2018; Gordon et al., 2015; Mayer & Beckh, 2016). Indicators at the low quality end of an item are often structural, about basic needs, health, safety and availability of materials, whereas indicators of medium and high quality often include more educational aspects, interaction and instruction, indicating process quality. Applying a stop-scoring procedure can result in limited information at the upper end of the scale, which may be particularly problematic for dimensions of quality (educational aspects, interaction and instruction) strongest related to child outcome. An alternative scoring procedure was used when collecting data for this study, ignoring the stop-scoring rule, meaning that all indicators was scored to provide a more complete

view of quality. Using this alternative scoring approach makes it possible to extract information on indicator level, contrary to previous studies using information based on item level data (Brunsek et al., 2017).

Studies of the scales' criterion validity have focused on the ability to predict child development from total scores and quality factors derived from item level values and have shown associations with cognitive and socio-emotional development (Burchinal et al., 2000; Sylva et al., 2006). More recent meta-analysis verified these findings, albeit finding moderate to small effect sizes (Brunsek et al., 2017; Keys et al., 2013; Ulferts et al., 2019). Studies attempting to predict cognitive development from quality of care often include preschool-aged children, but there is less evidence on toddler-aged children (Li et al., 2013; Melhuish, 2015; NICHD Early Childhood Research Network, 2006; Ulferts et al., 2019). For preschool-aged children the results are usually consistent in showing positive effects upon development of attending high-quality ECEC (Melhuish et al., 2015). For toddler age, there are less consistent associations between quality of care and cognitive development (Barnes & Melhuish, 2017; Eliassen et al., 2018; Ruzak et al., 2014).

In this study, we address the validity issues identified in the literature by taking a latent variable approach to predict child cognitive outcomes from newly developed quality factors derived from the indicator level of the quality rating scales (ITERS-R/ ECERS-R). With information on all quality levels, a latent approach helps us to reduce the dimensionality of data into an underlying concept of quality (cognitive stimulation). In other words, latent modeling allows us to better understand the underlying structures that produce relations between observed indicators, while simultaneously reducing error variance (measurement error), which often is an issue in observational data (see Beaujean, 2014). This study contributes to the existing literature on quality in ECEC by refining quality measures for both under and over three-year-olds in order to identify specific dimensions of the environment that are most useful for facilitating children's cognitive development. Moreover, this is, to our knowledge, one of the first studies to include both ITERS-R and ECERS-R to predict child outcomes for the same children over time.

Measuring quality in ECEC using environment rating scales

The most commonly used environment rating scales (ITERS & ECERS) were initially developed in the 1970s in the USA to judge the level of quality provided from the perspective of the children, with a strong focus on health and safety (Clifford et al., 2020). The two scales used in this study, (ITERS-R and ECERS-R), are revised versions of the original scales intended to capture development promoting aspects of caregiver-child interactions more effectively (Clifford et al., 2020). However, despite the revisions, the scales have been criticized for still emphasizing structural aspects of the environment (e.g., amount of play materials, books, safety aspects) over core interactional dimensions of process quality (Perlman et al., 2004). The third edition of one of the scales (the ECERS-3; Harms et al., 2015), contains new items on language and literacy, with less emphasis on the availability of materials and more on the staff's use of it with the children but was not available at the time of data collection for the current study. Neitzel et al. (2019) compared ECERS-R to ECERS-3 and found significant changes in terms of more focus on staff-child interactions and instruction (dimensions of high-quality). However, they found the scales to be associated, indicating that the scales are measuring common characteristics of the learning environment in ECEC (Neitzel et al., 2019).

The ITERS-R and ECERS-R scales were, like the original scales, designed to produce a unidimensional global quality measure, with subscales representing seven domain-specific dimensions of quality. However, the evidence on their psychometric properties is mixed with regard to identifying the unidimensional quality measure and the seven-factor structure in line with the intentions of the developers. Perlman et al. (2004) did not find evidence for the seven-factor structure and also argued for a unidimensional quality factor based on a subset of the items in the ECERS-R. Other researchers have found evidence of a two-factor structure (Cassidy et al., 2005; Early et al., 2007) and a three-factor structure (e.g., Brunsek et al., 2017; Gordon et al., 2013; Mayer and Beckh 2018). These studies based their analysis of psychometric properties on information at the item level. A recent study on the third edition, ECERS-3 (Early et al., 2018) found similar results in terms of factor structure and validity. They found a single factor did not adequately capture the variance among the items but instead identified a four-factor solution: Learning Opportunities, Gross Motor, Teacher Interactions, Math Activities. When predicting child outcomes, the total score and Teacher Interaction was associated with growth in executive functioning, the Learning Opportunities was associated with executive functions and math skills, and the Math Activities was associated with social skills. However, the associations were small and not domain specific (Early et al., 2018).

One issue identified is that the indicators, while intended to progress in an ordinal way from low to high quality, do not always follow an ordinal progression. The manual instructions require that low-quality indicators must be passed before progressing to consider higher quality indicators. Consequently, the resulting score on a single item depends on the number of lower-level indicators satisfied. This approach is referred to as a "stop-scoring approach". Issues with the stop-scoring approach are addressed by Fujimoto, Gordon Peng, and Hofer (2018), who found that none of the 36 items in the ECERS-R scale had categories that followed an ordinal progression with respect to quality. One reason for this may be that structural indicators (relating to safety and availability of materials, space and furnishing) are placed at the lower end, and interactions and instruction (process) are in the mid and high categories. Gordon et al. (2015) conducted an item response analysis of all indicators and found that the items did not fit with unitary dimensions and the results conflicted with the range-ordering of indicators. A panel of experts rated the domain specificity of the indicators, and showed that most indicators were relevant to multiple developmental domains (i.e., cognitive, socio-emotional, and health) – 26% of the indicators were rated relevant for all domains and 57% were relevant to both cognitive and socio-emotional development (Gordon et al., 2015). Moreover, the analysis showed that few indicators captured moderate to high quality. In a European context Mayer and Beckh (2016), replicating Gordon et al. (2013), found consistent results in that the indicators did not follow the specified low to high quality ordering, and possibly too many of the indicators are related to structural components (Cassidy et al., 2005).

