Australian Education Research Organisation

Technical report

Linking quality and child development in early childhood education and care

June 2024

The Australian Education Research Organisation (AERO) is Australia's national education evidence body, working to achieve excellence and equity in educational outcomes for all children and young people.

Acknowledgement

AERO's work is made possible by the joint funding it receives from Commonwealth, state and territory governments.

Part of this study was conducted at The University of Queensland (UQ) using data from the Effective Early Educational Experiences (E4Kids) study under licence from the University of Melbourne. E4Kids was funded by the Australian Research Council (ARC) Linkage Projects scheme (LP0990200), in collaboration with the Victorian Government Department of Education and Early Childhood Development, and the Queensland Government Department of Education and Training. The authors thank the early childhood education and care services, directors, teachers/staff, children and their families who participated in E4Kids.

This research was supported (partially or in full) by the ARC's Centre of Excellence for Children and Families over the Life Course (Project ID CE200100025). Sally Staton is supported by an ARC Discovery Early Career Fellowship (DE230100687), and Karen Thorpe by an ARC Laureate Fellowship (FL220100137).

The research undertaken within this project was approved by the UQ Human Research Ethics Committee (ID 2023/HE001018).

Acknowledgement of Country

AERO acknowledges the Traditional Custodians of the lands, waterways, skies, islands and sea Country across Australia. We pay our deepest respects to First Nations cultures and Elders past and present. We endeavour to continually value and learn from First Nations knowledges and educational practices.

Disclaimer

This report uses data from the Australian Early Development Census (AEDC). The AEDC is funded by the Australian Government Department of Education. The findings and views reported are those of the author and should not be attributed to the department or the Australian Government.

The results of these studies are based, in part, on data supplied to the Australian Bureau of Statistics (ABS) under the *Taxation Administration Act 1953*, *A New Tax System (Australian Business Number) Act 1999*, *Australian Border Force Act 2015*, *Social Security (Administration) Act 1999*, *A New Tax System (Family Assistance) (Administration) Act 1999*, *Paid Parental Leave Act 2010* and the Student Assistance Act 1973. Such data may only be used for the purpose of administering the *Census and Statistics Act 1905* or performance of functions of the ABS as set out in section 6 of the *Australian Bureau of Statistics Act 1975* (ABS Act). No individual information collected under the Census and Statistics Act is provided back to custodians for administrative or regulatory purposes. Any discussion of data limitations or weaknesses is in the context of using the data for statistical purposes and is not related to the ability of the data to support the Australian Taxation Office, Australian Business Register, Department of Social Services and/or Department of Home Affairs' core operational requirements.

Legislative requirements to ensure privacy and secrecy of these data have been followed. For access to the Person Level Integrated Data Asset (PLIDA; formerly MADIP) and/or Business Longitudinal Analysis Data Environment data under section 16A of the ABS Act or enabled by section 15 of the *Census and Statistics (Information Release and Access) Determination 2018*, source data are de-identified and so data about specific individuals has not been viewed in conducting this analysis. In accordance with the Census and Statistics Act, results have been treated where necessary to ensure that they are not likely to enable identification of a particular person or organisation.

Authors

Dr Peter Rankin, Associate Professor Sally Staton, Alicia Jones, Dr Azhar Hussain Potia, Dr Sandy Houen, Bridget Healey and ARC Laureate Professor Karen Thorpe.

Copyright

How to cite

Rankin, P., Staton, S., Jones, A., Potia, A. H., Houen, S., Healey, B., & Thorpe, K. (2024). *Linking quality and child development in early childhood education and care: Technical report*. Australian Education Research Organisation. https://www.edresearch.edu.au/research/technical-papers/linking-quality-and-child-development-early-childhood-education-and-care

Publication details

ISBN 978-1-923066-32-8 (online)
Cover image: AERO

Contents

Figures

Tables

Appendix tables

Abbreviations

Abbreviation	Full term
ABS	Australian Bureau of Statistics
AEDC	Australian Early Development Census
AERO	Australian Education Research Organisation
AIC	Akaike Information Criterion
ARID	Address Register ID
BIC	Bayesian Information Criterion
CBDC	Centre-based day care
CCMS	Child Care Management System
CLASS	Classroom Assessment Scoring System
E4Kids	Effective Early Educational Experiences
ECEC	Early childhood education and care
ECERS-R	Early Childhood Environment Rating Scale – Revised
ESL	English as a second language
FDC	Family day care
FFY	First Five Years
MADIP	Multi-Agency Data Integration Project
MB	Mesh Block
MBS	Medicare Benefits Schedule
NQS	National Quality Standard
OTE	Language background other than English
PLIDA	Person Level Integrated Data Asset
UQ	The University of Queensland
WJ III	Woodcock-Johnson III

Executive summary

Background to study

Internationally, investment in early childhood education and care (ECEC) has been a policy strategy to support each child's ongoing development, learning and wellbeing (*Equity and Child Rights*) and to deliver the skills and knowledge required for a thriving society and economy (*Human Capital, Brain Capital*). This investment is grounded in a substantial body of evidence from neuroscience, developmental science and economics. In Australia, this investment exceeds $14 billion per annum and continues to increase as each jurisdiction pledges further investments and the Australian Government rolls out its Early Years Strategy. Assessing the efficiency of 'how' these investments are implemented in practice, and of the methods of ongoing monitoring, are important steps in delivering effective early experiences for children that optimise early learning and improve life chances for all Australian children.

The Australian Education Research Organisation (AERO) aims to lay the groundwork for a stronger, smarter ECEC system that maximises the potential of education data for improving policy and practice, towards excellence and equity in learning and development for all Australian children. In 2023, AERO partnered with the Queensland Brain Institute, Child Development and Early Education Group at The University of Queensland (UQ), leveraging their expertise working with the Person Level Integrated Data Asset (PLIDA; formerly Multi-Agency Data Integration Project [MADIP]) database and the Effective Early Educational Experiences (E4Kids) longitudinal study. The Queensland Brain Institute prepared this technical report on the study findings.

Purpose of the study

This study responds to policy demand to better understand the effects of government investments in ECEC, and asked:

> How does the quality of ECEC provision across the diversity of formal ECEC programs predict child development outcomes as Australian children enter school?

We also considered development across time to ask the question:

> What is the 'value add' of ECEC service quality to children's developmental outcomes?

We applied these questions to all Australian children and to specific population subgroups who experience various forms of disadvantage on the basis of:

- remoteness – living in areas outside major cities of Australia
- income and welfare – living in circumstances of social and economic disadvantage
- individual variation – child language background and proficiency.

Design of the study

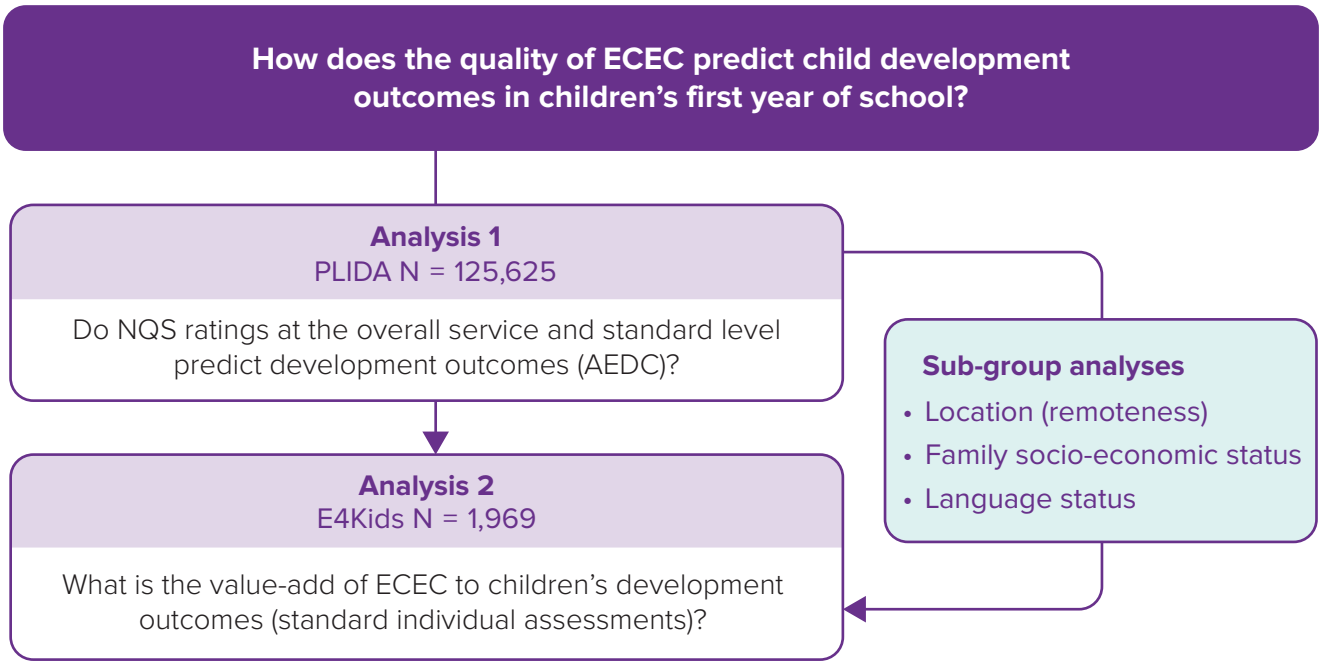
The study employed 2 pre-existing population datasets to interrogate the effect of ECEC quality: PLIDA and E4Kids.

Person Level Integrated Data Asset (PLIDA) is an Australian Government dataset integrating data across a range of Australian, state and territory government agencies overseen by the Australian Bureau of Statistics (ABS). PLIDA allows the creation of unique linkages containing data to answer important research. Data from the PLIDA The First Five Years (FFY) custom linkage project is used in this report. Key data incorporated in FFY included participation in ECEC from the childcare management system, ECEC quality from the National Quality Standard (NQS), child outcomes from the Australian Early Development Census (AEDC) and a range of social and demographic information. The 2018 AEDC dataset was the focus outcome of this report. It provides an Australian total population covering 96% of Australian children in their first year of full-time schooling, when the AEDC measures child development. Quality of ECEC is measured through the NQS ratings for family day care (FDC) and centre-based day care (CBDC; previously long day care) services attended with data available on hours of participation via childcare subsidy records. The limitations of the PLIDA FFY project are that ECEC data pertain solely to FDC and CBDC, as preschool data is held by state and territory jurisdictions. The data also do not provide a baseline assessment of each child's development prior to ECEC entry, limiting the ability to infer whether differences in outcomes in their first year of full-time schooling are due to the quality of ECEC or selection into a particular type and quality of ECEC service. The advantage is the large sample size and availability of a large range of variables that facilitate statistical adjustment for confounding explanatory mechanisms.

Effective Early Educational Experiences (E4Kids) was a study of 2,600 children from Queensland and Victoria, tracking from the preschool year in 2010 (age 4) through to 2014 (age 8). Detailed standard assessments of children's development and achievements were undertaken every year, providing a baseline in ECEC and outcomes in their first year of full-time schooling. Assessment of ECEC quality was through standard measures, the Classroom Assessment Scoring System (CLASS) and the Early Childhood Environment Rating Scale – Revised (ECERS-R). Parent reports provided a range of demographic variables that allowed for statistical adjustment for confounding explanatory mechanisms. The limitation of this study was the smaller sample size. The advantages were the detailed assessment of ECEC quality and child outcomes, as well as the availability of baseline developmental data to allow a 'value add' model that examines the unique contribution of ECEC to children's development.

In addressing our research questions, we first analysed the large PLIDA FFY dataset. We then analysed the E4Kids data to triangulate findings in a richer dataset and explored 'value add' effects. A range of analyses within each dataset were undertaken to explore the focus questions. The design is summarised in Figure 1.

Figure 1: Summary of design



Study outcomes: PLIDA FFY analyses

Box 1: The Australian Early Development Census

AEDC data is collected using the Australian version of the Early Development Instrument in a child's first year of full-time school. Based on their knowledge and observations of children in their class, teachers respond to approximately 100 questions across the 5 domains of the AEDC. These teachers' responses are combined into a score (range 0 to 10) for each child on the 5 AEDC domains. Using benchmarks calculated in 2009, children are then classified to be 'developmentally vulnerable' (score is in 0 to 10th age-adjusted percentile of 2009 AEDC cohort domain scores), 'developmentally at risk' (11th to 25th percentile) or 'developmentally on track' (greater than 26th percentile) for each of the following domains.

- **Physical Health and Wellbeing:** Children's physical readiness for the school day, physical independence, and gross and fine motor skills
- **Social Competence:** Children's overall social competence, responsibility and respect, approach to learning and readiness to explore new things
- **Emotional Maturity:** Children's pro-social and helping behaviours, anxious and fearful behaviour, aggressive behaviour, and hyperactivity and inattention
- **Language and Cognitive Skills (school-based):** Children's basic literacy; interest in literacy, numeracy and memory; advanced literacy; and basic numeracy
- **Communication Skills and General Knowledge:** Children's communication skills and general knowledge, based on broad developmental competencies and skills.

Analyses examined ECEC quality indexed by the NQS (2012 version) ratings as a predictor of AEDC (2018) 'developmentally vulnerable' and 'developmentally at-risk' outcome domain indicators in children's first year of full-time schooling (see Box 1 for a description of the AEDC). The quality of ECEC focused on the NQS ratings experienced 2 years before school (2016; complete case n = 89,988, imputed n = 125,625) and in the year before school (2017; complete case n = 82,357, imputed n = 116,356). Quality was assessed at the overall NQS rating level of each service (Significant Improvement Required through to Exceeding NQS/Excellent), as well as 7 NQS quality areas and a latent class analysis combining the 18 quality standards. Stratification effects were examined for service remoteness, income and income support of the child's caregivers, and language background and English-language proficiency of the child.

Key findings

1. **NQS overall ratings were consistently associated with AEDC domains.** Children in services rated Exceeding NQS or Excellent had lower rates of developmental vulnerability compared to children in services rated Meeting NQS, Working Towards NQS or Significant Improvement Required.
2. **Quality Areas 1 (Educational program and practice), 3 (Physical environment) and 5 (Relationships with children) had the greatest and most consistent association with AEDC domains.** Quality Area 6 (Collaborative partnerships with families and communities) was similarly associated but to a lesser extent. Quality Areas 2 (Children's health and safety), 4 (Staffing arrangements) and 7 (Leadership and service management) had some, but fewer, consistent associations with AEDC domains.
3. **Ratings of Exceeding NQS were consistently associated with lower rates of children assessed as developmentally vulnerable and developmentally at risk across the stratifications of remoteness, income and income support of the child's caregivers, and language background and ability of children.** However, the number of effects detected were related directly to sample size and were occasionally sporadic, in line with statistical principles.
4. **Latent class analyses of standards distinguished 6 types of ECEC service quality.** These map to the broader classes of Working Towards NQS, Meeting NQS or Exceeding NQS with variation within:
   a. Exceeding all Quality Areas (18.4% of all services)
   b. Exceeding in Quality Areas 1 and 5 (12.5% of all services)
   c. Exceeding in Quality Areas 6 and 7 (14.5% of all services)
   d. Meeting all Quality Areas (26.6% of all services)
   e. Working Towards, more Standards Meeting (19.6% of all services)
   f. Working Towards, more Standards Working Towards (8.3% of all services).
5. **Consistent with the analysis of overall quality and quality areas, children in services rated Working Towards NQS or Meeting NQS had higher rates of developmental vulnerability compared to those in services rated as Exceeding NQS.** Children in services in latent classes (iv, v and vi) had higher rates of developmental vulnerability for each AEDC domain when compared to the services typically exceeding all standards (i). Children in services that exceeded in Quality Area 6 (Collaborative partnerships with families and communities) and Quality Area 7 (Leadership and service management) also had higher rates of developmental vulnerability on the Communication Skills and General Knowledge domain, compared to children in services more likely to exceed all standards.

6. **The national assessment and rating system serves as a predictor of child outcomes, but not uniformly so.** Some quality areas of assessment were less consistently associated with developmental vulnerability across AEDC domains.

The results show that ECEC quality, as assessed by the NQS (2012 version), is associated with children's developmental vulnerability at entry to school, as assessed by the AEDC. In the absence of baseline developmental data, these results can be interpreted in 3 ways:

1. **Higher quality ECEC reduces the risk of children's developmental vulnerability in the first year of full-time schooling.** There is research evidence from Australia and internationally that concurs with this assertion, but the current study cannot infer causality.
2. **There is a selection effect in which children who are more developmentally vulnerable have less access to high-quality ECEC**, and this is not reduced by attendance at the ECEC programs they have access to. There is evidence from Australia that such selection effects occur, including selection by socio-demographic factors from analyses of the PLIDA FFY data. There is also evidence that children who are more developmentally vulnerable can and do access high-quality ECEC services – often through targeted models. Further examination of access to high-quality services, dosage and provider models is warranted to direct policies that enable children who commence life as developmentally vulnerable to access the highest quality ECEC services. The results also have implications for quality improvement investments.
3. **The results reflect a combination of ECEC effects and selection effects.** This is a likely scenario. It has been noted in prior research that class composition can affect outcomes, with high concentration of complex populations placing higher demand on educators and reducing interactional quality.

Study outcomes: E4Kids analyses

In E4Kids, the measures of quality and child outcomes were more detailed. Quality was measured with CLASS and ECERS-R (Third edition – activities, routines and furnishings). CLASS focuses on interactional (process) quality, while the subcategories of ECERS-R used in the study focused on program content and physical resources (structural quality). The child outcomes were measured using the Woodcock-Johnson III (WJ III) Tests of Cognitive Abilities and Tests of Achievement. Extending the analytic strategy applied in the PLIDA dataset, the quality of ECEC was modelled to explain changes in child outcomes between 2 years to assess the 'value add' from participating in high quality ECEC for 1,969 children. Latent class analyses were also used to derive and evaluate quality types across CLASS and ECERS-R.

Key findings

1. **Features of process quality and structural quality were associated with gains in cognitive development, identifying a value add of ECEC.** Children in services with higher Classroom Organization had greater improvement in WJ III outcomes (visual matching). Participation in services with higher ECERS-R activities (observed curriculum content) was also associated with greater improvement on the WJ III outcome understanding directions (listening ability, language development), and ECERS-R furnishings was associated with improved visual matching.
2. **Some aspects of structural quality may not benefit child outcomes measured by WJ III.** Children who participated in services with higher scores on ECERS-R routines (e.g., toileting, meals and sleep/rest practices) had lower gains on visual matching and understanding directions.

3. **Latent classes of CLASS and ECERS-R mapped to service type.** The latent class analysis delivered a 4-class solution. These quality profiles mapped closely to different types of provision (i.e., preschool, CBDC and FDC). The results likely reflect that different structural features (e.g., staffing and hours of operation) assessed through ECERS-R enable process quality, but may also map to selection effects into program type.

The results show that participation in higher-quality ECEC services, as assessed by CLASS and ECERS-R, was associated with improved gains in cognitive ability and achievement. This result is strengthened by adjusting for baseline outcomes and process and structural features of ECEC quality. Future research could include multiple observations of quality to further reduce confounding or selection, and could examine the role of ECEC routines in child development in greater nuance.

General discussion

In this study we analysed 2 key Australian datasets to examine the impact of ECEC provision on Australian children's developmental outcomes as they enter school.

Our results show that:

1. **The quality of ECEC matters.** Whether using a large sample (greater than 100,000) with broader scoped measures of quality and child outcomes, or a smaller sample (greater than 2,600) with more refined measurement, the analyses converged on the conclusion that higher quality ECEC supports a child's development outcomes in their first year of full-time schooling. Effects for children who live in more remote areas, in families who are socially and economically disadvantaged, or from language backgrounds other than English also showed positive effects. Evidence from E4Kids suggests that the effects of ECEC 'value add' are not simply due to selection of more developmentally advantaged children into poorer quality programs.
2. **The type of quality measure matters.** The combined evidence from both datasets suggests items focused on structural quality are less likely to predict child outcomes than process quality items – for example, ECERS-R routines and Quality Area 2 (Children's health and safety) are weaker or negative predictors. Examination of the content of these items shows a focus on issues such as hygiene rather than interactions with children.
3. **NQS ratings are sensitive predictors of developmental outcomes.** While there is evidence from the PLIDA data of selection effects in which children who are more developmentally advantaged are more likely to access higher quality programs as assessed by the NQS, our triangulation with the E4Kids data suggests a likely 'value add' effect. There are 3 specific findings that should be noted:
   a. Exceeding NQS/Excellent distinguishes the greatest effects in delivering improved developmental outcomes for children compared to other ratings (i.e., Meeting NQS, Working Towards NQS and Significant Improvement Required).
   b. Quality Areas 1, 3 and 5 are most strongly associated with improved developmental outcomes.
   c. Quality Areas 2, 4, 6 and 7 are positively associated with developmental outcomes, but less consistently.

The findings of positive effects of high quality ECEC on early development are not new. They add to the body of international literature that supports investment in ECEC to deliver the highest possible quality. Importantly, these findings confirm the value of investing in a quality assessment and rating system (NQS) and ongoing measures to improve quality of ECEC for all Australian children, particularly those experiencing life circumstances that increase the risk of develeopmental vulnerability.

The analyses presented here, and in a range of literature (e.g., Thorpe et al., 2020), also suggest staffing and routines are key places, or 'barometer events', to distinguish effective ECEC experiences. Evidence from studies of the assessment and rating process suggests observations of quality often focus on structural rather than interactional features in assessing Quality Area 2 (Children's health and safety) (Staton et al., 2020). This suggests there are opportunities for including a greater focus on process quality within areas less consistent in predicting child outcomes. Thus, this research confirms the role of the NQS in distinguishing developmental outcomes and highlights avenues for future refinement of the National Quality Framework.

Introduction to technical report

The following report details 2 sets of analyses undertaken by researchers at the Queensland Brain Institute and the Child Development and Early Education Group at The University of Queensland (UQ), in partnership with the Australian Education Research Organisation (AERO), examining the link between children's experience of early education and care quality and their developmental outcomes.

The first set of analyses indexed quality using the National Quality Standard (NQS; 2012) and developmental outcomes in the first year of full-time schooling by the Australian Early Development Census (AEDC). The analysis focused on children in the 2018 AEDC cohort and their experiences of quality in 2016 (2 years before school) and 2017 (the year before school). The NQS overall quality rating, as well as 7 component Quality Areas (Educational program and practice; Children's health and safety; Physical environment; Staffing arrangements; Relationships with children; Collaborative partnerships with families and communities; and Leadership and service management) were examined. Additionally, unique combinations of the 18 quality standards that make up the 7 quality areas were explored using latent class analysis. Finally, effects of quality were examined for several sub-population stratifications of interest, including children in outer-regional and remote areas, children who speak a language other than English, and children experiencing lower relative income. These analyses were informed via inverse probability weighting and multinomial regression to adjust for confounding from a range of child, family and socio-demographic characteristics that are likely to be associated with developmental outcomes and experiences of early childhood education and care (ECEC) quality.