Thus far, few studies have verified the factor structure, or a unidimensional global quality measure, intended by the developers of the ITERS-R and ECERS-R (Brunsek et al., 2017; Cassidy et al., 2005; Early et al., 2007; Perlman et al., 2004). Additionally, studies have identified issues with items within subscales and indicators within items (Fujimoto et al., 2018; Gordon et al., 2013). Despite these issues, the rating scales do capture information on important aspects of quality (Brunsek et al., 2017; Gordon et al., 2015;

Ulferts et al., 2019). Most studies have used information on item or subscale level (Cassidy et al., 2005; Early et al., 2007; Mayer & Beckh, 2016), with only a few having information on all indicators (Fujimoto et al., 2018; Gordon et al., 2015). However, no studies have used information derived from indicator level, based on all indicators by ignoring the stop-scoring rule, to develop alternative measures of quality. This would allow for all the higher level process indicators to be included.

Predicting cognitive development from elements of environment rating scales

The relevance of attending center-based care for early learning and development has been reviewed and meta-analyzed in several studies, showing the importance of high-quality group care (Brunsek et al., 2017; Ulferts et al., 2019; Rademacher et al., 2025). A comprehensive review of the impact of ECEC on child development found consistent positive results for preschool age children (aged three to five), with more mixed results for infants and toddlers (Melhuish et al., 2015). However, a recent literature review focusing specifically on process quality for children under three found positive links between quality and infant-toddler development (Cadima et al., 2020). Meta-analyses investigating the relationship between global quality measures and child development finds modest effects of quality (Brunsek et al., 2017; Keys et al., 2013 Perlman et al., 2016; Ulferts et al., 2019). A meta-analysis by Ulferts et al. (2019) including European studies on the impact of process quality on academic outcomes found an overall effect size of 0.11 SD for global measures, and 0.10 SD for domain-specific measures. Brunsek et al. (2017) conducted a meta-analysis of studies using ECERS and ECERS-R and found overall weak positive associations with language and social-emotional development. For language development, the pooled correlation was 0.05 for the total score, and 0.07 for the Language-Reasoning subscale. No significant results for the commonly used factors 'Provision for learning' and 'Teaching and interactions' were found (Brunsek et al., 2017).

Studies predicting child development from the environment rating scales have identified positive associations from the total score and alternative factor solutions (Gordon et al., 2013; Mayer Bech 2016; Sylva et al., 2006). Gordon et al. (2013) included both the total score and a three-factor solution and found few and small associations with child outcomes while controlling for child and family characteristics. Similarly, Mayer and Beckh (2016) found cognitive development to be positively related to the 'Language and interaction' and 'Space & Furnishing/ Activities/ Program' factor. However, these associations became non-significant when applying the Bonferroni-Holm correction (Mayer & Beckh, 2016). In the British EPPSE study, ECERS-R and an extension of the scale including curricular subscales ECERS-E were used to capture global process quality (Sylva et al., 2006). The results showed positive effects of the ECERS-E scale for cognitive/linguistic development, adjusting for demographic characteristics; however no significant associations were found using the ECERS-R. Based on the evidence, it can be concluded that quality factors that include more educational aspects (communication, interaction, instruction) predict children's cognitive development better than global measures of quality that include aspects related to children's basic needs (Sylva et al., 2006; Burchinal et al., 2021; Howard et al., 2024). Furthermore, studies exploring child-level measures targeting experiences of individual children, specifically related to developmental support, have confirmed this by showing promising results in predicting child outcomes (Burchinal et al., 2021, Justice et al., 2019).

Some Norwegian studies have included measures of ECEC quality when studying the relationship between ECEC and child development (Eliassen et al., 2018; Hansen & Broekhuizen 2020; Løkken et al., 2018). Eliassen et al. (2018) found no associations between quality and cognitive development at age three using the ITERS-R total score. Hansen and Broekhuizen (2020) found a small association with quality using items from the subscale 'Listening and Talking' from the ITERS-R on verbal development at age five, adjusted for verbal development at age three. Similar results were also found in a study by Løkken et al. (2018), predicting social-emotional developmental from the subscale 'Interaction' in the ITERS-R. However, no significant effects were found when using the original ECERS-R total score and subscales for cognitive development age five. Cultural considerations and contextualization are important when using quality rating scales developed in the US in other cultural contexts (Alyps & Hernandez-Torrano, 2026). The scales have been shown to capture aspects of the pedagogical quality valued in the Nordic tradition and curriculum, and the ITERS-R demonstrated a high degree of consistency with the Norwegian framework (Bjørnestad et al. 2019; Garvis et al. 2017).

Recent studies on the predictive ability of the scales have used quality factors based on information at the item level (Gordon et al., 2013; Mayer & Beckh, 2016). There are a few studies operationalizing domain-specific quality factors from information at the indicator level (Gordon et al., 2015), but currently there are no studies predicting child outcomes from latent variables derived from the indicator level. We investigate the potential of developing new quality factors and relating them to cognitive outcomes in a context of universal accessible ECEC in Norway.

The present study

In the present study, we explored quality factors from the ITERS-R and ECERS-R specifically related to cognitive stimulation through latent variable analysis at the indicator level, and then examined the associations between latent cognitive stimulation quality factors and children's cognitive development at toddler and preschool age in the context of the universally accessible ECEC system in Norway.

This study contributes to existing research literature in two ways. First, the study restructures the data from scales using information at the indicator level (and including all indicators by ignoring the 'stop-scoring' rule) to identify domain specific (cognitive stimulation) factors that might predict children's cognitive development. The expectation is that a specific quality measure might be more closely linked to specific child outcomes (i.e., a domain-specific approach) as suggested in previous research (Burchinal et al., 2021; Brunsek et al., 2017). Second, by including data from toddlerhood and preschool age, the study investigated the potential effect of quality during two sensitive developmental periods where investment is important for later attainment (Knudsen et al., 2006). Thus, the main research question was to determine the extent to which domain-specific cognitive stimulation quality factors derived from the ITERS-R or ECERS-R indicators were associated with cognitive development in toddler and preschool age in a universal accessible ECEC system. It was hypothesized, based on previous evidence, that there would be positive associations with cognitively stimulating dimensions of the environment on child cognitive development.