The second set of analyses examined quality via observational measures of process (Classroom Assessment Scoring System; CLASS) and structural quality (Early Childhood Environment Rating Scale – Revised; ECERS-R) and longitudinal cognitive ability and achievement outcomes using the E4Kids study. To corroborate the first set of analyses, latent class analysis was undertaken using CLASS and ECERS-R to identify if expanded profiles of quality emerge when combining multiple quality indices. Based on these results, longitudinal models of CLASS and ECERS-R predicting gains in cognitive outcomes were fit via multi-level regression. These models evaluated the added value that quality ECEC experiences provided to the cognitive development of children.

Methods: PLIDA FFY

Data source: PLIDA FFY

Sample: NQS and AEDC domains

This sample was a custom extract of the Person Level Integrated Data Asset (PLIDA, formerly MADIP) dataset created for The First Five Years (FFY): What Makes a Difference? (Australian Bureau of Statistics [ABS], 2019) project. The PLIDA links and integrates administrative, census and survey data across Australian Government departmental systems (ABS, 2021). In the current project, it includes detailed microdata on developmental outcomes, education, government payments, health, income and taxation, employment, population demographics and ECEC over time (ABS, 2021). We used the following collections in the PLIDA customised extract: the 2016 Census of Population and Housing (2016 Census); the 2018 Australian Early Development Census (2018 AEDC); Payment Summary, Personal Income Tax, and Social Security and Related Information; Child Care Management System (CCMS) and NQS for information on ECEC participation and service quality; the Data Exchange for additional information on First Nations status; and Medicare Benefits Schedule and Pharmaceutical Benefits Scheme for information on caregiver health service use.

The specific analytical sample comprised children in the 2018 AEDC cohort and their caregivers. Caregivers were those who are indicated as carers for the child in the CCMS between 2014 to 2015 or in the 2016 Census and the relationship started in, or before, 2016. The scoping file identifies these carers as 'PAR', although this does not imply biological parent. There were several inclusion criteria, specified in Table 1. Specifically, from 308,913 children in the 2018 AEDC, the final samples were: n = 125,625 children in 2016, and n = 116,356 in 2017 with some missing information on covariates (imputed sample); and n = 89,988 children in 2016, and n = 82,357 in 2017 with complete information on all variables (with slight variations for each outcome). The main factor that reduced sample size was the availability of linked NQS quality ratings, as shown in Table 1.

Table 1: Sample selection criteria – NQS and AEDC domains

Selection criteria	Children 2016	Children 2017
Included in AEDC 2018	308,913	308,913
Valid AEDC domains	293,910	293,910
Information on NQS quality	126,108	116,787
Information on hours of ECEC participation	126,108	116,768
Only 1 distinct observation on AEDC, NQS and ECEC participation (no multiple entries)	126,088	116,752
Had caregiver(s) with information on at least 1 covariate	125,822 (n caregivers = 221,884 of possible 956,315)	116,481 (n caregivers = 204,764 of possible 956,315)
Had information on NQS service type, year of rating and remoteness (imputed sample)	125,625 (n caregivers = 221,534)	116,356 children (n caregivers = 204,539)
Information on all covariates (complete sample)	89,988	82,357

Sample: NQS latent class and AEDC domains

The NQS latent class sample was the same as that previously described. However, the latent class analysis was fit at service level, instead of child level.

Outcomes

The outcomes modelled are 5 AEDC domains: Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Physical Health and Wellbeing, Social Competence, and Emotional Maturity. The Communication Skills and General Knowledge domain measures a child's communications skills and general knowledge based on broad developmental competencies and skills. The Language and Cognitive Skills (school-based) domain measures a child's basic literacy, advanced literacy, basic numeracy, and interest in literacy, numeracy and memory. The Social Competence domain measures a child's overall social competence, responsibility and respect, approach to learning and readiness to explore new things. The Emotional Maturity domain measures a child's prosocial and helping behaviours, anxious and fearful behaviour, aggressive behaviour, and hyperactivity and inattention. The Physical Health and Wellbeing domain measures a child's physical readiness for the school day, physical independence, and gross and fine motor skills.

The AEDC data is collected using the Australian version of the Early Development Instrument in a child's first year of full-time school. Based on their knowledge and observations of children in their class, teachers respond to approximately 100 questions across the 5 domains of the AEDC. These teacher responses are combined into a score (range 0 to 10) for each child on the 5 AEDC domains. Using benchmarks calculated in 2009, children are then classified to be 'developmentally vulnerable' (score is in 0 to 10th percentile of 2009 AEDC cohort domain scores, controlling for age variability in the first year of school), 'developmentally at risk' (11th to 25th percentile) or 'developmentally on track' (greater than 26th percentile) on each domain. We used these 3 categorical domain indicator ratings in this analysis with 'developmentally on track' as the baseline category. All analyses were carried out in accordance with the AEDC data usage guidelines, and AEDC domain indicator categories only examined for children with both valid AEDC domain scores and valid instruments. Specifically, domain scores are flagged as invalid for children who have been in class for less than 1 month, are less than 4 years old, or where teachers complete less than 75% of the items in any given domain. A valid instrument is a completed instrument for a child older than 3 years, where the child is not considered to have special needs, and with at least 4 valid domain scores. Children with special needs are not included within domain indicator categories because of the already identified substantial developmental needs of this group.

Of particular importance, the analysis of each AEDC domain included all children with a valid AEDC domain category indicator. Thus, some children did not have all 5 AEDC domains (due to missing or invalid data) and the sample size varied slightly for each domain. Appendix Table A2 details the sample size for each domain.

Treatment/key covariate: NQS and AEDC domains

Our key covariate was the assessed quality of the ECEC service each child experienced in 2016 and 2017. We investigated the NQS (2012 version) overall rating of quality and the ratings in 7 Quality Areas (1 – Educational program and practice; 2 – Children's health and safety; 3 – Physical environment; 4 – Staffing arrangements; 5 – Relationships with children; 6 – Collaborative partnerships with families and communities; and 7 – Leadership and service management). Each rating can take on 4 to 5 values, which we collapsed to Excellent or Exceeding NQS (referred to as Exceeding NQS within results), Meeting NQS, and Working Towards NQS or Significant Improvement Required (referred to as Working Towards NQS within results). For each year, we used the mode for the child's experience of quality across the year's 4 quarters. In the case of a tie, we randomly allocated which quality rating is used (e.g., first 2 quarters or last 2 quarters). For an NQS observation to be valid in this study, there must have been an overall rating and the record linked by: Address Register ID (ARID), year, quarter and service type; or Mesh Block (MB), year, quarter and service type. Links based on ARID, year and quarter or based on MB, year and quarter were considered insufficient. Appendix Table A1 outlines the NQS 2012 Quality Areas and contributing standards and elements.

Treatment/key covariate: Latent class NQS and AEDC domains

Our key covariate was the most likely latent class of ECEC quality defined at service level that each child experienced in 2016 and 2017. We investigated the 18 standards informing 7 Quality Areas (1 – Educational program and practice; 2 – Children's health and safety; 3 – Physical environment; 4 – Staffing arrangements; 5 – Relationships with children; 6 – Collaborative partnerships with families and communities; 7 – Leadership and service management). Each standard can take on 4 values, which we collapsed to Exceeding NQS, Meeting NQS, and Working Towards NQS or Significant Improvement Required (referred to as Working Towards NQS within results). Appendix Table A1 outlines the NQS Quality Areas and contributing standards and elements.

For each year (2016 and 2017), we used the mode for the child's experience of quality across the year's 4 quarters to determine the service and quality rating to include in the latent class analysis for that child. In the case of a tie, we randomly allocated which quality rating was used (e.g., first 2 quarters or last 2 quarters). For an NQS observation to be valid in this study, there must have been an overall rating and the record linked by ARID, year, quarter and service type or MB, year, quarter and service type. Links based on ARID, year and quarter and MB, year and quarter were considered insufficient.

From this child-level service rating, we examined distinct services (final approval date of rating and service approval number) for the latent class analysis such that a unique service quality rating was only included once. Thus, from the 125,625 children in the 2016 imputed sample there were 5,420 unique service ratings from 5,125 service approval numbers. Furthermore, from the 116,356 children in the 2017 imputed sample, there were 5,963 unique service ratings from the 5,474 unique service approval numbers. As all children had NQS ratings, there was not a separate latent class analysis at the service level of children with information on all other covariates.

Covariates: PLIDA FFY

A range of child, caregiver and ECEC service covariates were included in the models examining AEDC domains.

Child: Hours of participation in ECEC (2016 and 2017)

We included the number of hours children participated in any ECEC service (CBDC and FDC) in 2016, for the 2016 sample, and in 2017, for the 2017 sample. Additionally, we added information on teacher-reported preschool for the 2017 year for children with up to 600 hours of CBDC or FDC. The information is categorised into approximate 12-hour bins (typical maximum length ECEC can be open) multiplied over available weeks in each quarter, noting 600 hours is considered full-time preschool participation. Detailed categories are presented in Table 2.

Table 2: Hours of ECEC participation in CBDC and FDC for the 2016 and 2017 analyses of quality

2016 quality analysis (ECEC 2016)	2017 quality analysis (ECEC 2017)
0 hours to 300 hours	0 to 300 hours and flagged by teacher as attending preschool in AEDC
301 to 600 hours	300 to 600 hours and flagged by teacher as attending preschool in AEDC
601 to 1,200 hours	601 to 1,200 hours
1,201 to 1,800 hours	1,201 to 1,800 hours
1,801 to 2,400 hours	1,801 to 2,400 hours
2,401 to 3,000 hours	2,401 to 3,000 hours
>3,000 hours	>3,000 hours
	>0 to 600 hours flagged by teacher as **not** attending preschool in AEDC or preschool flag missing

Child: Type of ECEC quality rating came from (2016 and 2017)

An indicator of the service type the quality rating came from. Specifically, CBDC or FDC.

Child: Year ECEC quality rating came from (2016 and 2017)

An indicator of the year the NQS quality rating was undertaken. Specifically, for 2016 and 2017 respectively, the year of rating ranged from 2012 to 2016, and 2012 to 2017.

Child: Age at first entry in ECEC

An indicator of the child's age in years on the first available entry with greater than 0 hours in CBDC or FDC in the childcare management system. Age categories include less than 1 year, greater than 1 to 2, greater than 2 to 3, and greater than 3.

Child: Gender

From the AEDC; included as binary indicator of male or female.

Child: First Nations status

A binary indicator of whether the child was ever identified as Aboriginal and/or Torres Strait Islander origin based on status in half or more of 4 datasets (i.e., AEDC, CCMS, Census and data exchange).

Child: Speaks language other than English at home

From the AEDC, a binary indicator of whether the child was classified by their teachers as speaking a language other than English at home.

Child: Household size and number of children in family

An indicator was created that combined the number of persons usually resident in the dwelling for the child, based on 2016 Census, with the number of children in the child's family, based on the 2016 Census. Specifically, the indicator recoded into single adult and single/multiple children categories (combining 1 adult and 6 and 7 persons into 1 adult and 6 or 7 due to small sample numbers) or multiple adults and single/multiple children.

Child: Remoteness

An indicator of the remoteness of the child residence, as reported in the NQS, using the following 4 categories, based on the Australian Statistical Geography Standard Remoteness Structure: major cities of Australia, inner-regional Australia, outer-regional Australia and a combined category of remote Australia and very remote Australia.

Child: Tenure type of house

An indicator of the tenure type of the child's home during the 2016 Census. Several categories were collapsed, resulting in the following values: owned outright, owned with a mortgage, rented, being occupied rent-free and a combined category of being purchased under a shared equity scheme or being occupied under a life tenure scheme or other tenure type. Not applicable responses were treated as missing in this complete case analysis approach.

Caregiver: Education

An indicator of the highest level of education of the child's caregivers recorded on the 2016 Census. It took the values: postgraduate degree, undergraduate bachelor's degree, diploma, Year 12, Certificate 3 or 4, and Year 11 or less.

Caregiver: Combined disposable income percentile

Average disposable income over 2 financial years was estimated for each of the child's caregivers, using total income minus either deductions and losses and taxation and levies (preferred) or tax withheld (if insufficient information or zero income found in the preferred formula) following the approach defined in Tang et al. (2024) and then combined. Percentiles were calculated based on the available sample to distinguish 6 categories, including:

- greater than 80th to 100th percentile
- greater than 60th to less than or equal to 80th percentile
- greater than 40th to less than or equal to 60th percentile
- greater than 20th to less than or equal to 40th percentile
- greater than 2nd to less than or equal to 20th percentile
- less than or equal to 2nd percentile.

The 2nd percentile category was to capture caregivers with low taxable income, but who may have access to other sources of wealth. The estimation of disposable income did not include all sources of income as some payments were unavailable (e.g., Family Tax Benefit and Child Care Benefit [now called Child Care Subsidy] payments). Data was from the financial years 1 July 2014 to 30 June 2016 for the 2016 sample and 1 July 2015 to 30 June 2017 for the 2017 sample.

Caregiver: Income support

A binary indicator of whether any of the child's caregivers received income support over 2 financial years (July 2014 to 30 June 2016 for the 2016 sample and July 2015 to 30 June 2017 for the 2017 sample).

Caregiver: Unemployment

A binary indicator of whether any of the child's caregivers were unemployed (not employed in paid work) for both financial years, drawing on estimated personal exertion income during financial years 1 July 2014 to 30 June 2016 for the 2016 sample and 1 July 2015 to 30 June 2017 for the 2017 sample.

Caregiver: Migration

The longest migration history of the child's caregivers to Australia on the 2016 Census. Values included third-plus-generation migrant, second-generation migrant or first-generation migrant.

Caregiver: English-speaking ability

The highest level of English-speaking ability of the child's caregivers on the 2016 Census. It was a binary variable that took values of English only, or a combined category of speaks English very well, speaks English well, does not speak English well or no English at all.

Caregiver: Age when child was born

Oldest age of child's caregiver when the child was born. It consisted of 3 categories: less than or equal to 27 years of age, greater than 27 years of age and less than or equal to 35 years of age, and over 35 years of age.

Caregiver: Chronic health

A binary indicator describing whether any of the child's caregivers used health services for a chronic health condition based on Medicare Benefits Schedule (MBS) codes 6 and 13 in the years 2015 and 2016 for the 2016 sample, and 2016 and 2017 for the 2017 sample.

Caregiver: Mental health

A binary indicator describing whether any of the child's caregivers used health services for mental health based on MBS codes (1, 2, 3, 4 and 5) and Pharmaceutical Benefits Scheme codes (Rx Risk Codes 5, 8, 11, 38, 44 and 45) in the years 2015 and 2016 for the 2016 sample, and 2016 and 2017 for the 2017 sample.

Stratification analyses

Several stratification analyses were undertaken to examine whether the effects of quality varied for sub-populations of interest. These analyses treated each sub-population separately, estimating a weighting and statistical model for each group across all covariates as is recommended (Green & Stuart, 2014). The subgroups were defined as:

1. **Income and welfare status:** Combined disposable income and income support were combined to create 3 sub-populations for stratification analyses. The first group did not receive any income support and could have any range of disposable income. The second group received income support and had disposable income less than or equal to the 40th percentile. The third group received income support and had disposable income greater than the 40th percentile.
2. **Remoteness:** Remoteness of the service in which children participated was examined by estimating the effects of quality for 3 sub-populations: 1) children participating in services in major cities of Australia; 2) children participating in services in inner-regional Australia; and 3) children participating in services in outer-regional, remote or very remote Australian communities.
3. **English-speaking background and ability:** English-speaking background and ability of children was examined in 4 stratified sub-populations to test whether the associations between quality and AEDC domains differed across cohorts. The first sub-population comprised children for whom the child's teacher reported the child did not speak a language other than English. The second comprised all children whose teacher reported the child spoke a language other than English at home. The third comprised children whose teacher reported the child spoke a language other than English at home and who were also considered English as a second language (ESL), as defined by the AEDC. The fourth comprised children whose teacher reported the child spoke a language other than English at home and who were rated as 'poor/very poor' on either AEDC items B1 ('child's ability to use language effectively in English') or B2 ('child's ability to listen in English'). Thus, 3 sub-populations of English-speaking background and ability were examined. Further, as items B1 and B2 form part of the Communication Skills and General Knowledge domain, this domain was not examined for the fourth cohort defined on teacher ratings for those items.

Analytical plan: PLIDA FFY

Effect of NQS quality on AEDC domains

To estimate the effect of quality on AEDC domains, we used multinomial logistic regression with inverse probability weighting defined for both the average treatment effect and average treatment in the overlap (Li & Li, 2019). All analyses were run in R (R Core Team, 2022; version 4.2.1). Weights were defined using the WeightIt package (Greifer, 2023a; version 0.14.2) and multinomial logistic regression, and analytical models were implemented using the svyVGAM package (Lumley, 2021; version 1.2). Balance of weights and effective sample sizes across quality categories were evaluated with the cobalt package (Greifer, 2023b; version 4.5.1). Missing data on covariates, excluding outcomes and ECEC service quality, was imputed for 10 imputed datasets using multiple imputation via the mice package (van Buuren & Groothuis-Oudshoorn, 2011; version 3.16.0). Analyses were also run using complete cases with information on all variables.

Latent class analysis of NQS quality

Latent class analysis of the 18 quality standards (detailed in Appendix Table A1) was undertaken in R (R Core Team, 2022; version 4.2.1) using the poLCA package (Linzer & Lewis, 2011; version 1.6.0.1). We estimated 1 to 10 classes for the 2016 and 2017 sample. Model indices evaluated included Akaike Information Criterion (AIC; lower is better), Bayesian Information Criterion (BIC; lower is better), entropy (0–1; closer to 1 indicates better fit) and the diagonals of average posterior probabilities (values greater than 0.8 are desirable). Further, we evaluated model fit by considering whether the patterns illustrated meaningful variation across quality standards.

Effect of latent class NQS quality on AEDC domains

The most likely class for each service, and therefore children participating in that service, was exported and used as a covariate to model children's AEDC domains. The models were similar to those used for the quality areas, using a multinomial logistic regression with inverse probability weighting for both the average treatment effect and average treatment in the overlap. Models were estimated for both the imputed and complete case samples.

Results: PLIDA FFY

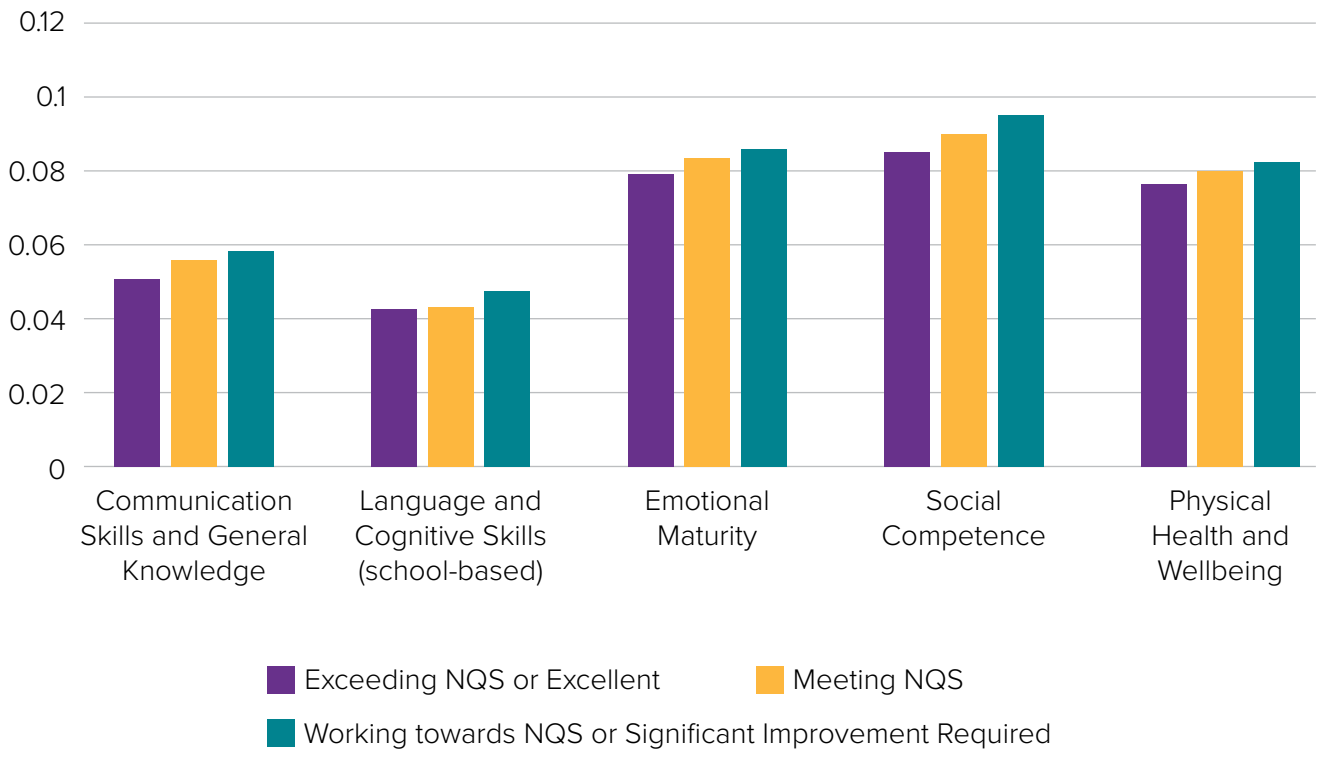
A summary of the statistical models follows.[1] Sample sizes for each domain (equivalent across quality ratings) are presented in Appendix Table A2. Descriptive statistics for sample covariates are presented in Appendix Tables A3 and A4.

Overall quality

Consistent in 2016 and 2017

Results for the overall quality rating analysis are presented in Table 3 and illustrated in Figure 2. These show that, compared to children participating in services rated Exceeding NQS, children in services rated Working Towards NQS were consistently more likely to be developmentally vulnerable for each AEDC domain. Specifically, Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Emotional Maturity, Social Competence, and Physical Health and Wellbeing. In addition, children in services rated Meeting NQS were consistently more likely be developmentally vulnerable on the Communication Skills and General Knowledge, Emotional Maturity, and Social Competence domains. Further, children in services rated Meeting NQS or Working Towards NQS were more likely to be developmentally at risk in the Communication Skills and General Knowledge domain.

Figure 2: Model implied rates of developmental vulnerability by overall quality for the 2016 imputed sample



1 Additional results are available in a series of supplementary tables. Please contact AERO for a copy.

Inconsistent across 2016 and 2017

Children in services rated Meeting NQS were more likely to be developmentally at risk in the Social Competence domain for the 2016 sample. However, this was not replicated in the 2017 sample, suggesting an inconsistent association. Children in services rated Meeting NQS or Working Towards NQS were more likely to be developmentally at risk in the Language and Cognitive Skills (school-based) domain in the 2017 sample, but this was not consistent in 2016 services rated Meeting NQS or Working Towards NQS.

Table 3: Association between NQS overall quality rating and AEDC domains, as represented by relative risk ratios

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	At risk	1.08 (1.04–1.13)***	1.04 (0.98–1.09)	1.01 (0.97–1.05)	1.05 (1.01–1.09)*	1.01 (0.97–1.05)
Meeting NQS 2017	At risk	1.07 (1.03–1.11)**	1.06 (1–1.11)*	1 (0.96–1.04)	1.03 (0.99–1.07)	1 (0.96–1.05)
Meeting NQS 2016	Vulnerable	1.1 (1.03–1.17)**	1.02 (0.95–1.09)	1.05 (1–1.11)*	1.06 (1.01–1.11)*	1.05 (0.99–1.1)
Meeting NQS 2017	Vulnerable	1.09 (1.03–1.16)**	1.06 (1–1.14)	1.07 (1.02–1.13)**	1.06 (1.01–1.11)*	1.02 (0.97–1.07)
Working Towards NQS# 2016	At risk	1.08 (1.03–1.13)***	1.06 (1–1.12)	1.03 (0.99–1.08)	1.06 (1.02–1.11)**	1.05 (1.01–1.1)*
Working Towards NQS# 2017	At risk	1.06 (1.01–1.11)*	1.07 (1.01–1.14)*	1 (0.96–1.05)	1.06 (1.01–1.11)*	1.07 (1.02–1.13)**
Working Towards NQS# 2016	Vulnerable	1.15 (1.08–1.23)***	1.12 (1.04–1.2)**	1.08 (1.02–1.15)**	1.12 (1.06–1.18)***	1.08 (1.02–1.14)**
Working Towards NQS# 2017	Vulnerable	1.13 (1.06–1.21)***	1.12 (1.04–1.21)**	1.12 (1.06–1.19)***	1.15 (1.09–1.21)***	1.1 (1.04–1.17)**

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Quality Area 1 – Educational program and practice

Consistent in 2016 and 2017

Results for the Quality Area 1 rating analysis are presented in Table 4. These show that, compared to children participating in services rated Exceeding NQS, children in services rated Working Towards NQS were consistently more likely to be developmentally vulnerable for each AEDC domain. Specifically, Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Emotional Maturity, Social Competence, and Physical Health and Wellbeing. In addition, children in services rated Meeting NQS, were consistently more likely be developmentally vulnerable on the Communication Skills and General Knowledge, Emotional Maturity, and Social Competence domains. Further, children in services rated Working Towards NQS were more likely to be developmentally at risk on the Communication Skills and General Knowledge, Social Competence, and Physical Health and Wellbeing domains. Also, children in services rated Meeting NQS were more likely to be developmentally at risk on the Communication Skills and General Knowledge domain.

Inconsistent across 2016 and 2017

Children in services rated Meeting NQS were more likely to be developmentally at risk in the Social Competence domain and developmentally vulnerable in the Physical Health and Wellbeing domain in 2016, but this was not consistent for the 2017 sample. Additionally, children in services rated Working Towards NQS were more likely to be developmentally at risk in the Language and Cognitive Skills (school-based) domain in 2016, but this was not consistent for the 2017 sample.

Table 4: Association between NQS Quality Area 1 (Educational program and practice) and AEDC domains, as represented by relative risk ratios

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	At risk	1.1 (1.06–1.15)***	1.03 (0.97–1.08)	1.02 (0.98–1.07)	1.07 (1.03–1.11)**	1.04 (0.99–1.09)
Meeting NQS 2017	At risk	1.11 (1.06–1.16)***	1.04 (0.99–1.11)	1 (0.95–1.04)	1.04 (1–1.09)	1.03 (0.98–1.07)
Meeting NQS 2016	Vulnerable	1.1 (1.03–1.17)**	1.01 (0.94–1.08)	1.06 (1.01–1.12)*	1.07 (1.02–1.13)**	1.09 (1.03–1.15)**
Meeting NQS 2017	Vulnerable	1.09 (1.02–1.16)*	1.04 (0.97–1.12)	1.08 (1.02–1.14)**	1.06 (1.01–1.12)*	1.05 (0.99–1.1)

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Working Towards NQS# 2016	At risk	1.09 (1.04–1.15)**	1.08 (1.01–1.15)*	1.04 (0.99–1.1)	1.08 (1.03–1.14)**	1.08 (1.03–1.15)**
Working Towards NQS# 2017	At risk	1.08 (1.02–1.14)**	1.07 (1–1.15)	1.02 (0.96–1.07)	1.07 (1.01–1.13)*	1.08 (1.02–1.14)**
Working Towards NQS# 2016	Vulnerable	1.15 (1.07–1.24)***	1.12 (1.03–1.22)**	1.11 (1.04–1.18)**	1.15 (1.09–1.23)***	1.14 (1.06–1.21)***
Working Towards NQS# 2017	Vulnerable	1.12 (1.03–1.21)**	1.14 (1.04–1.24)**	1.12 (1.04–1.2)**	1.17 (1.1–1.25)***	1.1 (1.03–1.18)**

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Quality Area 2 – Children's health and safety

Consistent in 2016 and 2017

Results for the Quality Area 2 rating analysis are presented in Table 5. These show that, compared to children participating in services rated Exceeding NQS, children in services rated Working Towards NQS were consistently more likely to be developmentally vulnerable on the Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), and Social Competence domains. Additionally, children in services rated Working Towards NQS were more likely to be developmentally at risk on the Language and Cognitive Skills (school-based) domains.

Inconsistent across 2016 and 2017

Children in services rated Meeting NQS were more likely to be developmentally vulnerable on the Social Competence domain in 2017, but this was not consistent in the 2016 sample. Further, children in services rated Working Towards NQS were more likely to be developmentally vulnerable in the Emotional Maturity domain, and developmentally vulnerable and developmentally at risk in the Physical Health and Wellbeing domain in 2017, but this was not consistent in the 2016 sample.

Table 5: Association between NQS Quality Area 2 (Children's health and safety) and AEDC domains, as represented by relative risk ratios

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	At risk	1.03 (0.99–1.07)	1.03 (0.97–1.08)	0.99 (0.95–1.03)	0.99 (0.96–1.03)	1.02 (0.98–1.07)
Meeting NQS 2017	At risk	1.03 (0.99–1.08)	1.03 (0.98–1.09)	1 (0.96–1.04)	0.99 (0.95–1.04)	1.01 (0.97–1.06)
Meeting NQS 2016	Vulnerable	1.01 (0.95–1.08)	0.98 (0.92–1.05)	1.02 (0.97–1.07)	1.03 (0.98–1.08)	1 (0.95–1.06)
Meeting NQS 2017	Vulnerable	1.02 (0.96–1.09)	1.06 (0.99–1.13)	1.04 (0.99–1.1)	1.06 (1.01–1.12)*	1.02 (0.97–1.07)
Working Towards NQS# 2016	At risk	1.04 (0.99–1.1)	1.07 (1–1.15)*	1 (0.95–1.05)	1.03 (0.98–1.09)	1.05 (1–1.11)
Working Towards NQS# 2017	At risk	1.04 (0.98–1.1)	1.09 (1.02–1.18)*	0.98 (0.93–1.04)	1.04 (0.98–1.1)	1.07 (1–1.13)*
Working Towards NQS# 2016	Vulnerable	1.12 (1.03–1.2)**	1.09 (1–1.18)*	1.05 (0.99–1.12)	1.07 (1–1.14)*	1.04 (0.98–1.12)
Working Towards NQS# 2017	Vulnerable	1.09 (1–1.18)*	1.12 (1.03–1.23)*	1.09 (1.02–1.17)*	1.13 (1.06–1.21)***	1.1 (1.02–1.18)*

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Quality Area 3 – Physical environment

Consistent in 2016 and 2017

Results for the Quality Area 3 rating analysis are presented in Table 6. These show that compared to children participating in services rated Exceeding NQS, children in services rated Working Towards NQS were consistently more likely to be developmentally vulnerable for each AEDC domain. Specifically, Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Emotional Maturity, Social Competence, and Physical Health and Wellbeing. In addition, children in services rated Meeting NQS were consistently more likely to be developmentally vulnerable on the Communication Skills and General Knowledge, Social Competence, and Physical Health and Wellbeing domains. Further, children in services rated Meeting NQS or Working Towards NQS were more likely to be developmentally at risk on the Communication Skills and General Knowledge and Physical Health and Wellbeing domains. Also, children in services rated Working Towards NQS were more likely to be developmentally at risk on the Social Competence domain.

Inconsistent across 2016 and 2017

Children in services rated Meeting NQS, compared to those in services rated Exceeding NQS, were more likely to be developmentally vulnerable for the Emotional Maturity domain for the 2017 sample, but this was not consistent in the 2016 sample.

Table 6: Association between NQS Quality Area 3 (Physical environment) and AEDC domains, as represented by relative risk ratios

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	At risk	1.06 (1.02–1.11)**	0.98 (0.93–1.04)	1.01 (0.97–1.06)	1.03 (0.99–1.07)	1.05 (1–1.1)*
Meeting NQS 2017	At risk	1.05 (1–1.09)*	1 (0.94–1.06)	1.01 (0.97–1.06)	1.01 (0.97–1.06)	1.04 (1–1.09)*
Meeting NQS 2016	Vulnerable	1.13 (1.06–1.2)***	1.03 (0.96–1.11)	1.05 (1–1.11)	1.06 (1.01–1.12)*	1.06 (1–1.12)*
Meeting NQS 2017	Vulnerable	1.1 (1.03–1.17)**	1.05 (0.98–1.13)	1.09 (1.03–1.15)**	1.07 (1.02–1.13)**	1.08 (1.03–1.14)**
Working Towards NQS# 2016	At risk	1.12 (1.06–1.19)***	1.06 (0.99–1.14)	1.05 (0.99–1.1)	1.08 (1.03–1.14)**	1.14 (1.08–1.21)***

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Working Towards NQS# 2017	At risk	1.09 (1.03–1.16)**	1.07 (0.99–1.16)	1.04 (0.98–1.1)	1.09 (1.03–1.16)**	1.13 (1.06–1.2)***
Working Towards NQS# 2016	Vulnerable	1.25 (1.15–1.36)***	1.2 (1.1–1.31)***	1.11 (1.03–1.19)**	1.15 (1.08–1.23)***	1.2 (1.12–1.28)***
Working Towards NQS# 2017	Vulnerable	1.23 (1.13–1.35)***	1.19 (1.08–1.31)***	1.18 (1.09–1.27)***	1.22 (1.14–1.31)***	1.22 (1.13–1.32)***

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Quality Area 4 – Staffing arrangements

Consistent in 2016 and 2017, but inconsistent quality gradient

Results for the Quality Area 4 rating analysis are presented in Table 7. These show that, compared to children participating in services rated Exceeding NQS, children in services rated Working Towards NQS were consistently more likely to be developmentally vulnerable on the Language and Cognitive Skills (school-based), Social Competence, and Physical Health and Wellbeing domains. Additionally, children in services rated Working Towards NQS were more likely to be developmentally at risk on the Physical Health and Wellbeing domain. Further, children in services rated Meeting NQS were consistently more likely to be developmentally vulnerable on the Communication Skills and General Knowledge and Social Competence domains. However, the association was not statistically reliable for those in services rated Working Towards NQS on the Communication Skills and General Knowledge domain. Further, children in services rated Meeting NQS were also more likely to be developmentally at risk for the Language and Cognitive Skills (school-based) domain.

Inconsistent across 2016 and 2017

Children in services rated Meeting NQS, compared to those in services rated Exceeding NQS, were more likely to be developmentally vulnerable on the Emotional Maturity domain in the 2017 sample, but not in the 2016 sample, suggesting inconsistent effects. Further, children were more likely to be developmentally at risk on the Social Competence domain in 2016, but not in the 2017 sample, and on the Communication Skills and General Knowledge domain in 2017, but not in the 2016 sample, suggesting inconsistent effects.

Table 7: Association between NQS Quality Area 4 (Staffing arrangements) and AEDC domains, as represented by relative risk ratios

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	At risk	1.04 (1–1.07)	1.06 (1.01–1.11)*	1.02 (0.98–1.06)	1.05 (1.01–1.09)*	1.01 (0.97–1.05)
Meeting NQS 2017	At risk	1.04 (1–1.08)*	1.08 (1.03–1.13)**	1 (0.96–1.04)	1.04 (1–1.08)	1.01 (0.97–1.05)
Meeting NQS 2016	Vulnerable	1.08 (1.03–1.15)**	1.02 (0.96–1.08)	1.05 (1–1.1)	1.09 (1.04–1.14)***	1.04 (0.99–1.09)
Meeting NQS 2017	Vulnerable	1.1 (1.04–1.16)***	1.02 (0.96–1.08)	1.11 (1.05–1.17)***	1.11 (1.06–1.17)***	1.03 (0.99–1.08)
Working Towards NQS# 2016	At risk	1.07 (0.99–1.15)	1.03 (0.94–1.13)	1.04 (0.97–1.11)	1.06 (0.99–1.13)	1.08 (1–1.17)*
Working Towards NQS# 2017	At risk	1.06 (0.98–1.15)	1.02 (0.92–1.13)	1.01 (0.94–1.09)	1.01 (0.94–1.09)	1.09 (1–1.19)*
Working Towards NQS# 2016	Vulnerable	1.09 (0.99–1.21)	1.14 (1.03–1.27)*	1.06 (0.97–1.16)	1.15 (1.07–1.25)***	1.12 (1.02–1.22)*
Working Towards NQS# 2017	Vulnerable	1.04 (0.93–1.16)	1.17 (1.04–1.31)**	1.06 (0.96–1.16)	1.17 (1.08–1.27)***	1.12 (1.01–1.23)*

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Quality Area 5 – Relationships with children

Consistent in 2016 and 2017

Results for the Quality Area 5 rating analysis are presented in Table 8. These show that in general, compared to children participating in services rated Exceeding NQS, children in services rated Working Towards NQS were more likely to be developmentally vulnerable for each AEDC domain – specifically, Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Emotional Maturity, Social Competence, and Physical Health and Wellbeing domains. Further, children in services rated Meeting NQS were more likely to be developmentally vulnerable on the Communication Skills and General Knowledge, Emotional Maturity, Social Competence, and Physical Health and Wellbeing domains. In addition, children in services rated Meeting NQS and Working Towards NQS, compared to those rated Exceeding NQS, were more likely to be developmentally at risk on the Communication Skills and General Knowledge and Language and Cognitive Skills (school-based) domains. Also, children in services rated Working Towards NQS were more likely to be developmentally at risk on the Emotional Maturity domain.

Inconsistent across 2016 and 2017

Children in services rated Meeting NQS, compared to those in services rated Exceeding NQS, were more likely to be developmentally vulnerable on the Physical Health and Wellbeing domain for the 2016 sample, but this was not statistically reliable in the 2017 sample. Additionally, children in services rated Working Towards NQS were more likely to be developmentally at risk for the Physical Health and Wellbeing domain in 2017, but this was not consistent for the 2016 sample in services rated Working Towards NQS.

Table 8: Association between NQS Quality Area 5 (Relationships with children) and AEDC domains, as represented by relative risk ratios

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	At risk	1.07 (1.03–1.11)***	1.05 (1–1.1)*	1.04 (1–1.07)	1.02 (0.98–1.05)	1.04 (1–1.08)*
Meeting NQS 2017	At risk	1.06 (1.03–1.11)***	1.06 (1.01–1.12)*	1.01 (0.97–1.05)	1.02 (0.98–1.06)	1.03 (1–1.08)
Meeting NQS 2016	Vulnerable	1.13 (1.07–1.19)***	1.04 (0.98–1.1)	1.05 (1–1.1)*	1.06 (1.01–1.11)*	1.05 (1–1.1)*

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Meeting NQS 2017	Vulnerable	1.13 (1.07–1.2)***	1.04 (0.98–1.1)	1.07 (1.01–1.12)*	1.06 (1.01–1.12)*	1.05 (1.01–1.1)*
Working Towards NQS# 2016	At risk	1.07 (1–1.14)*	1.14 (1.05–1.24)**	1.08 (1.02–1.15)*	1.06 (0.99–1.12)	1.06 (0.99–1.14)
Working Towards NQS# 2017	At risk	1.09 (1.01–1.17)*	1.18 (1.08–1.29)***	1.07 (1–1.15)*	1.06 (0.99–1.14)	1.08 (1–1.17)*
Working Towards NQS# 2016	Vulnerable	1.19 (1.09–1.31)***	1.25 (1.13–1.38)***	1.2 (1.11–1.29)***	1.25 (1.17–1.35)***	1.16 (1.07–1.25)***
Working Towards NQS# 2017	Vulnerable	1.15 (1.04–1.27)**	1.23 (1.11–1.37)***	1.2 (1.1–1.3)***	1.24 (1.15–1.34)***	1.16 (1.07–1.27)***

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Quality Area 6 – Collaborative partnerships with families and communities

Consistent in 2016 and 2017

Results for the Quality Area 6 rating analysis are presented in Table 9. These show that, compared to children participating in services rated Exceeding NQS, children in services rated Working Towards NQS were more likely to be developmentally vulnerable on the Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Social Competence, and Physical Health and Wellbeing domains. Additionally, children in services rated Working Towards NQS were more likely to be developmentally at risk on the Communication Skills and General Knowledge, and Physical and Health and Wellbeing domains. On the other hand, children in services rated Meeting NQS were more likely to developmentally vulnerable on the Communication Skills and General Knowledge domain and Social Competence domain.

Inconsistent across 2016 and 2017

Children in services rated Meeting NQS, compared to those in services rated Exceeding NQS, were more likely to be developmentally at risk on the Emotional Maturity domain for the 2017 sample, but this was not consistent for the 2016 sample.

Table 9: Association between NQS Quality Area 6 (Collaborative partnerships with families and communities) and AEDC domains, as represented by relative risk ratios

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	At risk	1.01 (0.98–1.05)	1 (0.96–1.05)	0.99 (0.95–1.03)	1.02 (0.98–1.06)	1.01 (0.98–1.05)
Meeting NQS 2017	At risk	1.02 (0.99–1.06)	1.03 (0.98–1.08)	0.98 (0.94–1.02)	1.01 (0.97–1.06)	1.03 (0.99–1.07)
Meeting NQS 2016	Vulnerable	1.08 (1.03–1.14)**	1 (0.95–1.06)	1.03 (0.98–1.08)	1.06 (1.01–1.11)*	1.03 (0.99–1.08)
Meeting NQS 2017	Vulnerable	1.08 (1.03–1.14)**	1.01 (0.95–1.07)	1.07 (1.01–1.13)*	1.08 (1.02–1.13)**	1.03 (0.98–1.07)
Working Towards NQS# 2016	At risk	1.09 (1.01–1.17)*	1.07 (0.97–1.18)	1.02 (0.96–1.09)	1.07 (1–1.14)	1.11 (1.03–1.21)**
Working Towards NQS# 2017	At risk	1.13 (1.04–1.23)**	1.06 (0.95–1.18)	1.03 (0.95–1.1)	1.06 (0.99–1.14)	1.13 (1.04–1.24)**
Working Towards NQS# 2016	Vulnerable	1.21 (1.09–1.35)***	1.26 (1.12–1.41)***	1.08 (0.99–1.17)	1.18 (1.1–1.28)***	1.17 (1.07–1.29)***
Working Towards NQS# 2017	Vulnerable	1.19 (1.07–1.33)**	1.29 (1.14–1.45)***	1.08 (0.99–1.18)	1.18 (1.09–1.28)***	1.18 (1.07–1.3)**

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Quality Area 7 – Leadership and service management

Consistent in 2016 and 2017

Results for the Quality Area 7 rating analysis are presented in Table 10. These show that, compared to children participating in services rated Exceeding NQS, children in services rated Working Towards NQS were more likely to be developmentally vulnerable on the Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), and Social Competence domains.

Inconsistent across 2016 and 2017

Children in services rated Working Towards NQS, compared to those in services rated Exceeding NQS, were more likely to be developmentally vulnerable on the Physical Health and Wellbeing domain for the 2016 sample, but this was not consistent in the 2017 sample. Additionally, children in services rated Meeting NQS in the 2017 sample were more likely to be developmentally at risk for the Communication Skills and General Knowledge domain, and developmentally vulnerable on the Social Competence domain, but these were not consistent for the 2016 sample.

Table 10: Association between NQS Quality Area 7 (Leadership and service management) and AEDC domains, as represented by relative risk ratios

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	At risk	1.03 (0.99–1.07)	1 (0.96–1.05)	0.99 (0.95–1.03)	1.03 (1–1.07)	0.99 (0.96–1.03)
Meeting NQS 2017	At risk	1.04 (1–1.08)*	1.01 (0.96–1.06)	0.99 (0.95–1.03)	1.01 (0.98–1.05)	0.98 (0.95–1.02)
Meeting NQS 2016	Vulnerable	1 (0.95–1.06)	0.95 (0.89–1)	0.98 (0.94–1.03)	1.03 (0.98–1.08)	0.99 (0.95–1.04)
Meeting NQS 2017	Vulnerable	1.02 (0.97–1.08)	0.99 (0.93–1.05)	1.03 (0.98–1.08)	1.06 (1.01–1.11)*	1.01 (0.97–1.06)
Working Towards NQS# 2016	At risk	1.04 (0.98–1.09)	1.04 (0.97–1.11)	0.99 (0.94–1.04)	1.05 (0.99–1.1)	0.99 (0.94–1.05)
Working Towards NQS# 2017	At risk	1.05 (0.99–1.11)	1.06 (0.98–1.14)	0.97 (0.91–1.02)	1.03 (0.97–1.09)	1.02 (0.96–1.09)

NQS quality rating	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Working Towards NQS# 2016	Vulnerable	1.09 (1.01–1.17)*	1.09 (1.01–1.18)*	1.02 (0.96–1.09)	1.08 (1.02–1.15)*	1.07 (1–1.14)*
Working Towards NQS# 2017	Vulnerable	1.1 (1.01–1.19)*	1.1 (1–1.2)*	1.04 (0.97–1.12)	1.12 (1.05–1.19)**	1.07 (1–1.15)

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Sensitivity analyses

Imputation: Complete case analysis

Using only cases with complete data resulted in fewer statistically significant associations. Specifically, of 611 statistically significant associations between AEDC domains and quality, 100 (16.4%) were no longer statistically significant for the complete case, while 33 (5.4%) were significant in only the complete case analysis. However, those that differed in statistical significance between methods had similar confidence intervals and point estimates (as did the 478 not differing in statistical significance), and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and methods of weighting. Thus, the results were relatively robust to imputation, which likely increased the ability to detect statistically significant associations due to a larger sample size.