Method

Participants and Procedure

The participants were from the longitudinal research project XX. The study was approved by the Norwegian Centre for Research Data, in accordance with national ethical guidelines. Using a stratified random sampling procedure, 158 ECEC centers in six regions in Norway were approached to participate in the study, and 93 centers agreed to participate. All parents with children born in 2011 and 2012 attending one of the selected centers were asked to participate and the parents provided written informed consent of participation on behalf of their children. The study includes 1200 children enrolled in 93 centers located in six different regions in Norway, enrolled in 206 toddler groups and 205 preschool groups. In Norway, typically children are moved from an infant/ toddler group to a preschool group within the same center the year that they turn three. This study included all children with non-missing outcome data, resulting in an analytical sample consisting of 1096 children (48% girls, 24% with one or both parents born outside of Norway), located in six different regions in Norway (region 1 = 19%; region 2 = 22%; region 3 = 14%; region 4 = 21%; region 5 = 20%; region 6 = 4%). The average age at first cognitive assessment as 2.96 years (*SD*=0.21) at 5.02 years (*SD*=0.12) at the second assessment Data for this study include environment quality ratings conducting during observations in the respective child groups, in addition to direct assessment of children's verbal and non-verbal cognitive development. The quality measures were conducted by certified ITERS-R and ECERS-R researchers who followed the procedures described by the Environment Rating Scales Institute (ERSI – ersi.info). All the ITERS-R and ECERS-R observers in this study were trained by the third author, who had been certified by the developers of the ITERS-R and ECERS-R. The observers reached a minimum 80% reliability at item and indicator level prior to conducting observations. The ratings were conducted during a one-day visit and were based on three to four hours of observation, followed by an interview with the preschool teacher to get additional information on quality indicators that could not be observed (e.g., frequency of activities and how the books available was chosen). The questions asked followed the procedure described in the ITERS-R/ ECERS-R manuals. The child assessments were conducted in surroundings familiar to the child in the ECEC center, and children were accompanied by a familiar staff member. Prior to the assessment, the data collector informed the parents and asked the staff members to inform and prepare the child. The staff members were told that they were participating to provide emotional support but were instructed not to assist the children with the tasks.

Measures

Cognitive outcomes. Cognitive development was assessed with (Norwegian translations of) two subscales from the British Ability Scale 3 (Elliot & Smith, 2011): naming vocabulary and picture similarities. These measures were chosen because of their predictive power for children's long-term educational and social development. Naming vocabulary measure cognitive verbal ability by assessing expressive language vocabulary. Children were shown a sequence of 36 pictures of everyday objects and asked to name them. Picture similarities measures non-verbal ability by assessing non-verbal reasoning on 35 items. Children were asked to match a picture on a card with one of four pictures in a booklet that had a common concept or element. For both scales, children were given

one chance on each item, and answers were scored as 1 (correct) or 0 (incorrect). Previous studies have used ability scores in accordance with the scoring manual that reflects both the numbers of items answered correct and the difficulty of items taken (Eliassen et al., 2018, Hansen & Broekhuizen, 2020). In this study, the cognitive outcome measures were modelled as latent variables, an approach shown to be appropriate when using BAS 3 in a Norwegian context (Elnes et al., 2024). By this approach, we reduce measurement error and estimate a cleaner measure of the underlying construct relating to cognitive abilities (see Beaujean, 2014).

ECEC quality. Quality in toddler-groups (0–3 years) was measured using the Infant/Toddler Environment Rating Scale-Revised (Harms et al., 2006) and in preschool groups (3–6 years) with the Early Childhood Environment Rating Scale-Revised (Harms et al., 2005). The scales include over 400 indicators covering aspects of child-care quality across a wide range of developmental, personal health and safety domains. The indicators make up 39 items in the ITERS-R and 43 items in the ECERS-R covering seven dimensions of quality. The indicators are ordered from low to high quality on a 1–7 point Likert-type scale (1 = inadequate, 3 = minimal, 5 = good, 7 = excellent). There are multiple indicators per quality level and the number of indicators differs between quality levels (1,3 or 5) within the items. To obtain an item score, the observer is directed in the manual to start scoring indicators on the lowest level of quality and stops when not all indicators are met at a specific level (1, 3 or 5). This "stop-scoring" procedure means that all indicators at a specific level need to be obtained before evaluation can proceed to the next quality level for the item. The items are further grouped into seven sub-scales, such as "space and furnishing" and "interactions". Given the high number of indicators and items, the subscales are intended to ease the interpretation of the results through distinct domains of quality. A subscale score is a mean of the items included in the subscale, and the total score is a mean of all items in the scale. The total score is often referred to as a measure of global quality. The data gathered for this study did not follow the 'stop-scoring procedure' but used an alternative scoring procedure, meaning that all indicators were scored. Rather than using the subscales, this procedure provides the possibility to identify alternative domain-specific factors, and to determine if they more accurately describe cognitive stimulation. Specifically, cognitively stimulating indicators were selected from the ITERS-R and ECERS and were used to model latent quality variables.

Statistical approach

Confirmatory factor analysis was used to model quality factors specific to cognitive stimulation for toddler and preschool groups from the ITERS-R and ECERS-R scales. The initial selection of indicators to include in the analysis was theoretically driven and included those thought to relate to cognitive stimulation. The initial selection of indicators was done by all the authors, who are a recognized team of experts in developmental and educational psychology, using a Delphi technique (Linstone & Turoff, 2002) to reach consensus. This selection included 131 out of 378 indicators from the ITERS-R, and 140 out of 383 indicators in the ECERS-R. The subscale Parents and Staff was not measured since the content is not directly related to children's learning environment. The subscale has also been excluded in the third edition of the scale – ECERS3 (Harms et al., 2015). The quality indicators that had little or no variance were excluded prior to the confirmatory analysis (ITERS-R: $N = 42$, ECERS-R: $N = 37$). The discarded indicators were mostly

found at indicator level three (ITERS-R: $N = 35$, ECERS-R: $N = 27$). Some theoretically relevant indicators related to availability of toys and material (ITERS-R: $N = 37$, ECERS-R: $N = 44$) and interactional aspects (ITERS-R: $N = 7$, ECERS-R: $N = 5$) were also excluded from the final factors due to low factor loadings. The final factors included 45 indicators from ITERS-R and 54 indicators from ECERS-R. During the development of the quality factors, a substantial reduction of indicators was observed from the initial theoretical model to the final model. The model-fitting process, undertaken to establish a model that is both theoretically and empirically grounded, resulted in an overall good model fit and high factor loadings The indicators and factor loadings included in the final solution are presented in the appendix, Tables 4 and 5.

The outcome measures, verbal (expressive language) and non-verbal (reasoning) cognitive development age three and five, were also modelled as latent variables using confirmatory factor analysis. Initially, all age specific items were included in separate analyses for each outcome measure. For verbal ability at age three, items 1–24 were included, and for age five, items 11–36 were included. For non-verbal ability at age three, items 1–18 were included and for age five, items 12–35 were included.

Several confirmatory factor models, using lavaan in R (Rosseel, 2012), were constructed. All variables were treated as ordinal in the factor analyses, using the robust weighted least squares estimator (WLSMV) (Nye & Drasgow, 2011). Model fit was reported with the following robust goodness-of-fit indices: root mean square of approximation (RMSEA), comparative fit index (CFI) and Tucker-Lewis index (TLI). Missing cases were dealt with using pairwise likelihood (PL) imputation in all factor analyses.

Analyses estimating the relation between the derived cognitive stimulation quality factors and children's cognitive development were conducted using Ordinary Least Square (OLS) regressions. Latent variables for cognitive stimulation quality and cognitive development were modelled prior to inclusion in the structural model and the factor scores were used as continuous variables in the regression analyses. The cognitive stimulation quality factors and total scores in toddler and preschool groups were regressed on verbal and non-verbal skills age three and five, respectively, to produce estimates on children's cognitive development. The total scores were included because there is no equivalent subscale related to cognitive development. In all models, potential regional confounders were controlled by including region as dummy variables (six different regions). Listwise deletion was used to handle missingness in all regression models. Reported regression estimates are standardized on both outcome and predictor variables, resulting in regression coefficients (β) that represent the predicted standard deviation difference in cognitive outcomes given a one-unit increase in quality.

Results

Descriptive statistics for the cognitive non-verbal and verbal outcomes, in addition to quality measures, are presented in Table 1. The cognitive stimulation quality factors developed in this study and the total scale score correlated strongly for ITERS-R (r=.81, p<.001) and ECERS-R (r=.81, p<.001). The measures of non-verbal and verbal skills correlated modestly at age three (r=.33, p<.001) and age five (r=.26, p<.001), showing that they measure two related but distinct aspects of cognition.

Measurement models. All measurement models produced statistics indicating adequate to good model fit (see Table 2). The only post-hoc modification done was that

Table 1 Descriptive statistics of child cognitive outcome and ECEC quality measures

Variables	*n*	Mean	SD
Non-verbal skills age three	1096	-0.03	0.44
Verbal skills age three	1078	-0.03	0.61
Non-verbal skills age five	1028	-0.02	0.37
Verbal skills age five	1029	-0.04	0.84
ITERS-R Cognitive Stimulation Factor	1070	0.05	0.61
ITERS-R Total Score	1087	3.96	0.77
ECERS-R Cognitive Stimulation Factor	915	-0.02	0.50
ECERS-R Total Score	915	4.15	0.85

Table 2 Robust goodness-of-fit indices for child outcome and quality factors

	RMSEA	CFI / TLI
Outcome factor		
Verbal skills age three	0.032 (0.028-0.036)	0.968/0.965
Non-verbal skills age three	0.023(0.017-0.029)	0.947/0.940
Verbal skills age five	0.020(0.015-0.025)	0.967/0.964
Non-verbal skills age five	0.021(0.016-0.026)	0.941/0.935
Quality factor		
ITERS-R Cognitive Stimulation	0.043 (0.037-0.049)	0.935/ 0.932
ECERS-R Cognitive Stimulation	0.041 (0.036-0.046)	0.926/0.924

RMSEA=Root Mean Squared Error of Approximation, CFI= Comparative Fit Index, TLI= Tucker-Lewis Index

Table 3 Regression analysis results, significant factors and total scores from ITERS-R and ECERS-R associated with child cognitive outcomes

	Dependent variable:			
	Non-verbal age three	Verbal age three	Non-verbal age five	Verbal age five
	B(SE)	B(SE)	B(SE)	B(SE)
ITERS-R Cognitive Stimulation	0.074 (0.032)*	0.065 (0.033)*		
ITERS-R Total Score	0.008 (0.031)	0.042 (0.032)		
ECERS-R Cognitive Stimulation			0.095 (0.037)*	0.110 (0.037)**
ECERS-R Total Score			0.094 (0.035)**	0.069 (0.036)
Observations	*1069*	*1063*	*906*	*905*

Covariates have been omitted from the presentation. *$p < .05$; **$p < .01$; ***$p < .001$

items 31 and 36 were discarded from verbal ability at age five due to low factor loadings. No post-hoc modifications were conducted for the quality factors (factor loadings are in Tables 4 and 5).

Regression estimates. Examining the relation between cognitive stimulation quality factors and children's cognitive outcomes, cognitive development varied as a function of cognitive stimulation quality after adjusting for region (see Table 3). Table 3 shows the standardized regression coefficient for the cognitive stimulating quality factors in toddler- and preschool groups and cognitive outcome at age three and five. For children under three, the cognitive stimulation factor derived from ITERS-R predicted non-verbal cognitive development ($\beta = 0.065$, p=.048) and verbal cognitive development age three ($\beta = 0.074$, p=.024), after adjusting for region. No significant associations were found between the ITERS-R total score and either non-verbal ($\beta = 0.042$, p=.185) or verbal skills age three ($\beta = 0.008$, p=.812), after adjusting for region. For children over

three years, the cognitive stimulation factor derived from ECERS-R predicted non-verbal cognitive development (β = 0.095, p= .011) and verbal cognitive development age five (β = 0.110, p=.003), after adjusting for region. For cognitive outcomes at age five years, the ECERS-R total score was significantly associated with non-verbal skills (β = 0.094, p=.007), but not with verbal skills (β = 0.069, p=.052) - although it approached significance.

Discussion

In this study, two latent quality factors were established using selected cognitively stimulating indicators from the ITERS-R and ECERS-R. The latent factors were found to have significant associations with children's cognitive development at toddler and preschool age in the context of a universally accessible ECEC system in Norway. The ITERS-R cognitive stimulation factor was positively associated with non-verbal and verbal cognitive development when the children were aged three years, and the ECERS-R cognitive stimulation factor was positively associated with non-verbal and verbal cognitive development at age five years. By contrast, the ITERS-R total score did not predict cognitive development for children aged three years, and ECERS-R total score only predicted non-verbal cognitive development at age five years.

The associations found between the cognitive stimulating quality factors and cognitive outcomes in this study were significant but small in magnitude (ranging from 0.06 to 0.11), which replicates results from prior studies (Burchinal et al., 2021; Early et al., 2018; Ulferts et al., 2019). The magnitude of the effect sizes can be considered small by conventional standards, with values ranging from 0.06 to 0.11 (in standard units), corresponding to small effects according to Cohen's classification of effect sizes (1998). However, interpreting the meaning of these effect sizes is not straightforward as their applicability can vary across research contexts (Belsky et al., 2007). Given that the positive effects of ECEC quality may be experienced by all children, even small effects can result in substantial cumulative impacts (Belsky et al., 2007). Moreover, as research suggests that these effects may persist into the school years (Eliassen et al., 2024; Elnes et al., 2026), small effect sizes can be understood as having modest but lasting implications for children's development. This suggests that such effects may accumulate over time, influencing developmental trajectories into adolescence (Sluiter et al., 2025) and pointing at that effect sizes in absolute terms may not be unimportant in practical terms.