Method of weighting: Average treatment effect vs average treatment effect in the overlap

There were no meaningful differences in the association of quality and AEDC domains when basing inverse probability weighting on the average treatment effect or average treatment effect in the overlap. Specifically, of 611 statistically significant associations between AEDC domains and quality, only 49 (8%) differed in statistical significance between methods. Further, those that differed in statistical significance had nearly identical confidence intervals and point estimates (as did those not differing in statistical significance), and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and imputation methodology. Thus, the results were robust to the method of weighting.

Weighting

Before weighting, the absolute maximum difference (bias) in covariates across all quality rating categories was 0.22 in the 2016 complete sample (median = 0.01) and 0.24 in the 2017 sample (median = 0.01). This indicates some bias in covariates between quality rating categories above the standard threshold of 0.05. However, after weighting, the absolute maximum difference (bias) in covariates between quality rating categories for the average treatment effect and average treatment effect in the overlap, respectively, was 0.04 and 0.02 in the 2016 sample (median = 0.001 and 0.001) and 0.048 and 0.02 in the 2017 sample (median = 0.001 and 0.001). Effective sample sizes after weighting were reduced by up to 41% in 2016 and up to 42% in 2017 across all quality rating areas, with a median reduction of 7% in 2016 and 9% in 2017. Further, examining average balance statistics across the imputed sample showed similar results. Specifically, a maximum difference of 0.2 and 0.24 before adjusting (median = 0.01, 0.01) for the 2016 and 2017 samples respectively. After adjusting there was a maximum difference of 0.03 and 0.02 (median = 0.001, 0.001), and 0.049 and 0.03 (median = 0.001, 0.001) for average treatment effect and average treatment effect in the overlap for 2016 and 2017, respectively. Likewise, the effective sample sizes after weighting for the imputed samples were reduced by up to 42% in 2016 and up to 47% in 2017 across all quality rating areas, with a median reduction of 7% in 2016 and 12% in 2017. Thus, the weighting was successful in reducing bias between covariates and quality rating categories to an acceptable level (less than 0.05) and did not drastically reduce statistical power given the large sample size.

Stratification 1: Income and welfare

Income was stratified by both disposable income and receipt of income support. Table 11 examines the AEDC domains by income and income support to show the nuance of this stratification. Specifically, rates of developmental vulnerability are higher for caregivers with lower income in the total sample. However, the association between income and developmental vulnerability for those not receiving income support instead shows a relatively stable and low rate of developmental vulnerability across the income gradient. On the other hand, there is a higher rate of developmental vulnerability for children receiving income support, and the highest is for children with caregivers in the lowest income percentiles who received income support. Thus, we stratified by: 1) did not receive income support and had any range of disposable income; 2) received income support and disposable income was less than or equal to the 40th percentile; and 3) received income support and had disposable income greater than the 40th percentile.

Table 11: Rates of developmental vulnerability by income support and combined disposable income percentile for the 2016 imputed sample

Income support	80th to 100th	60th to 80th	40th to 60th	40th to 20th	3rd to 20th	2nd
Communication Skills and General Knowledge						
Yes	5.8%	7.1%	8%	9.6%	13.6%	13.2%
No	2.9%	4.2%	4.9%	5.3%	5.5%	6%
Total	3%	4.5%	5.7%	7.1%	10.4%	8.9%
Language and Cognitive Skills (school-based)						
Yes	4.7%	6.8%	7.3%	9.2%	12.8%	12.2%
No	2%	2.7%	3.2%	3.7%	3.8%	3.8%
Total	2.1%	3.1%	4.1%	6%	9.3%	7.2%
Emotional Maturity						
Yes	13.5%	12.3%	13%	14%	14%	14.1%
No	5.9%	7%	7.2%	6.9%	6.9%	5.7%
Total	6.2%	7.6%	8.5%	9.9%	11.1%	9.1%
Social Competence						
Yes	14.7%	13.6%	14.4%	15.4%	17%	16.2%
No	6%	7.1%	7.7%	7.9%	7.8%	7.4%
Total	6.3%	7.8%	9.3%	11.1%	13.3%	11%
Physical Health and Wellbeing						
Yes	11.2%	11.2%	11.9%	13.1%	16.7%	16%
No	5.3%	6.1%	6.6%	6.7%	6.5%	7.6%
Total	5.5%	6.6%	7.8%	9.4%	12.6%	11%

Note: Percentages are median over 10 imputed samples.

Income and welfare: Association between quality and AEDC domains

Results of the stratification analysis by income and income support were generally consistent with non-stratified analyses. Specifically, children in services rated Working Towards NQS or Significant Improvement Required, or in some cases Meeting NQS, had higher rates of developmental vulnerability and developmentally at risk compared to children in services rated as Exceeding NQS for each AEDC domain to varying extents. Table 12 presents the results for overall quality. The results for each quality area are presented in the supplementary materials (Tables S25 to S56).[2]

2 Additional results are available in a series of supplementary tables. Please contact AERO for a copy.

Table 12: Association between NQS overall quality rating and AEDC domains for children by caregiver income category, as represented by relative risk ratios

NQS quality rating	Income stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	Low income	At risk	1.02 (0.94–1.1)	0.97 (0.88–1.07)	1.03 (0.94–1.11)	1.06 (0.97–1.15)	1.02 (0.94–1.12)
Meeting NQS 2017	Low income	At risk	1.06 (0.97–1.15)	1.03 (0.93–1.13)	1.03 (0.94–1.12)	1.01 (0.93–1.1)	1.06 (0.97–1.16)
Meeting NQS 2016	Mid income	At risk	1.03 (0.89–1.2)	0.99 (0.83–1.17)	0.9 (0.78–1.03)	0.94 (0.82–1.07)	1.03 (0.89–1.2)
Meeting NQS 2017	Mid income	At risk	1.06 (0.92–1.22)	1.06 (0.9–1.26)	0.91 (0.79–1.04)	1.07 (0.94–1.23)	1.04 (0.9–1.21)
Meeting NQS 2016	High income	At risk	1.12 (1.07–1.18)***	1.09 (1.02–1.16)*	1.01 (0.97–1.06)	1.05 (1–1.1)*	1 (0.95–1.05)
Meeting NQS 2017	High income	At risk	1.08 (1.03–1.14)**	1.08 (1.01–1.16)*	1 (0.95–1.05)	1.03 (0.98–1.08)	0.98 (0.93–1.03)
Meeting NQS 2016	Low income	Vulnerable	1.1 (0.99–1.22)	1.06 (0.95–1.17)	1.06 (0.97–1.17)	1.05 (0.96–1.15)	1.03 (0.94–1.12)
Meeting NQS 2017	Low income	Vulnerable	1.03 (0.93–1.14)	1.14 (1.03–1.27)*	1.03 (0.94–1.14)	1.05 (0.96–1.15)	1.07 (0.98–1.17)
Meeting NQS 2016	Mid income	Vulnerable	0.97 (0.79–1.19)	0.88 (0.71–1.09)	0.87 (0.74–1.02)	0.98 (0.84–1.15)	1.04 (0.88–1.23)

NQS quality rating	Income stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Meeting NQS 2017	Mid income	Vulnerable	1.1 (0.9–1.35)	1.1 (0.9–1.35)	1.06 (0.9–1.24)	1.03 (0.88–1.2)	0.96 (0.82–1.13)
Meeting NQS 2016	High income	Vulnerable	1.12 (1.03–1.22)**	1 (0.91–1.11)	1.08 (1.01–1.15)*	1.08 (1.02–1.15)*	1.07 (1–1.15)*
Meeting NQS 2017	High income	Vulnerable	1.14 (1.05–1.23)**	0.98 (0.89–1.08)	1.09 (1.02–1.16)**	1.06 (1–1.13)	1.01 (0.94–1.08)
Working Towards NQS# 2016	Low income	At risk	1.04 (0.95–1.14)	1.07 (0.97–1.19)	1.04 (0.96–1.14)	1.08 (0.99–1.18)	1.13 (1.03–1.24)**
Working Towards NQS# 2017	Low income	At risk	1.02 (0.93–1.12)	1.07 (0.96–1.19)	1.06 (0.96–1.16)	1.05 (0.96–1.15)	1.16 (1.05–1.28)**
Working Towards NQS# 2016	Mid income	At risk	1.02 (0.87–1.2)	1.08 (0.9–1.31)	0.98 (0.84–1.14)	0.99 (0.85–1.16)	1.02 (0.86–1.21)
Working Towards NQS# 2017	Mid income	At risk	0.96 (0.81–1.14)	1.06 (0.87–1.29)	0.85 (0.72–1.01)	1.04 (0.89–1.23)	1.08 (0.91–1.29)
Working Towards NQS# 2016	High income	At risk	1.1 (1.05–1.17)***	1.05 (0.97–1.13)	1.03 (0.98–1.09)	1.06 (1.01–1.12)*	1.03 (0.97–1.09)
Working Towards NQS# 2017	High income	At risk	1.1 (1.04–1.17)**	1.08 (0.99–1.17)	1.01 (0.96–1.08)	1.07 (1.01–1.14)*	1.05 (0.98–1.11)
Working Towards NQS# 2016	Low income	Vulnerable	1.17 (1.05–1.3)**	1.21 (1.08–1.34)***	1.12 (1.01–1.23)*	1.15 (1.04–1.26)**	1.06 (0.97–1.17)

NQS quality rating	Income stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Working Towards NQS# 2017	Low income	Vulnerable	1.11 (0.99–1.24)	1.24 (1.1–1.39)***	1.1 (0.99–1.23)	1.19 (1.08–1.31)***	1.13 (1.02–1.25)*
Working Towards NQS# 2016	Mid income	Vulnerable	1.02 (0.82–1.28)	1.11 (0.89–1.39)	0.97 (0.81–1.16)	1.14 (0.96–1.35)	1.21 (1.01–1.45)*
Working Towards NQS# 2017	Mid income	Vulnerable	0.96 (0.76–1.22)	1 (0.78–1.27)	1.16 (0.96–1.39)	1.14 (0.95–1.36)	1.05 (0.87–1.27)
Working Towards NQS# 2016	High income	Vulnerable	1.18 (1.08–1.29)***	1.04 (0.94–1.16)	1.09 (1.01–1.17)*	1.1 (1.03–1.18)**	1.09 (1.01–1.18)*
Working Towards NQS# 2017	High income	Vulnerable	1.2 (1.09–1.31)***	1.04 (0.93–1.17)	1.12 (1.04–1.22)**	1.13 (1.04–1.21)**	1.12 (1.03–1.21)**

Notes: P-value <0.05*, <0.01**, <0.001***. High income = caregivers not receiving income support, Mid income = receiving income support and above the 40th percentile of disposable income, Low income = receiving income support and disposable income in the 40th percentile or lower (low income). Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

No income support (high income)

Examining the consistent results across 2016 and 2017 for overall quality for the group with no income support, compared to children in services rated as Exceeding NQS, children in services rated Working Towards NQS or Significant Improvement Required had consistently higher rates of developmental vulnerability on the Communication Skills and General Knowledge, Emotional Maturity, Social Competence, and Physical Health and Wellbeing domains. They also had higher rates of developmentally at risk for the Communication Skills and General Knowledge domain and Social Competence domain. For children participating in services rated as Meeting NQS, there were consistently higher rates of developmental vulnerability for the Communication Skills and General Knowledge domain and Emotional Maturity domain, as well as consistently higher rates of developmentally at risk for the Communication Skills and General Knowledge domain and Language and Cognitive Skills (school-based) domain.

There were also inconsistent associations. In 2016, but not 2017, children in services rated Meeting NQS had higher rates of developmental vulnerability and developmentally at risk on the Social Competence domain compared to children in services rated Exceeding NQS.

Income support and disposable income in 40th percentile or below (low income)

Investigating the consistent results across 2016 and 2017 for overall quality for the group who received income support and were in the 40th percentile or below, compared to children in services rated as Exceeding NQS, children in services rated Working Towards NQS or Significant Improvement Required had consistently higher rates of developmental vulnerability on the Social Competence as well as the Language and Cognitive Skills (school-based) domains. They also had a higher rate of developmentally at risk on the Physical Health and Wellbeing domain.

There were also inconsistent associations across 2016 and 2017. In 2016, children in services rated Working Towards NQS or Significant Improvement Required had higher rates of developmental vulnerability on the Communication Skills and General Knowledge domain and Emotional Maturity domain compared to children in services rated as Exceeding NQS. In 2017, on the other hand, children in services Working Towards NQS or with Significant Improvement Required had higher rates of developmental vulnerability on the Physical Health and Wellbeing domain. Likewise, for the 2017 sample, children in services rated Meeting NQS, compared to Exceeding NQS, had higher rates for developmentally at risk on the Language and Cognitive Skills (school-based) domain.

Income support and disposable income greater than the 40th percentile (mid income)

There were only inconsistent results across 2016 and 2017 for overall quality for the group who received income support and were above the 40th percentile of disposable income. Specifically, for the 2016 sample, children in services rated as Working Towards NQS or Significant Improvement Required had higher rates of developmental vulnerability on the Physical Health and Wellbeing domain.

Summarising differences in the association between AEDC domains by income and income support

Differences in specific associations between NQS quality and AEDC could be compared in detail across the income stratifications and potentially extrapolated to imply meaningful differences in how NQS quality interacts with the income experiences of children to lead to different developmental outcomes.

However, it is also important to consider differences in model results in the context of statistical principles and general expectations when conducting an analysis that splits a sample. Specifically, stratification can reduce the power to detect statistically significant effects and increase the likelihood of spurious associations. Therefore, deviations from general statistical expectations are likely to identify where results may meaningfully differ by stratification.

Table 13 contains key information to help guide this investigation of consistent differences by stratification. Examining sample size, the high-income group has the largest number of children, followed by the low-income group, and then the mid-income group. Thus, the trend of statistically significant results also follows this gradient, with the most associations detected for the total sample, followed by high income, low income, and mid income with the least. Additionally, nearly all (greater than 80%) of the effects were present in another stratification of the total sample, suggesting few effects were unique to the specific stratification. Finally, inconsistent effects, where higher ratings of quality are associated with higher rates of developmentally vulnerable or developmentally at risk children, were present more frequently in smaller samples, which increases the likelihood of spurious associations. Thus, the pattern of statistically significant effects and their direction suggests that across quality areas, quality ratings, AEDC domains, outcome categories, imputed and complete case samples, years and weighting methodology, **stratification by income was not a reliable moderator that changed the association between NQS quality and AEDC domains**.

Nonetheless, there are some trends that may require additional investigation in future research. First, there was a relatively larger number of unique effects for the high-income group in the imputed sample (18.6%). This, on further investigation, was due to a higher number of statistically significant associations for children developmentally at risk on the Communication Skills and General Knowledge domain in the high-income group (92%) compared to the total sample (58%). In contrast, rates of statistical significance were approximately equal for children rated as developmentally vulnerable in the high-income group (84%) compared to the total sample (83%). Thus, this notable trend appears to be statistical rather than substantive in that higher quality ratings were associated with better developmental outcomes on Communication Skills and General Knowledge in both samples. The second trend was also a relatively higher number of unique statistically significant associations for the low-income group in the complete sample (22.7%). On further investigation, this was due to detecting a higher proportion of effects in 2016 (27.8%) compared to 2017 (13.4%), in greater frequency than in the total sample (2016 = 44.4%; 2017 = 40.6%). Thus, the unique effects also appear to be driven by sample selection rather than substantive differences in associations between NQS quality and AEDC domains for children in the low-income group. Finally, a sizeable number of the statistically significant associations in the complete sample for the mid-income group (24.4%) were inconsistent with the general pattern of higher NQS ratings, indicating lower rates of developmentally vulnerable and developmentally at risk children. This was driven by children in services rated as Meeting NQS, Working Towards NQS or Significant Improvement Required, across all quality areas and methodology, occasionally having lower rates of developmentally at risk for the Emotional Support domain. Given the small sample size (7% of total) and differential trend to all other stratifications and AEDC domains, it seems likely this pattern is spurious. Thus, these exceptions also point to overarching statistical principles explaining discrepancies.

Table 13: Summary of income stratification results

Stratification	Sample size*	Effective sample size	Statistically significant	Effect co-occurs	Effect unique	Inconsistent direction
Imputed sample						
Total	125,613	117,194 (93.3%)	339 (53%)	298 (87.9%)	41 (12.1%)	0 (0%)
High income	88,481	83,615 (94.5%)	285 (44.5%)	232 (81.4%)	**53 (18.6%)**	1 (0.4%)
Low income	28,119	25,706 (91.4%)	169 (26.4%)	150 (88.8%)	19 (11.2%)	1 (0.6%)
Mid income	9,013	8,360 (92.8%)	38 (5.9%)	33 (86.8%)	5 (13.2%)	6 (15.8%)
Complete sample						
Total	89,979	84,491 (93.9%)	272 (42.5%)	249 (91.5%)	23 (8.5%)	0 (0%)
High income	70,526	66,729 (94.6%)	211 (33%)	185 (87.7%)	26 (12.3%)	1 (0.5%)
Low income	13,044	11,961 (91.7%)	132 (20.6%)	102 (77.3%)	**30 (22.7%)**	2 (1.5%)
Mid income	6,409	5,891 (91.9%)	45 (7%)	28 (62.2%)	17 (37.8%)	**11 (24.4%)**

Notes: Compares sample size, effective sample size after weighting, statistically significant association between quality and AEDC domains, co-occurring effects across stratification, unique effects in each stratification, and associations inconsistent with a positive effect of quality. Bold text indicates deviations from general expectation and are addressed in text. *Sample size for Physical Health and Wellbeing domain presented for 2016 sample year. Sample size varies slightly for each AEDC domain (see Appendix Table A2).

Sensitivity analyses: Imputation and complete case analysis

Using only cases with complete data resulted in fewer statistically significant associations across income stratifications. Specifically, of 880 statistically significant associations between AEDC domains and quality, 178 (20.2%) were no longer statistically significant for the complete case, while 74 (8.4%) were significant in only the complete case analysis. However, those that differed in statistical significance between methods had similar confidence intervals and point estimates (as did the 628 not differing in statistical significance), and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and methods of weighting. Thus, the results were relatively robust to imputation, which likely increased the ability to detect statistically significant associations due to a larger sample size.

Sensitivity analyses: Method of weighting

There were no substantive differences in the association of quality and AEDC domains via income stratification when basing inverse probability weighting on the average treatment effect or average treatment effect in the overlap. Specifically, of 880 statistically significant associations between AEDC domains and quality across stratifications, 84 (9.5%) differed in statistical significance between methods. However, those that differed in statistical significance had nearly identical confidence intervals and point estimates (as did those not differing in statistical significance), and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and imputation methodology. Thus, the results were robust to the method of weighting.

Weighting

Before weighting, the absolute maximum difference (bias) in covariates across all quality rating categories and income stratifications was 0.29 in the 2016 complete sample (average median = 0.01) and 0.36 in the 2017 sample (average median = 0.01). This indicates some bias in covariates between quality rating categories above the standard threshold of 0.05. However, after weighting, the absolute maximum difference (bias) in covariates between quality rating categories for the average treatment effect and average treatment effect in the overlap, respectively, was 0.06 and 0.03 in the 2016 sample (average median = 0.002 and 0.002) and 0.05 and 0.04 in the 2017 sample (average median = 0.002 and 0.002). Effective sample sizes after weighting were reduced by up to 49.7% in 2016 and up to 52.1% in 2017 across all quality rating areas, with a median reduction of 7.4% in 2016 and 9% in 2017. Further, examining average balance statistics across the imputed sample showed similar results. Specifically, a maximum difference of 0.28 and 0.37 before adjusting (average median = 0.01, 0.01) for the 2016 and 2017 samples respectively. After adjusting, there was a maximum difference of 0.04 and 0.03 (average median = 0.001, 0.001), and 0.057 and 0.035 (average median = 0.002, 0.002) for average treatment effect and average treatment effect in the overlap for 2016 and 2017, respectively. Likewise, the effective sample sizes after weighting for the imputed samples were reduced by up to 45% in 2016 and up to 55% in 2017 across all quality rating areas, with a median reduction of 7% in 2016 and 10% in 2017. Thus, the weighting was generally successful in reducing bias between covariates and quality rating categories to an acceptable level (less than 0.05).

Stratification 2: Remoteness

Table 14 examines the AEDC domains by remoteness. Specifically, rates of developmental vulnerability increase along a gradient from lowest in major cities of Australia to highest in outer-regional, remote, and very remote Australia for all domains. Rates of developmentally at risk show a similar but less pronounced trend for Language and Cognitive Skills (school-based), Emotional Maturity, and Social Competence domains, while the trend is largely absent for Communication Skills and General Knowledge as well as Physical Health and Wellbeing domains.

Table 14: Rates of developmental vulnerability by remoteness for the 2016 imputed sample

AEDC category	Stratification	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Developmentally vulnerable	Total	6.3%	5.1%	8.8%	9.7%	8.6%
	Major cities	6.2%	4.7%	8.4%	9.5%	8.2%
	Inner regional	6.6%	6.3%	9.6%	10.4%	9.8%
	Outer regional, remote and very remote	6.9%	6.8%	10.5%	10.9%	10.2%
Developmentally at risk	Total	13.6%	7.9%	14.4%	14.7%	12.1%
	Major cities	13.6%	7.5%	14.3%	14.5%	12%
	Inner regional	13.6%	9.2%	14.6%	15.1%	12.4%
	Outer regional, remote and very remote	13.4%	9.2%	15.4%	15.5%	12.1%

Note: Percentages are median over imputed samples.