Given the small effects sizes, the results suggest that a cognitive stimulating environment has a modest, but reliable effect on three- and five-year old outcomes, with marginally stronger associations for the ECERS-R cognitive stimulation factor in preschool years compared to the ITERS-R cognitive stimulation factor in toddler age. This is aligned with previous research that, to a large extent, have been consistent about the positive effect of quality in preschool years (Melhuish et al., 2015). However, finding positive associations between quality in toddler care and child development has been more challenging. In a UK study, Barnes and Melhuish (2017) found quality in toddler care marginally predicted cognitive and verbal scores at age 51 months, but not non-verbal skills. Using the total score from ITERS-R, Eliassen et al. (2018) did not find quality predicted verbal or non-verbal cognitive skills age three years, while Hansen and Broekhuizen (2020) found only one specific item from ITERS-R to be significant associated with cognitive verbal skills age five years. Few studies have included quality ratings from both

toddler and preschool child groups. The findings from this study contribute by proposing a more robust and reliable way of modelling the relationship between ECEC quality, measured on the basis of a selected subset of indicators, and child development.

The cognitive stimulating quality factors did predict cognitive outcomes, albeit with small effect sizes, in contrast to the total scores from the standard rating scales. In this study theoretically relevant indicators were selected, which differs from previously used strategies. Previous research has shown that the indicators within the items do not follow an ordinal progression with respect to quality, which creates issues with capturing high-quality when applying the standard stop-scoring procedure (Fujimoto et al., 2018; Gordon et al., 2015; Mayer & Beckh, 2016). This approach was possible by using an alternative scoring procedure. In the initial model, level 3 indicators were initially included but then discarded when they showed low variation and/ or low factor loadings. All but two indicators included in the final factors were at level 5 or level 7, measuring good to excellent quality. This shows that, when using quality rating scales in a Norwegian context, there was a need for more fine-grained measures capturing variation in mid to high quality provisions. This issue is likely to arise in countries where state-funded provision with quality and/or inspection criteria with a focus on structural aspects of care limit very low quality provision. Based on the results in this study, it can be argued that indicator level information better captures the core aspects of quality in group ECEC.

Beside the validity issues, a refined quality measure can be easier to handle and use, both in terms of scoring and measuring high-quality since fewer items need to be assessed. After a period of optimism in using quality rating scales to predict child outcomes, there seems to be a shift in direction after several reports of weak associations or failure to find significant associations. There seems to be an agreement in the field of ECEC that staff-child interactions should be included in measures of quality based on developmental theories (Bronfenbrenner & Morris, 2006; Vygotsky, 1978). However, the challenge is to identify which dimensions of staff-child interactions and staff's educational practices are related to children's developmental outcomes in different domains. Interactions with responsive and sensitive staff members providing children with activities within their zone of proximal development is argued to be beneficial for cognitive development (Melhuish et al., 2015). The quality factors developed in this study included mainly indicators related to adults' support of children's interactions, communication and play, and can be a useful alternative way to summarize ITERS-R and ECERS-R data in order to understand ECEC quality at group level related to cognitive development. Thus, this study makes a theoretical contribution in conceptualizing ECEC quality in relation to cognitive development.

Overall, more research is needed to explore which dimensions of ECEC quality promote child development and how these can best be observed. The weak or no correlations between quality scores and child outcomes reported in previous research (Howard et al., 2024; Rademacher et al., 2025) may suggest that the measures used do not adequately capture meaningful aspects of ECEC quality. However, the findings from this study, together with those of other studies (Burchinal et al., 2021; Sylva et al., 2006), highlight the potential of quality measures targeting specific developmental domains, in addition to a more holistic approach that assesses the overall quality of children's play and learning environment and early childhood programs. Giving the recommendation of scoring all indicators in the new ECERS-3 (Harms et al., 2015), and the emphasis on

educational and instructional practices more related to child development (Neitzel et al., 2019), there is potential in developing quality factors more strongly related to child outcomes from this scale. We have in this study selected indicators of quality that were considered most appropriate for supporting children's cognitive development and relevant in a Norwegian context. Building on the results from this study, more refined quality measures capturing mid to high quality seems relevant and could be explored further in different ECEC systems and related to other developmental domains, such as social competence, and children's well-being.

There are limitations of this study. First, the data used in this study are observational, and therefore the results cannot be interpreted as causal. Second, analyses did not include covariates at the child level, e.g., home learning environment, parental education, income, and do not consider the potential moderating effects of family factors. Third, the study did not use the latest edition of the ECERS scale, ECERS-3, which has been revised following research findings as cited here, to theoretically better capture the teachers support for children's development of cognitive skills. However, a study by Early et al. (2018) found that the third edition had validity issues related to structure and prediction of child outcomes similar to those identified when using the ECERS-R and Neitzel et al. (2019) found ECERS-R to correlate with ECERS-3. Finally, there could be some limitations related to the measures of child outcomes (Burchinal, 2018). The measures used have previously shown good psychometric properties (Elliot & Smith, 2011) and valid to use in a Norwegian context (Elnes et al., 2024). And a replication on other child outcomes might result in different results for the total score and the cognitive stimulation factor developed in this study.

The results from this study have implications for policy and practice. Considering the importance of early years for development, and that the investments in this period are particularly productive (Knudsen et al., 2006; Melhuish & Barnes, 2021). The factors identified in this study offer a valuable complement to existing measures for documenting and understanding ECEC quality and its relationship to children's cognitive development. Specifically, staff-child interactions, including both the quantity and quality of communication, as well as the engagement of all children, were found to predict cognitive outcomes. Quality measures can inform professional development; however, they must be adapted to contextual and cultural conditions and integrated with complementary assessment approaches. When implemented thoughtfully, the cognitive stimulation measures can support professional development and contribute to an ECEC environment that supports children's learning, development, and well-being.