Remoteness: Association between quality and AEDC domains

Results of the stratification analysis by remoteness were generally consistent with non-stratified analyses. Specifically, children in services rated Working Towards NQS or Significant Improvement Required, or in some cases Meeting NQS, had higher rates of developmentally vulnerable and developmentally at risk children compared to children in services rated as Exceeding NQS for each AEDC domain, to varying extents. Table 15 presents the results for overall quality. The other quality areas are presented in the supplementary materials (Tables S57 to S88).[3]

3 Additional results are available in a series of supplementary tables. Please contact AERO for a copy.

Table 15: Association between NQS overall rating and AEDC domains for children by remoteness of their ECEC service, as represented by relative risk ratios

NQS quality rating	Remoteness stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	Remote	At risk	1.03 (0.89–1.2)	0.98 (0.83–1.17)	0.99 (0.86–1.14)	1.03 (0.89–1.19)	1.09 (0.94–1.28)
Meeting NQS 2017	Remote	At risk	0.93 (0.8–1.09)	0.99 (0.82–1.19)	0.9 (0.78–1.04)	1.02 (0.88–1.18)	0.87 (0.74–1.02)
Meeting NQS 2016	Inner region	At risk	1.07 (0.97–1.18)	0.91 (0.81–1.02)	1.03 (0.94–1.13)	0.97 (0.89–1.07)	1 (0.91–1.1)
Meeting NQS 2017	Inner region	At risk	1.08 (0.97–1.19)	0.92 (0.82–1.03)	1.02 (0.92–1.12)	0.99 (0.9–1.09)	1.03 (0.93–1.13)
Meeting NQS 2016	Major cities	At risk	1.09 (1.04–1.15)***	1.08 (1.01–1.15)*	1 (0.96–1.05)	1.06 (1.01–1.11)*	1 (0.95–1.05)
Meeting NQS 2017	Major cities	At risk	1.09 (1.03–1.14)***	1.1 (1.03–1.17)**	1 (0.96–1.05)	1.03 (0.99–1.08)	1.01 (0.96–1.06)
Meeting NQS 2016	Remote	Vulnerable	1.15 (0.92–1.42)	0.83 (0.67–1.03)	1.03 (0.87–1.23)	1.05 (0.88–1.24)	0.98 (0.82–1.16)
Meeting NQS 2017	Remote	Vulnerable	1 (0.8–1.26)	0.89 (0.71–1.11)	1.05 (0.87–1.27)	0.95 (0.79–1.14)	1.04 (0.86–1.25)
Meeting NQS 2016	Inner region	Vulnerable	1.03 (0.9–1.18)	0.83 (0.72–0.95)**	0.96 (0.86–1.08)	0.99 (0.88–1.1)	1.01 (0.9–1.13)

NQS quality rating	Remoteness stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Meeting NQS 2017	Inner region	Vulnerable	1.05 (0.91–1.21)	0.92 (0.79–1.06)	0.97 (0.86–1.09)	0.98 (0.88–1.1)	0.96 (0.86–1.08)
Meeting NQS 2016	Major cities	Vulnerable	1.12 (1.04–1.21)**	1.11 (1.02–1.2)*	1.08 (1.02–1.15)*	1.08 (1.02–1.14)**	1.07 (1.01–1.14)*
Meeting NQS 2017	Major cities	Vulnerable	1.1 (1.02–1.18)*	1.11 (1.03–1.21)**	1.1 (1.03–1.17)**	1.08 (1.02–1.15)**	1.04 (0.97–1.1)
Working Towards NQS# 2016	Remote	At risk	1.07 (0.9–1.27)	0.94 (0.77–1.16)	1.1 (0.94–1.3)	1 (0.84–1.18)	0.99 (0.82–1.19)
Working Towards NQS# 2017	Remote	At risk	0.95 (0.77–1.16)	0.98 (0.77–1.25)	0.93 (0.77–1.14)	0.96 (0.79–1.17)	0.92 (0.75–1.13)
Working Towards NQS# 2016	Inner region	At risk	0.96 (0.86–1.08)	0.91 (0.8–1.05)	1.06 (0.95–1.19)	1.04 (0.93–1.16)	0.89 (0.78–1)
Working Towards NQS# 2017	Inner region	At risk	0.93 (0.81–1.07)	0.92 (0.79–1.08)	1.03 (0.91–1.18)	0.96 (0.84–1.1)	0.94 (0.81–1.08)
Working Towards NQS# 2016	Major cities	At risk	1.1 (1.05–1.16)***	1.11 (1.04–1.19)**	1.01 (0.96–1.07)	1.07 (1.02–1.12)*	1.1 (1.04–1.16)***
Working Towards NQS# 2017	Major cities	At risk	1.11 (1.05–1.17)***	1.11 (1.04–1.2)**	1.01 (0.95–1.06)	1.08 (1.03–1.14)**	1.12 (1.06–1.18)***
Working Towards NQS# 2016	Remote	Vulnerable	1.19 (0.93–1.51)	1 (0.79–1.26)	1.26 (1.03–1.53)*	1.06 (0.87–1.29)	0.94 (0.77–1.15)

NQS quality rating	Remoteness stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Working Towards NQS# 2017	Remote	Vulnerable	1.17 (0.88–1.55)	1.15 (0.87–1.52)	1.24 (0.98–1.57)	1.03 (0.82–1.29)	0.93 (0.74–1.17)
Working Towards NQS# 2016	Inner region	Vulnerable	0.93 (0.79–1.09)	0.97 (0.83–1.14)	0.97 (0.85–1.12)	0.94 (0.82–1.07)	0.91 (0.8–1.05)
Working Towards NQS# 2017	Inner region	Vulnerable	1.01 (0.83–1.22)	0.97 (0.81–1.17)	1 (0.85–1.17)	1.03 (0.88–1.2)	0.99 (0.85–1.16)
Working Towards NQS# 2016	Major cities	Vulnerable	1.21 (1.13–1.31)***	1.19 (1.09–1.29)***	1.1 (1.03–1.17)**	1.17 (1.1–1.24)***	1.14 (1.07–1.22)***
Working Towards NQS# 2017	Major cities	Vulnerable	1.14 (1.05–1.23)***	1.15 (1.05–1.26)**	1.15 (1.07–1.23)***	1.18 (1.11–1.26)***	1.14 (1.06–1.22)***

Notes: P-value <0.05*, <0.01**, <0.001***. Remote = outer-regional, remote, or very remote Australia, Inner region = inner-regional Australia, Major cities = major cities of Australia. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Major cities

Examining the consistent results across 2016 and 2017 for overall quality for children in services located in major cities, compared to children in services rated as Exceeding NQS, children in services rated Working Towards NQS or Significant Improvement Required had consistently higher rates of developmental vulnerability on the Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Emotional Maturity, Social Competence, and Physical Health and Wellbeing domains. They also had higher rates of children developmentally at risk for the Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Emotional Maturity, and Physical Health and Wellbeing domains. For children participating in services rated Meeting NQS, there were also consistently higher rates of developmental vulnerability for each domain: Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Emotional Maturity, Social Competence, and Physical Health and Wellbeing. Likewise, there were consistently higher rates of children developmentally at risk on the Communication Skills and General Knowledge as well as Language and Cognitive Skills (school-based) domains.

There were also inconsistent associations. In 2016, but not 2017, children in services rated as Meeting NQS had higher rates of developmentally at risk on the Social Competence domain compared to children in services rated as Exceeding NQS.

Inner regional ('inner region')

There were only inconsistent results across 2016 and 2017 for overall quality for the children participating in services in inner-regional Australia. Specifically, services rated Meeting NQS had lower rates of children developmentally at risk on the Language and Cognitive skills (school-based) domain compared to services rated as Exceeding NQS. This inconsistent finding is explored in the summary of results across remoteness categories.

Outer-regional, remote and very remote ('remote')

There were only inconsistent results across 2016 and 2017 for overall quality for the group participating in services in outer-regional, remote and very remote Australia. Specifically, for the 2016 sample, children in services rated as Working Towards NQS or Significant Improvement Required had higher rates of developmental vulnerability on the Emotional Maturity domain compared to children in services rated as Exceeding NQS.

Summarising differences in the association between AEDC domains by income and income support

Differences in specific associations between NQS quality and AEDC could be compared in detail across the remoteness stratifications. They could potentially be extrapolated to imply meaningful differences in how NQS quality interacts with the location and remoteness of the services children attend to lead to different developmental outcomes.

However, it is also important to consider differences in model results in the context of statistical principles and general expectations when conducting an analysis that splits a sample. Specifically, stratification can reduce the power to detect statistically significant effects and increase the likelihood of spurious associations. Therefore, deviations from general statistical expectations are likely to identify where results may meaningfully differ by stratification.

Table 16 contains key information to help guide this investigation of consistent differences by stratification. Examining sample size, the largest number of children participated in ECEC services in major cities, followed by inner-regional services, and then outer-regional, remote and very remote services. Thus, the trend of statistically significant results also follows this gradient, with the most associations detected for the total sample, followed by major cities, inner-regional, and then outer-regional, remote and very remote locations with the least. Additionally, a majority of effects were present in another stratification or the total sample, suggesting few effects unique to the specific stratification. Finally, inconsistent effects, where higher ratings of quality are associated with higher rates of developmentally vulnerable or developmentally at risk children, were present more frequently in smaller samples, which increases the likelihood of spurious associations. Thus, the pattern of statistically significant effects and their direction suggests that across quality areas, quality ratings, AEDC domains, outcome categories, imputed and complete case samples, years and weighting methodology, **stratification by service remoteness was not a reliable moderator that changed the association between NQS quality and AEDC domains**.

Nonetheless, there were 2 notable and linked trends that may require additional investigation in future research. Specifically, there were many unique effects for services in major cities in the complete sample (21.6%). This was due to a higher number of statistically significant associations for developmentally vulnerable and developmentally at risk children on the Language and Cognitive Skills (school-based) domain in the major cities group (58% developmentally at risk; 61% developmentally vulnerable) compared to the total sample (23% developmentally at risk; 41% developmentally vulnerable). However, examining point estimates and confidence intervals shows that these effects were in a similar magnitude and direction. Thus, it seems this trend reflects a better ability to detect statistically significant effects for the Language and Cognitive Skills (school-based) domain, but these effects appear substantively similar in the broader sample. In a related manner, however, the second trend was a sizeable number of the statistically significant associations in the imputed sample for the inner-regional group (44.9 %) being inconsistent with the general pattern of higher NQS ratings, indicating lower rates of developmentally vulnerable and developmentally at risk children. This was driven primarily by children in services rated as Meeting NQS – across all quality areas and methodology – having lower rates of developmental vulnerability for the Language and Cognitive Skills (school-based) domain. Given these trends did not extend to other AEDC domains or children attending services rated Working Towards NQS or Significant Improvement Required, the most likely explanation seems to be the smaller sample size (16.5% of total) may have given rise to spurious or selection effects in the Meeting NQS rating. This also indicates that more effects were detected for Language and Cognitive Skills (school-based) for the major cities group due to the inconsistent association for inner-regional services being removed. Thus, overarching statistical principles primarily explain these exceptions, but future research may be required to confirm selection effects in inner-regional areas.

Table 16: Summary of remoteness stratification results

Stratification	Sample size*	Effective sample size	Statistically significant	Effect co-occurs	Effect unique	Inconsistent direction
Imputed sample						
Total	125,613	117,194 (93.3%)	339 (53%)	331 (97.6%)	8 (2.4%)	0 (0%)
Major cities	95,168	87,948 (92.4%)	387 (60.5%)	334 (86.3%)	53 (13.7%)	0 (0%)
Inner region	21,172	20,180 (95.3%)	69 (10.8%)	44 (63.8%)	25 (36.2%)	**31 (44.9%)**
Remote	9,273	8,719 (94%)	46 (7.2%)	35 (76.1%)	11 (23.9%)	7 (15.2%)
Complete sample						
Total	89,979	84,491 (93.9%)	272 (42.5%)	249 (91.5%)	23 (8.5%)	0 (0%)
Major cities	69,140	64,456 (93.2%)	306 (47.8%)	240 (78.4%)	**66 (21.6%)**	0 (0%)
Inner region	14,669	13,977 (95.3%)	76 (11.9%)	50 (65.8%)	26 (34.2%)	12 (15.8%)
Remote	6,170	5,759 (93.3%)	56 (8.8%)	43 (76.8%)	13 (23.2%)	4 (7.1%)

Notes: Compares sample size, effective sample size after weighting, statistically significant association between quality and AEDC domains, co-occurring effects across stratification, unique effects in each stratification and associations inconsistent with a positive effect of quality. Bold text indicates deviations from general expectations and are addressed in text. *Sample size for Physical Health and Wellbeing domain presented for 2016 sample year. Sample size varies slightly for each AEDC domain (see Appendix Table A2).

Sensitivity analyses: Imputation and complete case analysis

Using only cases with complete data resulted in fewer statistically significant associations across remoteness stratifications. Specifically, of 940 statistically significant associations between AEDC domains and quality ratings, 156 (16.6%) were no longer statistically significant for the complete case, while 92 (9.8%) were significant in only the complete case analysis. However, those that differed in statistical significance between methods had similar confidence intervals and point estimates (as did the 692 not differing in statistical significance), and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and methods of weighting. Thus, the results were relatively robust to imputation, which likely increased the ability to detect statistically significant associations due to a larger sample size.

Sensitivity analyses: Method of weighting

There were no substantive differences in the association of quality and AEDC domains via remoteness stratification when basing inverse probability weighting on the average treatment effect or average treatment effect in the overlap. Specifically, of 940 statistically significant associations between AEDC domains and NQS ratings across stratifications, 138 (14.7%) differed in statistical significance between methods. However, those that differed in statistical significance had nearly identical confidence intervals and point estimates (as did those not differing in statistical significance) and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and imputation methodology. Thus, the results were robust to the method of weighting.

Weighting

Before weighting, the absolute maximum difference (bias) in covariates across all quality rating categories and remote stratifications was 0.3 in the 2016 complete sample (average median = 0.01) and 0.34 in the 2017 sample (average median = 0.01). This indicates some bias in covariates between quality rating categories above the standard threshold of 0.05. After weighting, the absolute maximum difference (bias) in covariates between quality rating categories for the average treatment effect and average treatment effect in the overlap, respectively, was 0.16 and 0.09 in the 2016 sample (average median = 0.002 and 0.002) and 0.18 and 0.08 in the 2017 sample (average median = 0.003 and 0.002). Effective sample sizes after weighting were reduced by up to 80% in 2016 and up to 78% in 2017 across all quality rating areas, with a median reduction of 8% in 2016 and 11% in 2017. Further, examining average balance statistics across the imputed sample showed similar results. Specifically, a maximum difference of 0.3 and 0.31 before adjusting (average median = 0.01, 0.01) for the 2016 and 2017 samples respectively. After adjusting, there was a maximum difference of 0.15 and 0.09 (average median = 0.002, 0.001), and 0.18 and 0.08 (average median = 0.001, 0.001) for average treatment effect and average treatment effect in the overlap for 2016 and 2017 respectively. Likewise, the effective sample sizes after weighting for the imputed samples were reduced by up to 75% in 2016 and up to 55% in 2017 across all quality rating areas, with a median reduction of 8% in 2016 and 12% in 2017. Thus, the weighting was generally successful in reducing bias between covariates and quality rating categories to an acceptable level (less than 0.05), but on occasion weighting by average treatment effect did not completely balance all covariates on all quality ratings for the remoteness stratifications. The approach taken of also adjusting the regression models for covariates was, therefore, useful for increasing robustness for the results of stratification by remoteness.

Stratification 3: English-speaking background and ability

Table 17 examines the AEDC domains by English-speaking ability. Comparing the top-level categories of children who speak only English at home to children who speak a language other than English showed comparable rates of developmental vulnerability on Language and Cognitive Skills (school-based) and Social Competence domains. On the other hand, developmental vulnerability on the Communication Skills and General Knowledge domain was lowest for children who speak only English at home, while rates on Emotional Maturity as well as Physical Health and Wellbeing were lowest for children who speak a language other than English. Rates of developmentally at risk children were comparable, except rates of developmentally at risk were higher on the Communication Skills and General Knowledge domain for children who speak a language other than English at home.

Examining the 2 subsets of children who speak a language other than English at home, however, showed a consistent gradient. Children considered to be learning English as a second language had higher rates of being developmentally vulnerable and developmentally at risk on all AEDC domains. Likewise, children with poor proficiency in effective use of language in English and ability to listen in English had exceptionally high – indeed the highest – rates of being developmentally vulnerable and developmentally at risk across Language and Cognitive Skills (school-based), Emotional Maturity, Social Competence, and Physical Health and Wellbeing domains. Outcomes for Communication Skills and General Knowledge were not evaluated for the emerging proficiency stratification, as 2 items from this domain are used to identify the stratification. Indeed, due to comorbidity among developmental outcomes, this in part explains high rates of developmentally vulnerable and developmentally at risk children for this stratification.

Table 17: Rates of developmental vulnerability by English-speaking ability for the 2016 complete case sample

Stratification	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Developmentally vulnerable					
Total	5.1%	3.8%	7.3%	8.4%	7.7%
English only	4.2%	3.7%	7.4%	8.4%	8.2%
Other than English	9.1%	3.9%	6.7%	8.4%	6.1%
OTE ESL	12.5%	4.9%	7.5%	9.7%	6.5%
OTE emerging	–	28.2%	34.5%	43.6%	27.8%
Developmentally at risk					
Total	12.7%	6.8%	11.6%	13.8%	13.8%
English only	11.8%	6.9%	11.7%	14%	13.8%
Other than English	16.1%	6.4%	11.1%	13%	13.8%
OTE ESL	19.2%	7.7%	12.2%	14%	14.9%
OTE emerging	–	25.5%	23.6%	28.2%	31.5%

English-speaking ability: Association between quality and AEDC domains

Results of the stratification analysis by English-speaking ability were generally consistent with non-stratified analyses. Specifically, children in services rated Working Towards NQS or Significant Improvement Required, or in some cases Meeting NQS, had higher rates of developmentally vulnerable and developmentally at risk children compared to children in services rated as Exceeding NQS for each AEDC domain, to varying extents. Table 18 presents the results for overall quality. The other quality areas are presented in the supplementary materials (Tables S89 to S104).[4]

4 Additional results are available in a series of supplementary tables. Please contact AERO for a copy.

Table 18: Association between NQS overall quality rating and AEDC domains by children's English-speaking background and ability, as represented by relative risk ratios

NQS quality rating	English-speaking stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Exceeding NQS^ (2016/2017)	–	–	Baseline	Baseline	Baseline	Baseline	Baseline
Meeting NQS 2016	English only	At risk	1.1 (1.04–1.16)***	1.08 (1.01–1.16)*	1 (0.95–1.05)	1.06 (1.01–1.11)*	1 (0.95–1.06)
Meeting NQS 2017	English only	At risk	1.05 (1–1.11)	1.05 (0.98–1.12)	0.96 (0.91–1.01)	1.01 (0.96–1.07)	1 (0.95–1.06)
Meeting NQS 2016	Other than English	At risk	1.04 (0.93–1.17)	0.93 (0.78–1.1)	1.01 (0.9–1.14)	0.98 (0.87–1.11)	1 (0.88–1.14)
Meeting NQS 2017	Other than English	At risk	1.12 (1–1.26)	1.04 (0.87–1.24)	1.05 (0.93–1.18)	1.04 (0.92–1.18)	1.03 (0.91–1.17)
Meeting NQS 2016	OTE ESL	At risk	1.04 (0.91–1.19)	0.89 (0.72–1.09)	1.07 (0.92–1.24)	1.1 (0.94–1.29)	1.12 (0.95–1.32)
Meeting NQS 2017	OTE ESL	At risk	1.08 (0.94–1.24)	1.04 (0.84–1.29)	1.05 (0.9–1.22)	1.06 (0.9–1.24)	1.07 (0.91–1.26)
Meeting NQS 2016	OTE emerging	At risk	–	0.9 (0.58–1.38)	1.07 (0.69–1.67)	1.16 (0.73–1.87)	1.06 (0.67–1.66)
Meeting NQS 2017	OTE emerging	At risk	–	1.33 (0.85–2.06)	1.1 (0.72–1.69)	1.16 (0.73–1.83)	0.79 (0.5–1.24)

NQS quality rating	English-speaking stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Meeting NQS 2016	English only	Vulnerable	1.04 (0.95–1.14)	1.02 (0.93–1.13)	1.1 (1.03–1.18)**	1.04 (0.97–1.11)	1.03 (0.96–1.1)
Meeting NQS 2017	English only	Vulnerable	1.07 (0.98–1.17)	1.01 (0.92–1.11)	1.1 (1.03–1.17)**	1.04 (0.98–1.11)	1.02 (0.95–1.09)
Meeting NQS 2016	Other than English	Vulnerable	1.19 (1.02–1.38)*	1.13 (0.9–1.43)	1.03 (0.86–1.22)	1.11 (0.95–1.3)	1.09 (0.92–1.29)
Meeting NQS 2017	Other than English	Vulnerable	1.1 (0.95–1.29)	1.16 (0.93–1.46)	1.12 (0.95–1.34)	0.99 (0.85–1.16)	1.03 (0.87–1.22)
Meeting NQS 2016	OTE ESL	Vulnerable	1.15 (0.97–1.37)	1.16 (0.88–1.53)	1.09 (0.87–1.36)	1.12 (0.93–1.36)	1.09 (0.87–1.35)
Meeting NQS 2017	OTE ESL	Vulnerable	1.06 (0.89–1.26)	1.12 (0.86–1.45)	1.18 (0.94–1.47)	1.03 (0.86–1.25)	1.06 (0.86–1.3)
Meeting NQS 2016	OTE emerging	Vulnerable	–	1.12 (0.71–1.78)	1 (0.63–1.6)	1.59 (1.02–2.49)*	1.4 (0.89–2.19)
Meeting NQS 2017	OTE emerging	Vulnerable	–	1.78 (1.15–2.76)**	1.38 (0.88–2.16)	1.41 (0.92–2.16)	1.23 (0.8–1.88)
Working Towards NQS# 2016	English only	At risk	1.07 (1.01–1.14)*	1.05 (0.97–1.13)	1.03 (0.97–1.09)	1.04 (0.98–1.1)	1.02 (0.96–1.09)
Working Towards NQS# 2017	English only	At risk	1.04 (0.97–1.11)	1.05 (0.97–1.15)	0.98 (0.92–1.05)	1.04 (0.97–1.11)	1.04 (0.97–1.11)

NQS quality rating	English-speaking stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Working Towards NQS# 2016	Other than English	At risk	1.16 (1.03–1.3)*	1.01 (0.85–1.19)	0.99 (0.88–1.12)	1.01 (0.89–1.14)	1.14 (1–1.3)
Working Towards NQS# 2017	Other than English	At risk	1.22 (1.08–1.37)**	0.96 (0.8–1.15)	1 (0.89–1.14)	1.04 (0.92–1.19)	1.14 (1–1.31)*
Working Towards NQS# 2016	OTE ESL	At risk	1.12 (0.98–1.29)	0.95 (0.77–1.16)	0.94 (0.81–1.09)	1.11 (0.95–1.3)	1.23 (1.04–1.44)*
Working Towards NQS# 2017	OTE ESL	At risk	1.09 (0.94–1.26)	0.87 (0.69–1.08)	0.98 (0.84–1.15)	1.01 (0.85–1.18)	1.13 (0.95–1.33)
Working Towards NQS# 2016	OTE emerging	At risk	–	0.96 (0.62–1.5)	1.14 (0.73–1.79)	1.6 (0.98–2.61)	1.63 (1.02–2.59)*
Working Towards NQS# 2017	OTE emerging	At risk	–	1.02 (0.65–1.59)	1.16 (0.75–1.8)	1.5 (0.93–2.4)	1.3 (0.82–2.07)
Working Towards NQS# 2016	English only	Vulnerable	1.08 (0.98–1.19)	1.13 (1.02–1.26)*	1.14 (1.06–1.23)***	1.09 (1.02–1.18)*	1.1 (1.02–1.18)*
Working Towards NQS# 2017	English only	Vulnerable	1.07 (0.96–1.2)	1.07 (0.95–1.2)	1.16 (1.07–1.26)***	1.11 (1.02–1.2)*	1.1 (1.01–1.19)*
Working Towards NQS# 2016	Other than English	Vulnerable	1.31 (1.12–1.52)***	1.21 (0.96–1.52)	1 (0.84–1.2)	1.19 (1.02–1.39)*	1.22 (1.03–1.45)*
Working Towards NQS# 2017	Other than English	Vulnerable	1.23 (1.05–1.43)**	1.2 (0.96–1.51)	1.03 (0.85–1.24)	1.13 (0.96–1.32)	1.17 (0.99–1.4)

NQS quality rating	English-speaking stratification	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
Working Towards NQS# 2016	OTE ESL	Vulnerable	1.25 (1.05–1.48)*	1.14 (0.87–1.49)	1.07 (0.86–1.33)	1.18 (0.97–1.42)	1.33 (1.07–1.64)**
Working Towards NQS# 2017	OTE ESL	Vulnerable	1.13 (0.95–1.35)	1.09 (0.84–1.41)	0.99 (0.79–1.25)	1.13 (0.93–1.36)	1.15 (0.93–1.42)
Working Towards NQS# 2016	OTE emerging	Vulnerable	–	1.06 (0.67–1.69)	0.98 (0.61–1.56)	1.83 (1.15–2.92)*	1.78 (1.13–2.79)*
Working Towards NQS# 2017	OTE emerging	Vulnerable	–	1.35 (0.87–2.1)	1.1 (0.7–1.75)	1.75 (1.14–2.7)*	1.7 (1.11–2.62)*

Notes: P-value <0.05*, <0.01**, <0.001***. English only = children who do not speak a language other than English at home, Other than English = speak a language other than English at home, OTE ESL = speak a language other than English at home and considered to be learning English as a second language, OTE emerging = speak a language other than English at home and rated as poor proficiency in ability to use language effectively in English or listening ability in English. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. ^Includes Excellent rating. #Includes Significant Improvement Required rating. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2016 is 2 years before school, 2017 is the year before school. Results combined from complete case analysis.

Speak English only at home ('English only')

Examining the consistent results across 2016 and 2017 for overall quality for children who speak only English at home, compared to children in services rated as Exceeding NQS, children in services rated Working Towards NQS or Significant Improvement Required had consistently higher rates of developmental vulnerability on the Emotional Maturity, Social Competence, and Physical Health and Wellbeing domains. Likewise, children in services rated Meeting NQS, compared to services rated as Exceeding NQS, had higher rates of developmental vulnerability on the Emotional Maturity domain.

There were also inconsistent associations. In 2016, but not 2017, children in services rated Working Towards NQS or Significant Improvement Required, compared to children in services rated as Exceeding NQS, had higher rates of developmental vulnerability on the Language and Ccognitive Skills (school-based) domain. Likewise, in 2016, children in services rated Meeting NQS compared to children in services rated as Exceeding NQS had higher rates of children developmentally at risk on the Communication Skills and General Knowledge, Language and Cognitive Skills (school-based) and Social Competence domains.

Speak language other than English at home ('Other than English')

Regarding consistent results across 2016 and 2017 for children who speak a language other than English at home, children in services rated Working Towards NQS or Significant Improvement Required for overall quality had higher rates of being developmentally vulnerable and developmentally at risk on the Communication Skills and General Knowledge domain.

Inconsistent results across 2016 and 2017 for overall quality were also observed. In 2016, children in services rated Working Towards NQS or Significant Improvement Required, compared to those rated as Exceeding NQS, had higher rates of developmental vulnerability on the Social Competence domain and Physical Health and Wellbeing domain. Likewise, in 2017, children in services rated Working Towards NQS or Significant Improvement Required had a higher rate of developmentally at risk on the Physical Health and Wellbeing domain.

Speak language other than English at home and considered to be learning English as a second language ('OTE ESL')

There were only inconsistent results across 2016 and 2017 for overall quality for the children who speak a language other than English at home and were considered to be learning English as a second language. Specifically, in 2016, children in services rated Working Towards NQS or Significant Improvement Required had higher rates of developmental vulnerability, compared to children in services rated as Exceeding NQS, on the Communication Skills and General Knowledge as well as Physical Health and Wellbeing domains. Additionally, in 2017, children in services rated Working Towards NQS or Significant Improvement Required had a higher rate of developmentally at risk on the Physical Health and Wellbeing domain.

Speak language other than English at home and poor proficiency in English ('OTE emerging')

There were consistent trends across 2016 and 2017 for overall quality for the children who speak a language other than English at home and were rated as having poor proficiency in ability to use language effectively in English or listening ability in English. Specifically, children participating in services rated as Working Towards NQS or Significant Improvement Required, compared to services rated as Exceeding NQS, had higher rates of developmental vulnerability on the Emotional Maturity domain and Physical Health and Wellbeing domain.

There were also inconsistent associations across 2016 and 2017. In 2016, children participating in services rated as Working Towards NQS or Significant Improvement Required, compared to services rated as Exceeding NQS, had a higher rate of being developmentally at risk on the Physical Health and Wellbeing domain. Likewise, in 2016, children in services rated Meeting NQS had higher rates of developmental vulnerability on the Social Competence domain, while in 2017 they had higher rates of developmental vulnerability on the Language and Cognitive Skills (school-based) domain.

It is also important to note that as this cohort is defined using items from the Communication Skills and General Knowledge domain, this outcome was not assessed. The cohort is also more likely to have confounding and biased associations between AEDC domains and NQS quality due to selecting on a possible intermediate outcome.

Summarsing differences in the association between AEDC domains by English-speaking ability and background

Differences in specific associations between NQS quality and AEDC domains could be compared in detail across the English-speaking stratifications and potentially extrapolated to imply meaningful differences in how NQS quality of the services children attend interacts with their English-speaking ability and background to lead to different developmental outcomes.

However, it is also important to consider differences in model results in context of statistical principles and general expectations when conducting an analysis that splits a sample. Specifically, stratification can reduce the power to detect statistically significant effects and increase the likelihood of spurious associations. Therefore, deviations from general statistical expectations are likely to identify where results may meaningfully differ by stratification.

Table 19 contains key information to help guide this investigation of consistent differences by stratification. Examining sample size, the largest number of children only spoke English at home, followed by children who spoke a language other than English at home, and then children who also considered English as a second language, and finally those speaking a language other than English at home who had emerging English proficiency. Thus, the trend of statistically significant results also follows this gradient, with the most associations detected for the total sample, followed by English speaking only, other than English, and English as a second language and emerging proficiency with the fewest. Additionally, most (greater than 85%) of the effects were present in another stratification or the total sample, suggesting few effects unique to the specific stratification. Finally, inconsistent effects, where higher ratings of quality are associated with higher rates of developmentally vulnerable or developmentally at risk children, were present more frequently in smaller samples, which increases the likelihood of spurious associations. Thus, the pattern of statistically significant effects and their direction suggests that across quality areas, quality ratings, AEDC domains, outcome categories, years and weighting methodology, **stratification by English-speaking ability was not a reliable moderator that changed the association between NQS quality and AEDC domains**.

Nonetheless, there was one notable trend that may require additional investigation in future research. Specifically, although having a much smaller sample (1.2% of total) and overlapping substantially, children speaking a language other than English with emerging proficiency had a similar number of statistically significant results as children speaking a language other than English who were considered English as a second language (12% of total). As there were no obvious distinguishing trends across methodology factors, it is likely that the higher rates of developmentally vulnerable and developmentally at risk children in the emerging proficiency cohort (Table 17) may increase the statistical power to an equivalent level as the broader English as a second language cohort. Thus, general statistical principles seem to be a likely explanation, but future work distinguishing the emerging proficiency cohort before entry to ECEC (to remove intermediate outcome confounding) may reveal a stronger association with quality for this stratification.

Table 19: Summary of English-speaking background and ability stratification results

Stratification	Sample size*	Effective sample size	Statistically significant	Effect co-occurs	Effect unique	Inconsistent direction
Complete sample (including Communication Skills and General Knowledge)						
Total	89,979	84,491 (93.9%)	272 (42.5%)	225 (82.7%)	47 (17.3%)	0 (0%)
English only	72,307	68,997 (95.4%)	170 (26.6%)	162 (95.3%)	8 (4.7%)	2 (1.2%)
Other than English	17,672	15,456 (87.5%)	139 (21.7%)	119 (85.6%)	20 (14.4%)	1 (0.7%)
OTE ESL	10,793	9,345 (86.6%)	89 (13.9%)	78 (87.6%)	11 (12.4%)	2 (2.2%)
Complete sample (excluding Communication Skills and General Knowledge)						
Total	89,979	84,491 (93.9%)	184 (35.9%)	162 (88%)	22 (12%)	0 (0%)
English only	72,307	68,997 (95.4%)	133 (26%)	125 (94%)	8 (6%)	2 (1.5%)
Other than English	17,672	15,456 (87.5%)	85 (16.6%)	72 (84.7%)	13 (15.3%)	1 (1.2%)
OTE ESL	10,793	9,345 (86.6%)	64 (12.5%)	57 (89.1%)	7 (10.9%)	2 (3.1%)
OTE emerging	1,090#	827 (75.8%)	**64 (12.5%)**	54 (84.4%)	10 (15.6%)	1 (1.6%)

Notes: Compares of sample size, effective sample size after weighting, statistically significant association between quality and AEDC domains, co-occurring effects across stratification, unique effects in each stratification, and associations inconsistent with a positive effect of quality. Bold text indicates deviations from general expectation and are addressed in text. *Sample size for Physical Health and wellbeing domain presented for 2016 sample year. Sample size varies slightly for each AEDC domain (see Appendix Table A2). # rounded to nearest 10.

Sensitivity analyses: Method of weighting

There were no substantive differences in the association of quality and AEDC domains via English-speaking background and proficiency stratification when basing inverse probability weighting on the average treatment effect or average treatment effect in the overlap. Specifically, of 462 statistically significant associations between AEDC domains and quality across stratifications, 76 (16.5%) differed in statistical significance between methods. However, those that differed in statistical significance had nearly identical confidence intervals and point estimates (as did those not differing in statistical significance), and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and imputation methodology. Thus, the results were robust to the method of weighting.

Weighting

Before weighting, the absolute maximum difference (bias) in covariates across all quality rating categories and English background and proficiency stratifications was 0.23 in the 2016 complete sample (average median = 0.02) and 0.43 in the 2017 sample (average median = 0.02). This indicates some bias in covariates between quality rating categories above the standard threshold of 0.05. After weighting, the absolute maximum difference (bias) in covariates between quality rating categories for the average treatment effect and average treatment effect in the overlap, respectively, was 0.08 and 0.05 in the 2016 sample (average median = 0.004 and 0.003) and 0.15 and 0.07 in the 2017 sample (average median = 0.004 and 0.003). Effective sample sizes after weighting were reduced by up to 58% in 2016 and up to 70% in 2017 across all quality rating areas, with a median reduction of 12% in 2016 and 17% in 2017. Thus, the weighting was generally successful in reducing bias between covariates and quality rating categories to an acceptable level (less than 0.05), but on occasion weighting by average treatment effect did not completely balance all covariates on all quality ratings for English-speaking background. This was likely due to the small number of children in the other than English emerging proficiency stratification. The approach taken of also adjusting the regression models for covariates was, therefore, useful for increasing robustness for the results of stratification by English-speaking background.

Latent class analysis: PLIDA FFY

Number of classes and profiles

As additional classes from 1 to 10 were included, the model fit improved. This can be seen by continuous decreases in the value of the AIC and BIC, with entropy remaining high (greater than 0.9) for both the 2016 and 2017 sample (see Table 20). The decreases in AIC and BIC, however, slowed after 4 classes were included. This suggests, at minimum, a 4-class solution would provide a good fit based on the AIC, BIC and entropy. The 6-class solution was chosen as the best fit after consideration of patterns of the classes.

Table 20: Fit statistics for 1 to 10 latent classes in the analysis of the 18 standards of the NQS quality areas in 2016 and 2017 samples

Number of classes	2016			2017		
	Entropy	AIC	BIC	Entropy	AIC	BIC
1	–	173,141	173,379	–	1914,97	191,738
2	0.95	132,477	132,959	0.95	145,046	145,534
3	0.93	118,907	119,633	0.94	129,014	129,750
4	0.91	114,888	115,858	0.91	124,464	125,448
5	0.91	112,119	113,333	0.94	121,371	122,602
6	0.95	110,890	112,348	0.95	120,080	121,560
7	0.94	109,818	111,521	0.93	118,806	120,533
8	0.94	109,075	111,021	0.95	118,141	120,115
9	0.92	108,436	110,627	0.92	117,495	119,717
10	0.92	108,159	110,594	0.93	117,168	119,638

As additional classes were fit, however, the class profiles began to reflect changes in relative performance across all quality standards, compared to a mix of low and high performance in different quality areas or standards. Therefore, we chose to pursue a 6-class solution that illustrated differences in higher-rated services as well as capturing variability in services with lower ratings. The 6-class solutions also demonstrated high (greater than 0.85) average posterior probabilities, suggesting adequate class separation, ranging from 0.96 to 0.87 in 2016, and 0.96 to 0.88 in 2017.

Profiles for the classes are presented in Figure 3 for 2016 and Figure 4 for 2017. As the profiles were similar in both years, they are described and labelled together. The first class was labelled 'exceeding all' and captured services that were characterised by high rates of Exceeding NQS across the 18 quality standards. The second class was labelled 'Exceeding Quality Areas 1 and 5', and also had high rates of Exceeding NQS ratings across all quality standards, but particularly in Quality Area 1 (Educational program and practice) and Quality Area 5 (Relationships with children). The third class, labelled 'Exceeding Quality Areas 6 and 7' was similar, with higher rates of Exceeding NQS across the quality standards, but with increased prevalence of Exceeding NQS for standards in Quality Area 6 (Collaborative partnerships with families and communities) and Quality Area 7 (Leadership and service management). Fourth was the class 'meeting all', which had a very high proportion of Meeting NQS ratings across all quality standards and very few ratings of Working Towards NQS or Exceeding NQS across the standards. The fifth class was labelled 'Working Towards, more Standards Meeting'. Services in this class had a relatively high proportion of Meeting NQS ratings across quality standards, but also a moderate number of ratings of Working Towards NQS. Finally, the sixth class, 'Working Towards, more Standards Working Towards', had the highest prevalence of Working Towards NQS ratings across all quality standards.

Figure 3: Latent class profiles for the 6-class solution across the 18 NQS standards, 2016 sample

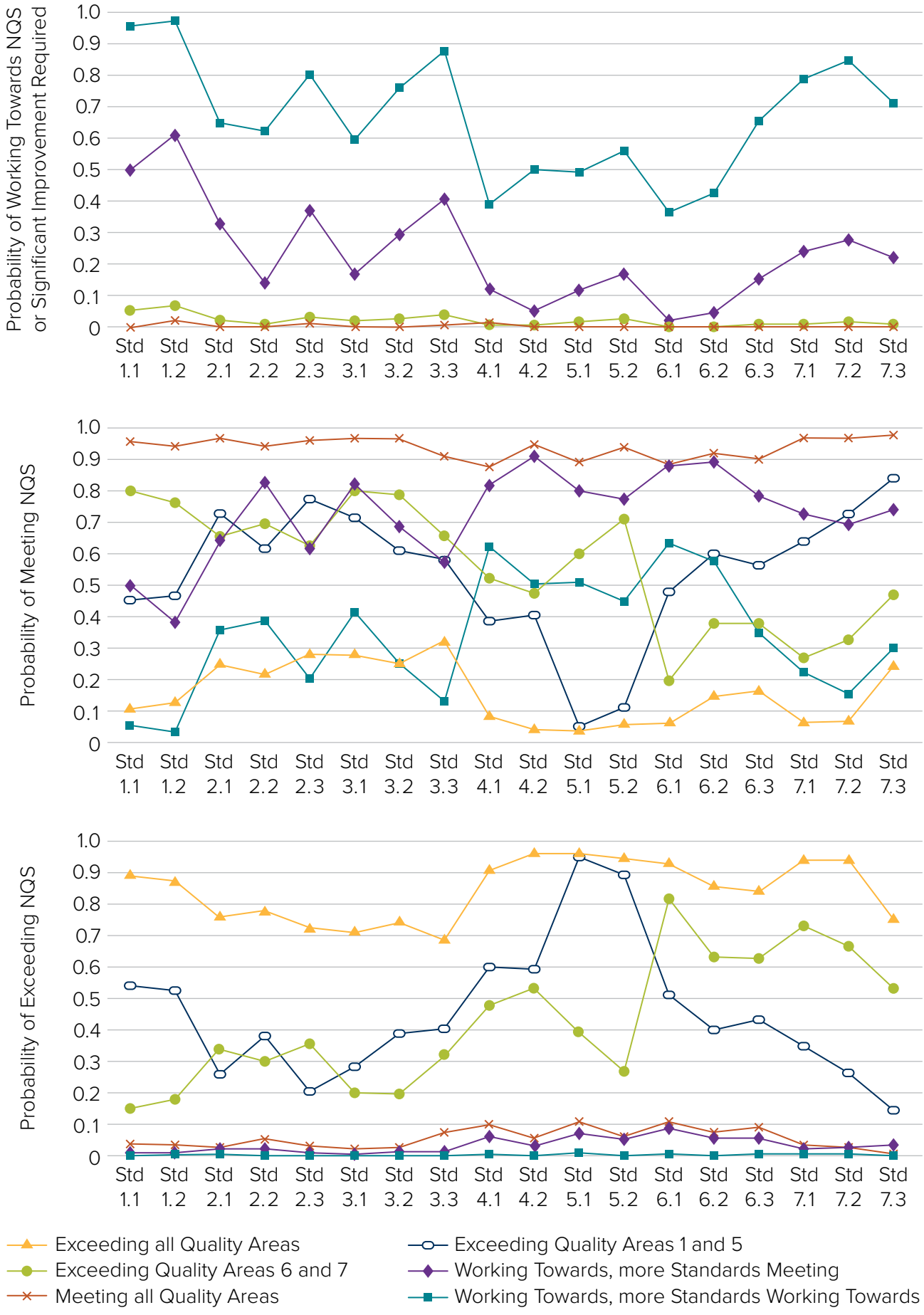
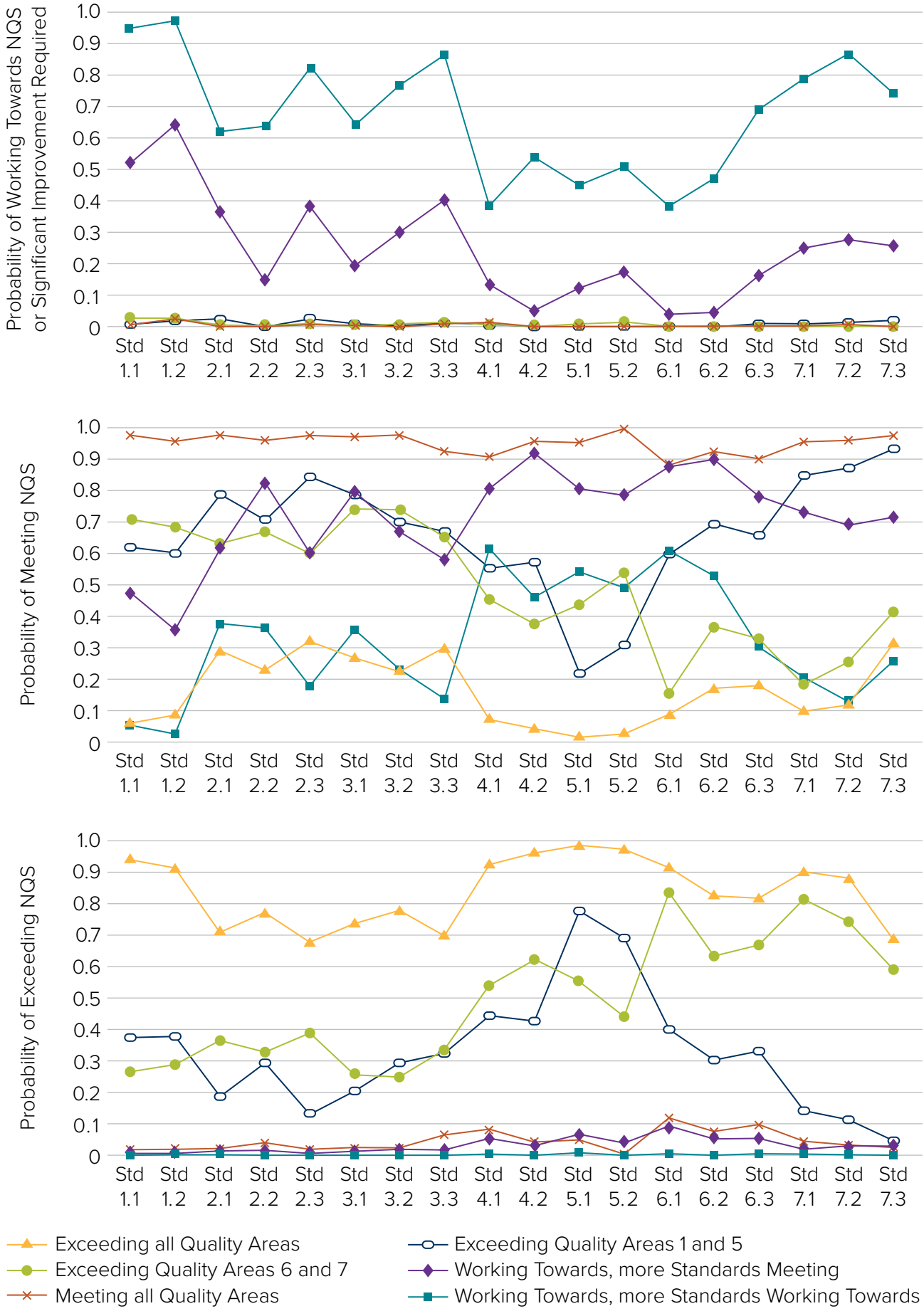


Figure 4: Latent class profiles for the 6-class solution across the 18 NQS standards, 2017 sample



Details on the number of services most likely to be in each class profile and service characteristics are presented in Table 21 for 2016 and 2017. Notable insights include a higher ratio of CBDC services in the classes more likely to exceed NQS standards, while a higher proportion of FDC services aligned with the profiles that were Working Towards NQS or Meeting NQS. Nonetheless, a sizeable number of services were represented in all class profiles.