Conclusion

In summary, the findings suggest that the quality of care is important both during toddlerhood and preschool period, but that more refined measures of quality can be developed if the focus is targeted on the outcome of interest. The results suggest that there is a reasonable level of stability in the associations between factors of quality focusing specifically on cognitive stimulation in ECEC and cognitive development during two developmental periods: toddler and preschool age.

Appendix

Table 4 Table of factor loadings for ITERS-R cognitive stimulation factor

Indicator	Estimate	Item*
Indicator 7.7.1	0.699	Meals/ Snacks
Indicator 12.5.1	0.818	Helping children understand language
Indicator 12.5.4	0.806	Helping children understand language
Indicator 12.7.1	0.796	Helping children understand language
Indicator 12.7.3	0.675	Helping children understand language
Indicator 13.5.1	0.795	Helping children use language
Indicator 13.5.2	0.797	Helping children use language
Indicator 13.5.3	0.823	Helping children use language
Indicator 13.7.1	0.737	Helping children use language
Indicator 13.7.2	0.587	Helping children use language
Indicator 13.7.3	0.577	Helping children use language
Indicator 13.7.4	0.500	Helping children use language
Indicator 14.5.4	0.349	Using books
Indicator 17.5.3	0.356	Art
Indicator 17.7.2	0.417	Art
Indicator 19.5.1	0.380	Blocks
Indicator 19.7.2	0.397	Blocks
Indicator 20.7.3	0.421	Dramatic play
Indicator 21.3.2	0.545	Sand/ water play
Indicator 21.5.2	0.474	Sand/ water play
Indicator 21.5.3	0.497	Sand/ water play
Indicator 25.5.3	0.832	Supervision of play and learning
Indicator 25.5.4	0.843	Supervision of play and learning
Indicator 25.7.1	0.845	Supervision of play and learning
Indicator 25.7.2	0.782	Supervision of play and learning
Indicator 25.7.3	0.733	Supervision of play and learning
Indicator 26.7.2	0.441	Peer interaction
Indicator 27.5.1	0.864	Staff-child interaction
Indicator 27.7.1	0.705	Staff-child interaction
Indicator 27.7.2	0.785	Staff-child interaction
Indicator 28.5.1	0.791	Discipline
Indicator 28.5.2	0.747	Discipline
Indicator 28.5.3	0.735	Discipline
Indicator 28.5.4	0.694	Discipline
Indicator 28.7.1	0.418	Discipline
Indicator 28.7.2	0.563	Discipline
Indicator 29.5.1	0.367	Schedule
Indicator 29.7.1	0.699	Schedule
Indicator 29.7.2	0.593	Schedule
Indicator 30.5.1	0.493	Free play
Indicator 30.5.2	0.793	Free play
Indicator 30.7.1	0.769	Free play
Indicator 30.7.2	0.570	Free play
Indicator 31.5.3	0.388	Group play activities
Indicator 31.7.2	0.645	Group play activities

Estimates reported are standardized factor loadings. Variable names refer to item number, level of quality and question number

* Copyright regulations prevent providing further detail of the indictors

Table 5 Table of factor loadings for ECERS-R cognitive stimulation factor

Indicator	Estimate	Item*
Indicator 4.5.1	0.538	Room arrangement for play
Indicator 4.5.3	0.335	Room arrangement for play
Indicator 4.7.1	0.426	Room arrangement for play
Indicator 4.7.2	0.465	Room arrangement for play
Indicator 10.7.1	0.486	Meals/ snacks
Indicator 10.7.3	0.794	Meals/ snacks
Indicator 11.7.2	0.375	Nap/ rest
Indicator 13.7.1	0.657	Health practices
Indicator 15.5.1	0.324	Books and pictures
Indicator 16.5.1	0.823	Encouraging children to communicate
Indicator 16.7.1	0.719	Encouraging children to communicate
Indicator 17.3.2	0.644	Using language to develop reasoning skills
Indicator 17.5.1	0.417	Using language to develop reasoning skills
Indicator 17.5.2	0.491	Using language to develop reasoning skills
Indicator 17.7.1	0.687	Using language to develop reasoning skills
Indicator 17.7.2	0.747	Using language to develop reasoning skills
Indicator 18.5.1	0.798	Informal use of language
Indicator 18.5.3	0.803	Informal use of language
Indicator 18.5.4	0.860	Informal use of language
Indicator 18.7.1	0.890	Informal use of language
Indicator 18.7.2	0.799	Informal use of language
Indicator 20.5.1	0.365	Art
Indicator 20.5.2	0.406	Art
Indicator 21.7.1	0.457	Music/ movement
Indicator 21.7.3	0.398	Music/ movement
Indicator 22.7.1	0.492	Blocks
Indicator 24.7.4	0.565	Dramatic play
Indicator 25.5.4	0.483	Nature/ science
Indicator 25.7.1	0.369	Nature/ science
Indicator 26.5.1	0.360	Math/ number
Indicator 26.5.4	0.577	Math/ number
Indicator 28.5.2	0.368	Promoting acceptance of diversity
Indicator 29.5.3	0.514	Supervision of gross motor activities
Indicator 29.7.1	0.552	Supervision of gross motor activities
Indicator 29.7.2	0.708	Supervision of gross motor activities
Indicator 30.5.1	0.642	General supervision of children
Indicator 30.5.2	0.772	General supervision of children
Indicator 30.5.4	0.767	General supervision of children
Indicator 30.7.1	0.872	General supervision of children
Indicator 30.7.2	0.877	General supervision of children
Indicator 31.5.3	0.690	Discipline
Indicator 31.7.1	0.752	Discipline
Indicator 31.7.2	0.504	Discipline
Indicator 32.7.2	0.791	Staff-child interactions
Indicator 33.5.2	0.834	Interactions among children
Indicator 33.7.1	0.652	Interactions among children
Indicator 33.7.2	0.457	Interactions among children
Indicator 34.7.1	0.513	Schedule
Indicator 34.7.2	0.641	Schedule
Indicator 35.5.2	0.844	Free play
Indicator 35.7.1	0.836	Free play
Indicator 35.7.2	0.419	Free play

Table 5 (continued)

Indicator	Estimate	Item*
Indicator 36.5.3	0.561	Group time
Indicator 36.7.2	0.752	Group time

Estimates reported are standardized factor loadings. Variable names refer to item number, level of quality and question number

* Copyright regulations prevent providing further detail of the indictors

Acknowledgments
This work was supported by the The Research Council of Norway (BePro-220570).

Author contributions
Joakim Evensen Hansen: conception, data collection, analysis and interpretation of results, manuscript preparation. Erik Eliassen: conception, analysis and interpretation of results, manuscript preparation. Elisabeth Bjørnestad: conception, data collection, interpretation of results, manuscript preparation. Edward Melhuish: conception, interpretation of results, manuscript preparation. Jacqueline Barnes: conception, interpretation of results, manuscript preparation.

Funding
This work was supported by the The Research Council of Norway (BePro-220570).

Data availability
The data sets generated and/or analyzed during the current study are part of an ongoing research project and are not publicly available due to national data protections. Contact the corresponding author for more information.

Declarations

Competing interests
None of the authors have a competing interests to declare.

Author details
[1]Norwegian Reading Centre, University of Stavanger, Stavanger 4016, Norway
[2]Department of Child Health and Development, Norwegian Institute of Public Health, Oslo, Norway
[3]Department of Early Childhood Education, Oslo Metropolitan University, Oslo, Norway
[4]Department of Education, University of Oxford, Oxford, UK
[5]Department of Psychological Sciences, Birkbeck, University of London, London, UK
[6]University of Oxford, Oxford, UK
[7]Beijing City University, Beijing, China
[8]Beijing Normal University, Beijing, China
[9]Zhejiang University, Hangzhou, China

Received: 8 November 2024 / Accepted: 15 April 2026
Published online: 19 June 2026

References

Alpys, A., & Hernández-Torrano, D. (2026). Scoping Review of Early Childhood Environment Rating Scale (ECERS): Assessing and Validating Quality in Early Childhood Education Settings. *Early Education and Development*, *37*(3), 644–664. https://doi.org/10.1080/10409289.2025.2585294

Barnes, J., & Melhuish, E. C. (2017). Amount and timing of group-based childcare from birth and cognitive development at 51 months: A UK study. *International Journal of Behavioral Development*, *41*, 360–370. https://doi.org/10.1177/0165025416635756

Bjørnestad, E., Baustad, A. G., Alvestad, M. (2019). To what extend does the ITERS-R address pedagogical quality as described in the Norwegian Framework plan?. Phillipson, Sivanes; Garvis, Susanne (Red.). Teachers' and Families' Perspectives in Early Childhood Education and Care Early Childhood Education in the 21st Century Vol. II, 1st Edition. p. 155-171. Routledge. https://doi.org/10.4324/9780203730546

Beaujean, A. A. (2014). *Latent variable modeling using R: A step-by-step guide*. Routledge.

Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., & Owen, M. T. and (2007). Are There Long-Term Effects of Early Child Care? *Child Development*, *78*, 681–701. https://doi.org/10.1111/j.1467-8624.2007.01021.x

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon, R. M. Lerner, & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (6th ed., pp. 793–828). John Wiley & Sons Inc.

Brunsek, A., Perlman, M., Falenchuk, O., McMullen, E., Fletcher, B., et al. (2017). The relationship between the Early Childhood Environment Rating Scale and its revised form and child outcomes: A systematic review and meta-analysis. *PLOS ONE*, *12*(6), e0178512. https://doi.org/10.1371/journal.pone.0178512

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., Riggins, R., Zeisel, S., Neebe, E., & Bryant, D. (2000). Relating Quality of Center-Based Child Care to Early Cognitive and Language Development Longitudinally. *Child Development*, *71*, 339–357.

Burchinal, M., Garber, K., Foster, T. J., Bratsch-Hines, M. E., Franco, X., & Peisner-Feinberg, E. (2021). Relating early care and education quality to preschool outcomes: The same or different models for different outcomes? *Early Childhood Research Quarterly*, *55*, 35–51.

Cadima, J., Nata, G., Barros, S., Coelho, V., & Barata, C. (2020). Literature review on early childhood education and care for children under the age of 3. *OECD Education Working Papers*, No. 243, OECD Publishing, Paris. https://doi.org/10.1787/a9cef727-en

Cassidy, D., Hestenes, L. L., Hegde, A. V., Hestenes, S. L., & Mims, S. U. (2005). Measurement of quality in preschool childcare classrooms: An exploratory and confirmatory factor analysis of the early childhood environment rating scale-revised. *Early Childhood Research Quarterly*, *20*, 345–360.

Clifford, R., Yazejian, N., Cryer, D., & Harms, T. (2020). Forty years of measuring quality with the Environment Rating Scales. *Early Childhood Research Quarterly*, *51*, 164–166.

Dearing, E., McCartney, K., & Taylor, B. A. (2009). Does higher quality early child care promote low-income children's math and reading achievement in middle childhood? *Child development*, *80*(5), 1329–1349. https://doi.org/10.1111/j.1467-8624.2009.01336.x

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., Cai, K., Clifford, R. M., Ebanks, C., Griffin, J. A., Henry, G. T., Howes, C., Iriondo-Perez, J., Jeon, H. J., Mashburn, A. J., Peisner-Feinberg, E., Pianta, R. C., Vandergrift, N., & Zill, N. (2007). Teachers' education, classroom quality, and young children's academic skills: results from seven studies of preschool programs. *Child development*, *78*, 558–580. https://doi.org/10.1111/j.1467-8624.2007.01014.x

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.

Eliassen, E., Zachrisson, H. D., & Melhuish, E. (2018). Is cognitive development at three years of age associated with ECEC quality in Norway? *European Early Childhood Education Research Journal*, *26*, 110–197.

Eliassen, E., Brandlistuen, R. E., & Wang, M. V. (2024). The effect of ECEC process quality on school performance and the mediating role of early social skills. *European Early Childhood Education Research Journal*, *32*(2), 325–338. https://doi.org/10.1080/1350293X.2023.2254531

Elliot, C. D., & Smith, P. (2011). British ability scales – third edition (BAS-3). London: GL Assessment Limited.