Table 21: Proportion of service types and services in each latent class of NQS Quality Standards in 2016 and 2017

Latent class	2016			2017		
	FDC	CBDC	Total	FDC	CBDC	Total
Exceeding all	46 (11.6%)	949 (18.9%)	995 (18.4%)	43 (8.5%)	1,161 (21.3%)	1,161 (21.3%)
Exceeding QA 1 and 5	19 (4.8%)	661 (13.2%)	680 (12.5%)	27 (5.3%)	734 (13.5%)	734 (13.5%)
Exceeding QA 6 and 7	61 (15.4%)	727 (14.5%)	788 (14.5%)	71 (14%)	912 (16.7%)	912 (16.7%)
Meeting all	77 (19.5%)	1,366 (27.2%)	1,443 (26.6%)	81 (16%)	1,393 (25.5%)	1,393 (25.5%)
Working Towards, more Meeting	100 (25.3%)	962 (19.1%)	1,062 (19.6%)	139 (27.5%)	905 (16.6%)	905 (16.6%)
Working Towards, more Working Towards	92 (23.3%)	360 (7.2%)	452 (8.3%)	145 (28.7%)	352 (6.5%)	352 (6.5%)
All	395 (7.3%)	5,025 (92.7%)	–	506 (8.5%)	5,457 (91.5%)	–

Links between latent classes and AEDC domains

Consistent in 2016 and 2017

Results for links between latent classes and AEDC domains are shown in Table 22 for the imputed sample and average treatment effect (see supplementary Tables S105 to S107 for all other models[5]). Compared to children in the class 'Exceeding all Quality Areas', children in the class 'Exceeding Quality Areas 6 and 7' were more likely to be developmentally vulnerable and developmentally at risk on the Communication Skills and General Knowledge domain. Also, children in this class were more likely to be developmentally at risk on the Physical Health and Wellbeing domain. In addition, children in the class 'Meeting all Quality Areas' were more likely to be developmentally at risk on the Communication Skills and General Knowledge domain. Further, children in the class 'Working Towards, more Standards Working Towards' were more likely to be developmentally vulnerable on the Communication Skills and General Knowledge, Language and Cognitive Skills (school-based), Social Competence, and Physical Health and Wellbeing domains. Also, children in this class were more likely to be developmentally at risk on the Communication Skills and General Knowledge as well as the Physical Health and Wellbeing domains. Additionally, children in the class 'Working Towards, more Standards Meeting' were more likely to be developmentally vulnerable on the Communication Skills and General Knowledge and Social Competence domains.

Inconsistent across 2016 and 2017

Children in the class 'Exceeding Quality Areas 1 and 5', compared to children in the class 'Exceeding all', were more likely to be developmentally at risk on the Communication Skills and General Knowledge domain and developmentally vulnerable on the Emotional Maturity domain in the 2017 sample, but this was not consistent in the 2016 sample. Additionally, children in the class 'Exceeding Quality Areas 6 and 7' were more likely to be developmentally at risk on the Language and Cognitive skills (school-based) domain in the 2017 sample, and not in the 2016 sample, and developmentally vulnerable in the 2016 sample, and not in the 2017 sample, suggesting inconsistent effects. Also, in this class children were more likely to be developmentally vulnerable on the Emotional Maturity domain in the 2017 sample, but this was not consistent in the 2016 sample. Further, children in the class 'Meeting all' were more likely to be developmentally vulnerable on the Communication Skills and General Knowledge, Emotional Maturity, and Social Competence domains in the 2017 sample but not in the 2016 sample, suggesting inconsistent effects. In addition, children in the class 'Working Towards, more Standards Meeting' were more likely to be developmentally vulnerable on the Emotional Maturity as well as Physical Health and Wellbeing domains in the 2017 sample, but these were not consistent in the 2016 sample. Also, children in the class 'Working Towards, more Standards Meeting' were more likely to be developmentally at risk on the Language and Cognitive skills (school-based) and developmentally vulnerable on the Emotional Maturity domains in the 2017 sample, but these were not consistent in the 2016 sample. Whereas children in this class were also more likely to be developmentally at risk on the Social Competence domain in the 2016 sample, but this was not reliable in the 2017 sample.

5 Additional results are available in a series of supplementary tables. Please contact AERO for a copy.

Table 22: Association between latent classes of NQS Quality Standards in 2016 and 2017 and AEDC domains, as represented by relative risk ratios

Year	NQS latent class	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
–	Exceeding all	–	Baseline	Baseline	Baseline	Baseline	Baseline
2016	Exceeding QA 1 and 5	At risk	1.02 (0.95–1.08)	1 (0.92–1.09)	0.95 (0.89–1.01)	1.04 (0.98–1.11)	1 (0.94–1.07)
2017	Exceeding QA 1 and 5	At risk	1.09 (1.02–1.16)*	1.07 (0.98–1.16)	1.02 (0.95–1.09)	1.03 (0.96–1.1)	1 (0.94–1.07)
2016	Exceeding QA 1 and 5	Vulnerable	0.95 (0.86–1.05)	1 (0.9–1.11)	0.99 (0.91–1.08)	0.99 (0.92–1.08)	0.99 (0.91–1.07)
2017	Exceeding QA 1 and 5	Vulnerable	1 (0.91–1.11)	1.06 (0.95–1.18)	1.11 (1.02–1.21)*	1.05 (0.97–1.14)	0.99 (0.91–1.08)
2016	Exceeding QA 6 and 7	At risk	1.13 (1.07–1.2)***	1.05 (0.97–1.13)	1.06 (1.01–1.13)*	1.02 (0.97–1.08)	1.08 (1.02–1.15)**
2017	Exceeding QA 6 and 7	At risk	1.09 (1.03–1.16)**	1.11 (1.03–1.19)**	1.01 (0.96–1.07)	1 (0.95–1.07)	1.08 (1.02–1.14)*
2016	Exceeding QA 6 and 7	Vulnerable	1.17 (1.07–1.27)***	1.11 (1.01–1.21)*	1.07 (1–1.15)	1.03 (0.97–1.11)	1.09 (1.01–1.17)*
2017	Exceeding QA 6 and 7	Vulnerable	1.12 (1.03–1.22)**	1.1 (1–1.2)	1.13 (1.05–1.21)**	1.07 (1–1.15)	1.06 (0.99–1.14)
2016	Meeting all	At risk	1.07 (1.02–1.13)**	1 (0.94–1.07)	1.01 (0.95–1.06)	1.02 (0.97–1.08)	0.99 (0.94–1.05)
2017	Meeting all	At risk	1.08 (1.02–1.14)**	1.04 (0.97–1.12)	0.99 (0.94–1.05)	1.02 (0.96–1.07)	1 (0.95–1.06)
2016	Meeting all	Vulnerable	1.08 (1–1.16)	0.96 (0.88–1.04)	1.04 (0.97–1.11)	1.06 (1–1.13)	1.03 (0.96–1.1)
2017	Meeting all	Vulnerable	1.09 (1.01–1.18)*	0.99 (0.91–1.08)	1.09 (1.01–1.17)*	1.08 (1.01–1.16)*	1.04 (0.98–1.11)
2016	Working Towards, more Meeting	At risk	1.06 (1–1.12)*	1.04 (0.96–1.11)	1.03 (0.97–1.09)	1.03 (0.97–1.09)	1.04 (0.98–1.11)

Year	NQS latent class	AEDC domain indicator category	Communication Skills and General Knowledge	Language and Cognitive Skills (school-based)	Emotional Maturity	Social Competence	Physical Health and Wellbeing
2017	Working Towards, more Meeting	At risk	1.06 (1–1.13)	1.1 (1.01–1.19)*	1.02 (0.96–1.09)	1.05 (0.99–1.12)	1.07 (1–1.14)*
2016	Working Towards, more Meeting	Vulnerable	1.13 (1.04–1.23)**	1.07 (0.98–1.18)	1.06 (0.99–1.14)	1.1 (1.03–1.18)**	1.06 (0.99–1.14)
2017	Working Towards, more Meeting	Vulnerable	1.13 (1.04–1.24)**	1.06 (0.96–1.17)	1.13 (1.04–1.22)**	1.16 (1.08–1.25)***	1.1 (1.02–1.19)*
2016	Working Towards, more Working Towards	At risk	1.12 (1.03–1.22)**	1.11 (0.99–1.24)	1 (0.92–1.09)	1.09 (1–1.18)*	1.12 (1.02–1.23)*
2017	Working Towards. more Working Towards	At risk	1.15 (1.04–1.27)**	1.17 (1.03–1.33)*	1.01 (0.92–1.11)	1.06 (0.97–1.16)	1.11 (1–1.24)*
2016	Working Towards, more Working Towards	Vulnerable	1.25 (1.11–1.41)***	1.3 (1.14–1.48)***	1.1 (0.99–1.22)	1.21 (1.1–1.33)***	1.27 (1.14–1.41)***
2017	Working Towards, more Working Towards	Vulnerable	1.21 (1.05–1.39)**	1.4 (1.2–1.62)***	1.12 (1–1.26)*	1.25 (1.12–1.38)***	1.23 (1.08–1.4)**

Notes: P-value <0.05*, <0.01**, <0.001***. Vulnerable = Developmentally vulnerable, At risk = Developmentally at risk. QA = Quality Area. Model adjusted for covariates and with inverse probability weighting for the average treatment effect. Years denote when quality of ECEC experience occurred; 2017 is the year before school, 2016 is 2 years before school. Results combined from 10 datasets using multiple imputation. Coefficient and 95% Confidence Interval reported. Sample size varies slightly for each domain based on presence of valid domain category indicator.

Sensitivity analyses: Imputation and complete case analysis

Using only cases with complete data resulted in fewer statistically significant associations across the latent class models. Specifically, of 155 statistically significant associations between AEDC domains and latent class of quality, 30 (19.4%) were no longer statistically significant for the complete case, while 9 (5.8%) were significant in only the complete case analysis. However, those that differed in statistical significance between methods had similar confidence intervals and point estimates (as did the 628 not differing in statistical significance), and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and methods of weighting. Thus, the results were relatively robust to imputation, which likely increased the ability to detect statistically significant associations due to a larger sample size.

Sensitivity analyses: Method of weighting

There were no substantive differences in the association of latent class of quality and AEDC domains via income stratification when basing inverse probability weighting on the average treatment effect or average treatment effect in the overlap. Specifically, of 155 statistically significant associations between AEDC domains and quality latent class, 17 (11%) differed in statistical significance between methods. However, those that differed in statistical significance had nearly identical confidence intervals and point estimates (as did those not differing in statistical significance), and they were distributed among sample years (2016 and 2017), AEDC domains, developmentally vulnerable and developmentally at risk AEDC categories, and imputation methodology. Thus, the results were robust to the method of weighting.

Weighting

Before weighting, the absolute maximum difference (bias) in covariates across all latent classes was 0.2 in the 2016 complete sample (median = 0.01) and 0.28 in the 2017 sample (median = 0.01). This indicates some bias in covariates between quality rating categories above the standard threshold of 0.05. However, after weighting, the absolute maximum difference (bias) in covariates between latent classes for the average treatment effect and average treatment effect in the overlap, respectively, was 0.04 and 0.03 in the 2016 sample (median = 0.001 and 0.001) and 0.06 and 0.05 in the 2017 sample (median = 0.001 and 0.002). Effective sample sizes after weighting were reduced by up to 39% in 2016 and up to 49% in 2017 across all quality rating areas, with a median reduction of 10% in 2016 and 12% in 2017. Further, examining average balance statistics across the imputed sample showed similar results. Specifically, a maximum difference of 0.21 and 0.37 before adjusting (median = 0.01, 0.01) for the 2016 and 2017 samples respectively. After adjusting, there was a maximum difference of 0.04 and 0.03 (median = 0.001, 0.001), and 0.07 and 0.05 (median = 0.001, 0.002) for average treatment effect and average treatment effect in the overlap for 2016 and 2017 respectively. Likewise, the effective sample sizes after weighting for the imputed samples were reduced by up to 41% in 2016 and up to 57% in 2017 across all quality rating areas, with a median reduction of 10% in 2016 and 13% in 2017. Thus, the weighting was generally successful in reducing bias between covariates and latent classes of quality to an acceptable level (less than 0.05).

Summary of key findings: PLIDA FFY

This analysis investigated the link between children's experience of ECEC quality indexed by the NQS ratings and AEDC domain indicator categories in the first year of full-time schooling. The 2018 AEDC cohort was examined, focusing on quality experienced 2 years before school (2016; complete n = 89,988, imputed n = 125,625) and in the year before school (2017; complete n = 82,357, imputed n = 116,356).

Key findings include:

- **NQS Quality Areas**
  - NQS quality ratings were associated with AEDC domains. Children in services rated Exceeding NQS or Excellent had lower rates of developmental vulnerability compared to children in services rated Meeting NQS, and those in services rated Working Towards NQS or Significant Improvement Required.
  - Overall NQS quality and Quality Areas 1 (Educational program and practice), 3 (Physical environment) and 5 (Relationships with children) had the greatest and most consistent association with AEDC domains. Quality Area 6 (Collaborative partnerships with families and communities) was similarly associated, but to a lesser extent.
  - Quality Areas 2 (Children's health and safety), 4 (Staffing arrangements), and 7 (Leadership and service management) had fewer consistent associations with AEDC domains.

- **AEDC domains**
  - Children were consistently more likely to be developmentally vulnerable for each AEDC domain if the service was rated Working Towards NQS or Significant Improvement Required compared to services rated Exceeding NQS or Excellent.
  - Children most consistently exhibited higher rates of being developmentally vulnerable on the Communication Skills and General Knowledge, Emotional Maturity, and Physical Health and Wellbeing AEDC domains when the service quality was rated Meeting NQS, as compared to services rated Exceeding NQS or Excellent.
  - Children consistently had elevated rates of being developmentally at risk on the Communication Skills and General Knowledge domain if the service quality was rated Working Towards NQS, Significant Improvement Required or Meeting NQS when compared to services rated Exceeding NQS or Excellent.
  - The link between quality ratings and AEDC domains was most frequently and consistently observed when comparing rates of children who were developmentally vulnerable compared to rates of children who were developmentally at risk.

- **Latent class analysis of NQS standards**
  - The latent class analyses of the 18 NQS standards distinguished 6 types of ECEC service quality. These map to the broader classes of Working Towards NQS, Meeting NQS and Exceeding NQS, with variation within. They include:
    1. Exceeding all Quality Areas
    2. Exceeding in Quality Areas 1 and 5
    3. Exceeding in Quality Areas 6 and 7
    4. Meeting all Quality Areas
    5. Working Towards, more Standards Meeting
    6. Working Towards, more Standards Working Towards.
  - Consistent with the analysis of overall quality and quality areas, services rated Working Towards NQS or Meeting NQS had higher rates of developmental vulnerability compared to services rated as Exceeding NQS. Children in services in latent classes 4, 5 and 6 had higher rates of developmental vulnerability for each AEDC domain when compared to the services typically exceeding on all standards (1). Children in services that exceeded in Quality Area 6 (Collaborative partnerships with families and communities) and Quality Area 7 (Leadership and service management) also had higher rates of developmental vulnerability on the Communication Skills and General Knowledge domain, compared to children in services more likely to exceed all standards.

Overall finding

The national assessment and rating system serves as a predictor of child outcomes but not uniformly so. Some quality areas of assessment were less consistently associated with developmental vulnerability across AEDC domains.

Next steps

The current analyses of NQS quality and AEDC domains were strengthened by an extensive sample size and the availability of a large range of variables that facilitated statistical adjustment for confounding explanatory mechanisms. Moreover, the results were robust to the weighting methodology and accounting for missing data. Additionally, adequate balance across quality ratings was achieved in nearly all analyses, including within stratifications with substantially fewer children. Thus, a range of confounding was accounted for successfully. However, future research would be strengthened by the addition of baseline assessments of each child's development prior to the assessment of the ECEC services' quality. Without this information, the current analysis has limited ability to confirm that differences in outcomes in the first year of full-time schooling are due to quality of ECEC or selection into ECEC of particular types and quality. Additionally, the statistical models were limited in the ability to account for nesting and cross-classification when children moved between services, though using modern inference techniques. The models could, therefore, be extended as new analytical capability is unlocked to account for more underlying processes. Further, the incidence validity (whom the results are relevant for) could be increased by including data on preschool participation that resides with state and territory jurisdictions. Finally, the services in the study were rated by the NQS 2012 version, and the NQS ratings of services came from 2012 through to 2017. Updating the analyses with information based on the latest ratings and the 2018 version of the NQS would further strengthen their relevance and provide an excellent sensitivity test.

Methods: E4Kids

Data source: E4Kids

Sample

The sample for these analyses comes from the Early Effective Education Experiences (E4Kids) study. E4Kids was a 5-year longitudinal investigation of N = 2,606 children recruited from Australian ECEC services. It focused on rooms providing for children aged 3 to 4 years. The protocol for this study has been published in detail (Tayler et al., 2013; Tayler et al., 2016).

E4Kids captured a broad range of Australian ECEC services across geographical locations and socio-economic areas. It did this by explicit and implicit sampling protocols based on large-scale studies (Adams & Wu, 2003). The first stage identified 4 locations – 2 major cities of Australia (Melbourne, Victoria and Brisbane, Queensland), 1 inner-regional location (Shepparton, Victoria), and 1 remote location (Mt Isa, Queensland; see Pink, 2011 for location classifications). The second stage stratified based on type of program (CBDC, Pre-K, FDC, limited hours care), service capacity and relative socio-economic advantage and disadvantage status of the locality. Thus, E4Kids provides a broad and holistic view of the early education experience of a diverse profile of Australian children.

In 2010, the first year of the study, 3- to 4-year-old children were recruited from N = 142 ECEC services, comprising CBDC (n = 92), standalone Pre-K (also known as preschool; n = 40), FDC (n = 7) and limited hours care (n = 3). For all classrooms in these services with a study child in 2010, observations were made using CLASS (Pre-K; Pianta et al., 2008a; Pianta et al., 2008b) and ECERS-R (Harms et al., 1998). Trained fieldworkers also individually tested children using the WJ III Tests of Cognitive Abilities and Tests of Achievement (Woodcock et al., 2001). Further information about the children was obtained from surveys completed by the primary caregiver (e.g., parent), educators and ECEC directors.

In the subsequent years, the initial sample were tracked into ECEC and school services via supplied email and phone contacts with a range of complementary and expanded information collected. All participation in the study was voluntary and could be withdrawn at any time. The E4Kids study was approved by the University of Melbourne Human Research Ethics Committee (ID 0932660.2), and the current analysis was approved by the UQ Human Research Ethics Committee (ID 2023/HE001018).

For the present study, we examined 2 subsets of the E4Kids sample. First, we selected rooms from the E4Kids study with information on service type (excluding limited hours care; 1.2% of sample), CLASS and partial information on ECERS-R for a latent class analysis. Thus, we examined the profiles of 242 rooms from 130 services. Second, we examined a sample of 1,969 children with information on CLASS in 2010 (n services = 140; n rooms = 249) and their cognitive outcomes in 2010 (n with = 1,776) or 2011 (n with = 1,715) to evaluate whether the quality of ECEC they experience in 2010 predicts gains in outcomes into the following year.

Treatment/key covariate: CLASS and ECERS-R

The CLASS observational measure (Pianta et al., 2008a; Pianta et al., 2008b) was the first key covariate. We examine the 3 subscales of Emotional Support (4 dimensions), Classroom Organization (3 dimensions), and Instructional Support (3 dimensions). Each subscale is on a 7-point scale ranging from 1 (lowest quality) to 7 (highest quality). An observation was considered valid in this study if there were between 2 and 6 CLASS completed observation cycles of 20 minutes each (for example, 1–3 hours per room). Most observations (latent class analysis: 2 = 1.2%; 3 = 6.6%; 4 = 31%; 5 = 40%; 6 = 21%; outcome analysis: 2 = 1.2%; 3 = 7.6%; 4 = 31.3%; 5 = 38.6%; 6 = 21.3%) applied to our sample had 4 or more cycles. Observations in the current sample were drawn from 32 research staff (31 for latent class sample) who were trained and certified as reliable adhering to the standard CLASS protocols (Pianta et al., 2008a; Pianta et al., 2008b; Tayler et al., 2016). Psychometric results (for example, measurement invariance across cycles) of the subscales for these observations were excellent and are published elsewhere (Thorpe et al., 2020), while in-field assessment of fieldworkers against a gold standard CLASS-coder was conducted in 2011 and agreement within one rating was high (96.4%, Cloney et al., 2017). Information on the items contributing to CLASS are provided in Table 23.

The ECERS-R observational measure (Harms et al., 1998) was the second key covariate. A subset of key items was collected. Specifically, observations in this sample were from 26 trained research staff (26 for latent class sample) regarding the quality of routines (6 items), quality of furnishings (8 items), and quality of activities (10 items). Each quality rating ranges from 1 to 7, with the average score across ratings used. Information on the items contributing to ECERS-R is provided in Table 23.

Table 23: Structure of CLASS and ECERS-R measures of ECEC quality

Measure	Domain	Dimensions/items
CLASS, Pre-K	Emotional Support	Positive climate; Negative climate (reverse scored); Teacher sensitivity; Regard for student perspectives
	Instructional Support	Concept development; Quality of feedback; Language modelling
	Classroom Organization	Behaviour management; Productivity; Instructional learning formats
ECERS-R, Third Edition	Furnishings	8 items: Indoor space; Furniture for routine care, play and learning; Furnishings for relaxation and comfort; Room arrangement for play; Space for privacy; Child-related display; Space for gross motor play; Gross motor equipment
	Routines	6 items: Greeting/departing; Meals/snacks; Nap/rest; Toileting/ diapering; Health practices; Safety practices
	Activities	10 items: Fine motor; Art; Music/movement; Blocks; Sand/water; Dramatic play; Nature/science; Math/number; Use of TV, video, and/or computers; Promoting acceptance of diversity

Outcomes

Child cognition and achievement were measured using the WJ III Tests of Cognitive Abilities and Tests of Achievement (Woodcock et al., 2001). Specifically, we utilised 5 W scores from the WJ III test battery; Cognitive 1-W Verbal Comprehension (measuring lexical knowledge and language development), 5-W Concept Formation (induction), 6-W Visual Matching (perceptual speed), Achievement 4-W Understanding Directions (listening ability, language development) and 10-W Applied Problems (quantitative reasoning, math achievement, math knowledge) tests (Woodcock et al., 2001). These were administered in person by trained fieldworkers.

Covariates: E4Kids

A range of child, caregiver and ECEC service covariates were included in the models examining cognitive outcomes for the E4Kids sample.