Elnes, M., Hansen, J. E., Lervåg, A., Hatlevik, O. E., & Reikerås, E. K. L. (2024). Verbal and non-verbal skills in early childhood: dimensionality, developmental trajectories, and gender differences. *Frontiers in psychology*, *15*, 1330334. https://doi.org/10.3389/fpsyg.2024.1330334

Elnes, M., Hatlevik, O. E., & Reikerås, E. (2026). ECEC quality matters for boys: gender differences in the long-term associations between ECEC quality and academic outcomes. *European Early Childhood Education Research Journal*, *34*(1), 90–105. https://doi.org/10.1080/1350293X.2025.2477497

Fujimoto, K. A., Gordon, R. A., Peng, F., & Hofer, K. G. (2018). Examining the Category Functioning of the ECERS-R Across Eight Data Sets. *AERA Open*. https://doi.org/10.1177/2332858418758299

Garvis, S., Sheridan, S., Williams, P., & Mellgren, E. (2017). Cultural considerations of ECERS-3 in Sweden: a reflection on adaption. *Early Child Development and Care*, *188*(5), 584–593. https://doi.org/10.1080/03004430.2017.1377192

Gordon, R. A., Fujimoto, K., Kaestner, R., Korenman, S., & Abner, K. (2013). An assessment of the validity of the ECERS-R with implications for measures of child care quality and relations to child development. *Developmental psychology*, *49*, 146–160. https://doi.org/10.1037/a0027899

Gordon, R., Hofer, K., Fujimoto, K. A., Risk, N. M., Kaestner, R., & Korenman, S. (2015). Identifying High Quality Preschool Programs: New Evidence on the Validity of the ECERS-R in Relation to School Readiness Goals. *Early Education and Development*, *26*, 1086–1110. https://doi.org/10.1080/10409289.2015.1036348

Hansen, J. E., & Broekhuizen, M. L. (2020). Quality of the Language-Learning Environment and Vocabulary Development in Early Childhood. *Scandinavian Journal of Educational Research*, *65*, 302–317. https://doi.org/10.1080/00313831.2019.1705894

Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early childhood environment rating scale – revised edition*. Teachers College, Columbia University.

Harms, T., Cryer, D., & Clifford, R. M. (2006). *Infant/toddler environment rating scale – revised edition*. Teachers College.

Harms, T., Clifford, D. M., & Cryer, D. (2015). *Early childhood environment rating scale third edition*. Teachers College.

Hestenes, L. L., Cassidy, D., Hegde, A. V., & Lower, J. K. (2007). Quality in Inclusive and Non-inclusive Infant and Toddler Classrooms. *Journal of Research in Childhood Education*, *22*, 69–84.

Howard, S. J., Lewis, K. L., Walter, E., et al. (2024). Measuring the Quality of Adult–Child Interactions in the Context of ECEC: a Systematic Review on the Relationship with Developmental and Educational Outcomes. *Educ Psychol Rev*, *36*, 6. https://doi.org/10.1007/s10648-023-09832-3

Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., Ruzek, E. A., & Howes, C. (2013). Preschool center quality and school readiness: quality effects and variation by demographic and child characteristics. *Child development*, *84*, 1171–1190. https://doi.org/10.1111/cdev.12048

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 10155–10162. https://doi.org/10.1073/pnas.0600888103

Li, W., Farkas, G., Duncan, G. J., Burchinal, M. R., & Vandell, D. L. (2013). Timing of high-quality child care and cognitive, language, and preacademic development. *Developmental psychology*, *49*, 1440–1451. https://doi.org/10.1037/a0030613

Løkken, I. M., Broekhuizen, M. L., Barnes, J., Moser, T., & Bjørnestad, E. (2018). Interaction-quality and children's social competence in Norwegian ECEC. *Varhaiskasvatuksen Tiedelehti*, *7*, 338–361. https://doi.org/10.1186/s40723-018-0048-z

Mariano, M., Caetano, S., Silva, A. R., Surkan, P., Martins, S., & Cogo-Moreira, H. (2019). Psychometric Properties of the ECERS-R Among an Epidemiological Sample of Preschools. *Early Education and Development*, *30*, 511–521.

Mayer, D., & Beckh, K. (2016). Examining the validity of the ECERS–R: Results from the German National Study of Child Care in Early Childhood. *Early Childhood Research Quarterly*, *36*, 415–426. https://doi.org/10.1016/j.ecresq.2016.01.001

Melhuish, E., & Barnes, J. (2021). Preschool programs for the general population. Melhuish E, topic ed. In R. E. Tremblay, M. Boivin, & de R. Peters V., (Eds.), *Encyclopedia on Early Childhood Development [online]*. Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development,1-6

Melhuish, E. C., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., & Leseman, P. (2015). *A review of research on the effects of early childhood Education and Care (ECEC) upon child development*. and CARE project. Curriculum Quality Analysis and Impact Review of European Early Childhood Educationhttp://ecec-care.org/resources/publications/

Neitzel, J., Early, D., Sideris, J., LaForrett, D., Abel, M. B., Soli, M., & Kortright Wood, J. (2019). A comparative analysis of the Early Childhood Environment Rating Scale–Revised and Early Childhood Environment Rating Scale, Third Edition. *Journal of Early Childhood Research*, *17*(4), 408–422. https://doi.org/10.1177/1476718X19873015

NICHD Early Child Care Research Network (2003). Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development. *Child Development*, *74*, 1454–1475.

Nye, C. D., & Drasgow, F. (2011). Assessing Goodness of Fit: Simple Rules of Thumb Simply Do Not Work. *Organizational Research Methods*, *14*, 548–570. https://doi.org/10.1177/1094428110368562

Perlman, M., Zellman, G., & Le, V. (2004). Examining the psychometric properties of the Early Childhood Environment Rating Scale-Revised (ECERS-R). *Early Childhood Research Quarterly*, *19*, 398–412.

Rademacher, A., Bäker, N., von Düring, U., et al. (2025). The effects of early childhood education and care quality on academic competences in early primary school: A systematic review and meta-analysis. *European Journal Of Psychology Of Education*, *40*, 110. https://doi.org/10.1007/s10212-025-01005-w

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1–36.

Slot, P., Leseman, P., 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.

Sluiter, R. M. V., Fukkink, R. G., & Fekkes, M. (2025). The Impact of Process Quality in Early Childhood Education and Care on Socio-Emotional Development: A Meta-Analysis of Longitudinal Studies. International journal of environmental research and public health, 22(5), 775. https://doi.org/10.3390/ijerph22050775

Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K., & Totsika, V. (2006). Capturing Quality in Early Childhood through Environmental Rating Scales. *Early Childhood Research Quarterly*, *21*, 76–92.

Turoff, M., & Linstone, H. A. (2002). *The Delphi method-techniques and applications*. Addison-Wesley.

Ulferts, H., Wolf, K. M., & Anders, Y. (2019). Impact of Process Quality in Early Childhood Education and Care on Academic Outcomes: Longitudinal Meta-Analysis. *Child Development*, *90*, 1474–1489. https://doi.org/10.1111/cdev.13296

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., Vandergrift, N., & NICHD Early Child Care Research Network. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child development*, *81*, 737–756. https://doi.org/10.1111/j.1467-8624.2010.01431.x

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.

Publisher's note