Child: Gender

A binary indicator of the child's gender as either female or male.

Child: First Nations status

A binary indicator of the child's First Nations status as either First Nations or non-First Nations.

Child: Born overseas

A binary indicator of the whether the child was born overseas or in Australia.

Child: Developmental delay

A binary indicator of whether the child was considered to have a developmental delay.

Child: English-speaking background

An indictor of the child's English-speaking background, including English main and only language; English main language, but also speaks other language; and English not main language.

Child: Age in years at baseline cognitive test, and time between baseline and follow-up

Child age in years (calculated at monthly intervals, centred at the grand mean) at the time of the cognitive test was included in the model, as well as the time between the baseline test and follow up test.

Child: Difficult temperament

Difficult temperament was assessed via the 12-item Short Temperament Scale for Children (Sanson et al., 1994). This scale comprises 3 subscales (persistence, reactivity and sociability), derived from aggregation of 4 items. Caregivers were asked to rate their child's behaviour (e.g., 'If this child is upset, it is hard to comfort him/her') on a 6-point Likert scale (1 = almost never to 6 = almost always). Lower persistence, higher reactivity and lower sociability indicate a more difficult child temperament, so the persistence and sociability scales are reverse coded and averaged with reactivity to derive the difficult temperament scores (Prior et al., 1989).

Caregiver: Education

An indicator of the primary caregiver's highest level of education: postgraduate degree, bachelor degree, diploma, Year 12 or TAFE (technical college) certificate, or Year 10 or lower.

Caregiver: Health Care Card

A binary measure of whether the caregiver had a low-income Health Care Card.

Caregiver: Born overseas

A binary indicator of the whether the child was born overseas or in Australia.

Caregiver: First Nations status

A binary indicator of the caregivers' First Nations status as either First Nations or non-First Nations.

Caregiver: Stressful life events

The presence and effect on the child of 11 stressful life events were assessed (Holmes & Rahe, 1967) – for example, death of someone close to the caregiver. It was coded as 0 = event not experienced, or event experienced but no negative effect on child, or event experienced and negative effect on child, and 1 = event experienced and had serious negative effect on child. The maximum score across the 11 events was retained.

Caregiver: Psychological distress

Caregiver psychological distress was measured by the Kessler-10 (Kessler et al., 2002). This measure is the sum of 10 items that measure psychological distress in the past 30 days. For example, 'How often did you feel hopeless?' (1 = none of the time; 5 = all of the time).

Caregiver: Home learning environment

The home learning environment was measured using the average of 12 items that related to learning materials and interaction at home. For each item, caregivers rated the frequency of an activity over the last week on a 0–7 scale (0 = no days; 7 = 7 days) – for example, read to the study child from a book.

ECEC: Service type

A categorical indicator of service type, including preschool, CBDC, preschool program for 3-year-old children in the second year before school, FDC and limited hours care.

Analytical plan: E4Kids

Latent class analysis of CLASS and ECERS-R quality

Latent class analysis of CLASS domains (Emotional Support, Instructional Support and Classroom Organization) and ECERS-R items (furnishings, 8 items; routines, 6 items; activities, 10 items) was undertaken in Mplus (Muthén & Muthén, 2017; version 8.8). We estimated 1 to 4 classes. Model indices evaluated included AIC (lower is better), BIC (lower is better), entropy (0–1; closer to 1 indicates better fit), the diagonals of average posterior probabilities (values greater than 0.8 are desirable), and Lo-Mendell-Rubin likelihood ratio test of model fit (Lo et al., 2001; direct test if the number of classes – e.g., 3 – is an equivalent, or worse, fit to one less class – e.g., 2). We initially ran 100 random sets of starting values in the initial stage and 20 optimisations in the final stage. We then doubled these numbers until the log-likelihood was repeated and then doubled once more to repeat the log-likelihood a final time. Further, we evaluated model fit by considering whether the patterns illustrated meaningful variation across CLASS and ECERS-R measures.

We ran models, including service type in the latent class model, or including service type as a covariate (3-step approach to incorporate class uncertainty; Asparouhov & Muthén, 2014) to be predicted from latent class. This was to evaluate that in the absence of service type directly influencing latent class profiles, there remained, or did not remain, an association between the classes and service type.

Some modifications were made due to low sample numbers. Specifically, occasional care services (n = 3; 1.2%) were excluded. Additionally, the activities items regarding the 'use of TV, video, and/or computers' was excluded due to very low availability in the sample (n = 128; 52.9%).

Effect of CLASS and ECERS-R quality on gains in cognitive ability and achievement

To estimate the effect of CLASS and ECERS-R on gains in cognitive ability and achievement outcomes, we used multilevel regression models. The models included a random intercept for the room the child was in and always adjusted for baseline outcome of the child, age of the child at baseline, and time difference between tests. Further models were fit that adjusted for the full range of covariates. Additionally, as there were high correlations between CLASS domains (r=0.58 to 0.84), the models were estimated separately for each domain. Thus, these models are the 'value add' specification used to understand how characteristics of the ECEC experience predict gains in outcomes (e.g., Curby et al., 2013).

All analyses were run in R (R Core Team, 2022; version 4.2.1) using the lme4 package (Bates et al., 2015). Missing data was imputed using multiple imputation via the mice package (van Buuren & Groothuis-Oudshoorn, 2011; version 3.16.0), with 50 datasets imputed and appropriate grouping structure specified for room-level variables.

Results: E4Kids

E4Kids: Latent class analysis

Sample description

Descriptive statistics of the rooms included in the latent class analysis are presented in Appendix Table A5. Briefly, the sample captured a broad range of scores on CLASS and ECERS-R, as well as having a reasonable diversity of service types.

Number of classes and profiles

As additional classes, from 1 to 4, were included, the model fit improved on some statistics (Table 24). This can be seen by continuous decreases in the value of the AIC and BIC, with entropy remaining high (greater than 0.9) beyond 3-classes. The decreases in AIC and BIC, however, were marginal beyond 2-classes. Additionally, the Lo-Mendell-Rubin test statistics strongly suggested more than a 2-class solution did not improve the fit. This suggests a 2 or 3-class solution is likely sufficient to provide a good fit based on the AIC, BIC, entropy and Lo-Medell-Rubin test.

Table 24: Fit statistics for 1 to 5 latent classes in the analysis of CLASS and ECERS-R indices

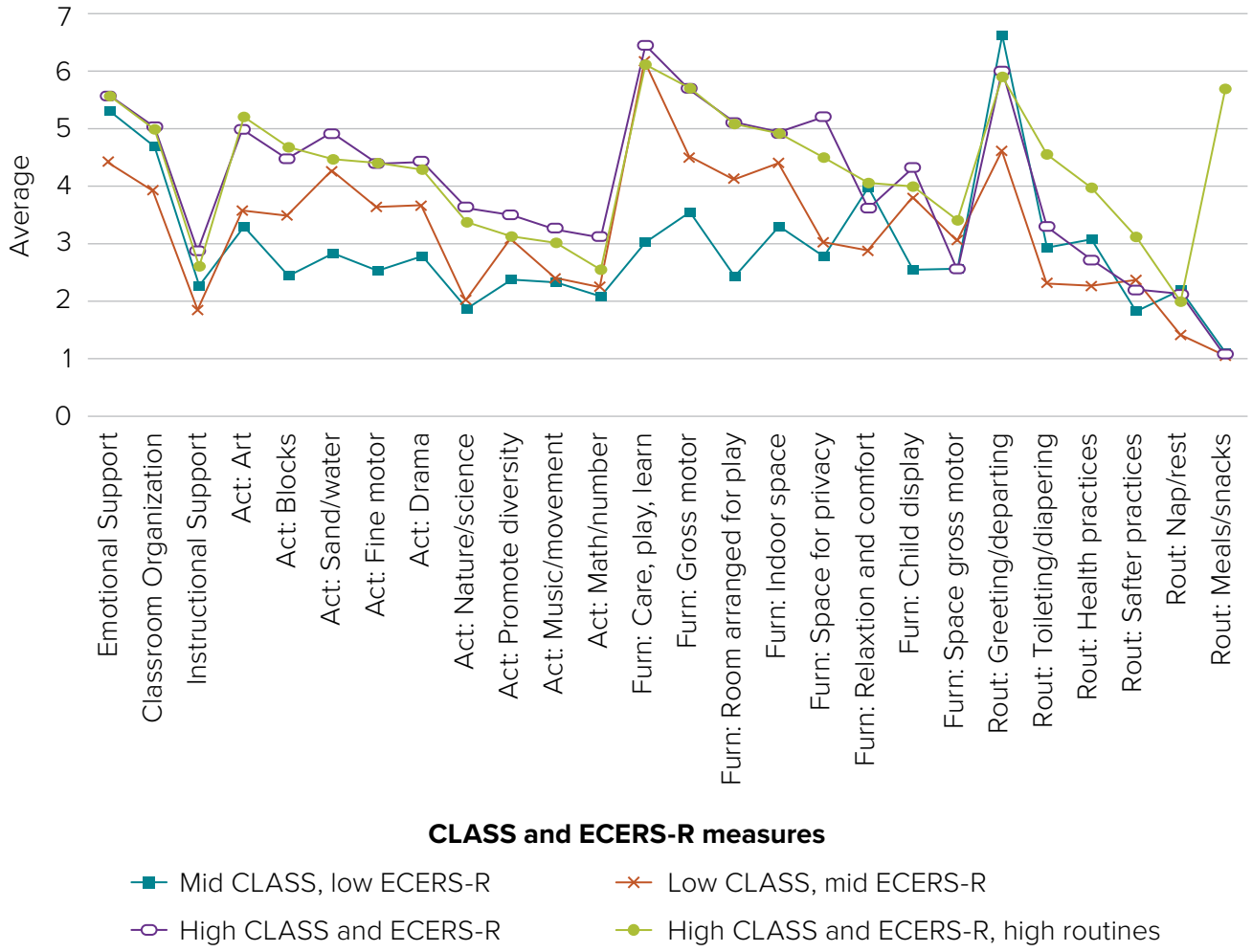
Number of classes	AIC	BIC	Entropy	Lo-Mendell-Rubin p-value*
1	23,729	23,920	–	–
2	23,155	23,451	0.84	0.053
3	22,774	23,175	0.97	0.728
4	22,527	23,033	0.92	0.744

Notes: *A lower p-value indicates the number of classes fit better than 1-less class.

As additional classes were fit, however, distinct profiles emerged that captured variability across the quality indicators in meaningful ways. Specifically, the 2-class solution split the sample along both CLASS and ECERS-R, the 3-class solution further split the sample by ECERS-R, and the 4-class solution split the sample by ECERS-R and CLASS in distinct areas. Therefore, we chose to pursue a 4-class solution that illustrated differences in both CLASS and ECERS-R. The 4-class solution also demonstrated high (greater than 0.94) average posterior probabilities suggesting adequate class separation, ranging from 0.94 to 0.98.

Profiles for the 4 classes are presented in Figure 5. The first class was labelled '*mid CLASS, low ECERS-R*'. Services in this class had mid-range scores for Emotional Support, Instructional Support and Classroom Organization, but this contrasted with the lowest scores on ECERS-R activities and furnishings, and low scores on routines. Services in the second class, labelled '*low CLASS, mid ECERS-R*', had the lowest scores on CLASS domains, and mid-range scores across ECERS-R activities, routines and furnishings. The third class was labelled '*high CLASS and ECERS-R*' due to high scores across both indices of quality. Finally, the fourth class was labelled, '*high CLASS and ECERS-R, high routines*' due to high scores across ECERS-R and CLASS, accentuated by the highest scores for ECERS-R routines and particularly high scores for the meals/snacks item.

Figure 5: Latent class analysis profile of ECERS-R and CLASS, 4-class solution



Notes: Act = activities; Furn = furnishings and Rout = routines.

Details on the service types most likely to be in each class profile are presented in Table 25. Notable insights include that nearly all (91.5%) FDC services were likely to be in the 'mid CLASS, low ECERS-R' class, and that they were the most frequent service associated with that profile (79.2%). Likewise, 57% of CBDC services were most likely to be in the 'low CLASS, mid ECERS-R' profile and constituted 69.8% of the class. Further, preschool services for 3-year olds were most likely to be in the 'high CLASS and ECERS-R' profile (71.1%), and formed a sizeable portion of the total (31.6%). Finally, preschool services were most frequently in the 'high CLASS and ECERS-R' profile, but with a less noticeable trend. Of interest, the 'high CLASS and ECERS-R, high routines' profile was distributed among service types. Nonetheless, this strong association between service type and class profile indicates that latent class analysis of ECERS-R and CLASS quality is unlikely to increase the variability in these indices of quality, and instead reflected service type.

As a sensitivity test, we ran the 4-class solution again but predicted service type instead of including it within the latent class model. This revealed approximately equal class profiles and these profiles were strongly predictive of FDC, CBDC and preschool (standard and 3-year old provision). Indeed, the average absolute difference in service type allocation to the latent classes was 0.9%, with a maximum difference of 5.1%. Thus, the identified variation across CLASS and ECERS-R was indicative of differences brought about in part by how service type interacts with the quality observed.

Table 25: Association between latent class profiles and ECEC service type

CLASS label	Proportion of latent class				Proportion of service type			
	FDC	K	K3YO	CBDC	FDC	K	K3YO	CBDC
Mid CLASS, low ECERS-R	**79.2%**	4.3%	0%	16.4%	**91.5%**	3.3%	0%	8.7%
Low CLASS, mid ECERS-R	0%	30.2%	0%	**69.8%**	0%	36.1%	0%	**57%**
High CLASS and ECERS-R	1.7%	40.4%	31.6%	26.3%	2.7%	43.4%	**71.1%**	19.3%
High CLASS and ECERS-R, high routines	7.9%	29.4%	22.8%	39.8%	6.7%	16.7%	27.1%	15.4%

Notes: FDC = Family day care; K = Preschool; K3YO = Preschool for 3-year olds; CBDC = Centre-based day care. Bold text indicates sizeable associations between class profile and service types.

Summary of latent class analysis: Quality profiles index service type

The latent class analysis of ECERS-R and CLASS indices within the E4Kids sample indicated that distinct profiles of quality were largely illustrative of service type differences or tended to divide quality indices along a gradient of low to high. This provides unique insight into how the quality indices may function differently across service types, and the potential need for tailored observations and measures for research. Additionally, however, it has implications for statistical analysis aiming to examine links between quality and outcomes. Specifically, additional variance, and thus predictive power, is unlikely to be found as the latent classes collapse quality indices while increasing co-linearity with service type. Therefore, adjusting for service type and then including the broader range of quality (for example, scores from 1 to 7 for CLASS domains) is more appropriate to understand how variation in quality is associated with variation in children's outcomes. This insight guides the analysis presented in the next section examining associations between CLASS and ECERS-R quality and gains in cognitive ability and achievement.

E4Kids: Association between cognitive ability and achievement and CLASS and ECERS-R

Sample description and overview

Descriptive statistics for child-level sample outcomes covariates are presented in Appendix Table A6. Briefly, the sample captured a broad range of socio-demographic and developmental variability. Additionally, Appendix Table A7 presents room-level descriptive statistics. These show higher scores on CLASS Emotional Support and Classroom Organization, and relatively lower scores for Instructional Support consistent with existing and international findings. There was also a diversity of services, with CBDC comprising the majority of rooms, but a sizeable number of FDC and preschool rooms present. Table 26 presents a summary of the analytical models, and the supplementary material contains the full model specifications (supplementary Tables S108 to S111).[6]

6 Additional results are available in a series of supplementary tables. Please contact AERO for a copy.

CLASS: Classroom Organization, Emotional Support and Instructional Support

Children who participated in services with higher Classroom Organization exhibited greater gains in visual matching (β = 0.86; effect size in standard deviations = 0.051), and this was consistent across adjusted and unadjusted models. Children who participated in services with higher Classroom Organization also had greater gains in verbal comprehension (β = 0.76; effect size = 0.051), understanding directions (β = 0.74; effect size = 0.045), and applied problems (β = 1.04; effect size = 0.04), while children in services with higher Emotional Support had greater gains in applied problems (β = 0.87; effect size = 0.034), but these were only statistically significant in the unadjusted models.

The non-statistically significant effects for the CLASS domains were all positive. Thus indicating, on balance, that the effects of these quality domains are likely positive, but the magnitude and direction cannot be inferred across all domains and outcomes with the present sample size and requires additional research.

ECERS-R: Furnishings, routines and activities

Children who participated in services with higher scores on activities exhibited greater gains in understanding directions (β = 1.3; effect size in standard deviations = 0.073), and this was consistent across adjusted and unadjusted models and the CLASS domain included in the model. Additionally, children who participated in services with higher scores on furnishing had larger increases in visual matching (β = 0.8; effect size = 0.051), though this was only statistically significant in the adjusted model.

Contrary to expectation, however, children who participated in services with higher scores on routines exhibited lower gains in visual matching (β = -0.56; effect size = -0.047) and understanding directions (β = -0.64; effect size = -0.055), and this was consistent across adjusted and unadjusted models and the CLASS domain included in the model. The model with ECERS-R routines was re-run including only routines as a sensitivity test. This further suggested the negative association between higher scores on routines and lower gains in cognitive ability and achievement, but the effects were not statistically significant. Thus, adjusting for other aspects of process and structural quality appears necessary to identify significant negative associations for routines. High and low scores on routines coexist within a range of scores on other quality measures. Statistical or design adjustments, therefore, seem a necessary condition for future research exploring the effects of routines.

The non-statistically significant effects for the ECERS-R activities, furnishings and routines domains were generally all positive, but several negative associations with wide confidence intervals were present. Thus, the effects of these quality domains were more uncertain and the magnitude and direction would be better inferred across ECERS-R domains and outcomes with a larger sample.

Table 26: Association between CLASS and ECERS-R measures of ECEC quality and children's gains in measures of cognitive ability and achievement

Quality index	Verbal comprehension		Concept formation		Visual matching		Understanding directions		Applied problems	
	Unadj	Adj	Unadj	Adj	Unadj	Adj	Unadj	Adj	Unadj	Adj
Classroom Organization	0.76 (0.16–1.35)*	0.56 (-0.01–1.13)	0.75 (-0.16–1.67)	0.33 (-0.53–1.19)	1.06 (0.4–1.72)**	0.86 (0.18–1.53)*	0.74 (0.04–1.44)*	0.23 (-0.43–0.9)	1.04 (0.18–1.9)*	0.86 (-0.03–1.74)
Furnishings	0.05 (-0.63–0.74)	0.18 (-0.46–0.82)	0.1 (-0.97–1.16)	0.3 (-0.7–1.3)	0.63 (-0.14–1.4)	0.71 (-0.07–1.48)	-0.25 (-1.08–0.58)	0.16 (-0.62–0.95)	0.28 (-0.69–1.25)	0.5 (-0.48–1.48)
Routines	-0.32 (-0.81–0.17)	-0.31 (-0.78–0.16)	0.28 (-0.46–1.02)	0.35 (-0.34–1.04)	-0.71 (-1.24–-0.18)**	-0.66 (-1.19–-0.13)*	-0.71 (-1.3–-0.13)*	-0.68 (-1.23–-0.12)*	-0.63 (-1.36–0.09)	-0.67 (-1.4–0.06)
Activities	0.1 (-0.61–0.82)	-0.03 (-0.74–0.68)	0.54 (-0.62–1.69)	0.03 (-1.09–1.14)	-0.42 (-1.25–0.41)	-0.39 (-1.26–0.48)	1.3 (0.41–2.18)**	1.3 (0.45–2.16)**	-0.27 (-1.35–0.81)	-0.25 (-1.38–0.89)
Emotional Support	0.59 (-0.02–1.21)	0.34 (-0.25–0.93)	0.81 (-0.13–1.74)	0.34 (-0.56–1.24)	0.66 (-0.02–1.35)	0.38 (-0.32–1.08)	0.57 (-0.13–1.28)	0.01 (-0.66–0.69)	0.87 (0.02–1.72)*	0.64 (-0.25–1.52)
Furnishings	0.1 (-0.58–0.79)	0.22 (-0.42–0.86)	0.13 (-0.93–1.19)	0.31 (-0.69–1.31)	0.71 (-0.07–1.49)	0.79 (0.01–1.57)*	-0.2 (-1.03–0.63)	0.19 (-0.59–0.97)	0.34 (-0.63–1.31)	0.56 (-0.42–1.53)
Routines	-0.28 (-0.77–0.21)	-0.27 (-0.75–0.2)	0.28 (-0.46–1.02)	0.35 (-0.33–1.04)	-0.64 (-1.17–-0.1)*	-0.59 (-1.13–-0.05)*	-0.68 (-1.27–-0.09)*	-0.64 (-1.19–-0.09)*	-0.59 (-1.31–0.13)	-0.64 (-1.37–0.09)
Activities	0.07 (-0.66–0.8)	-0.03 (-0.75–0.69)	0.45 (-0.71–1.62)	-0.01 (-1.14–1.12)	-0.44 (-1.3–0.42)	-0.37 (-1.26–0.52)	1.26 (0.37–2.15)**	1.33 (0.46–2.19)**	-0.33 (-1.42–0.76)	-0.28 (-1.43–0.87)
Instructional Support	0.51 (-0.06–1.08)	0.23 (-0.32–0.78)	0.62 (-0.25–1.49)	0.03 (-0.81–0.87)	0.48 (-0.15–1.11)	0.2 (-0.44–0.85)	0.58 (-0.11–1.27)	0.22 (-0.43–0.88)	0.33 (-0.5–1.16)	0.21 (-0.65–1.06)

Quality index	Verbal comprehension		Concept formation		Visual matching		Understanding directions		Applied problems	
	Unadj	Adj	Unadj	Adj	Unadj	Adj	Unadj	Adj	Unadj	Adj
Furnishings	0.1 (-0.58–0.79)	0.23 (-0.41–0.87)	0.14 (-0.92–1.21)	0.33 (-0.67–1.33)	0.72 (-0.07–1.5)	0.8 (0.02–1.59)*	-0.21 (-1.04–0.62)	0.18 (-0.6–0.96)	0.37 (-0.6–1.34)	0.59 (-0.38–1.57)
Routines	-0.26 (-0.75–0.23)	-0.26 (-0.73–0.22)	0.32 (-0.41–1.06)	0.4 (-0.28–1.08)	-0.6 (-1.14–-0.05)*	-0.56 (-1.1–-0.02)*	-0.67 (-1.25–-0.08)*	-0.67 (-1.22–-0.12)*	-0.5 (-1.22–0.23)	-0.58 (-1.31–0.16)
Activities	0.1 (-0.62–0.83)	0.01 (-0.7–0.72)	0.52 (-0.64–1.68)	0.06 (-1.06–1.17)	-0.39 (-1.25–0.47)	-0.32 (-1.21–0.56)	1.28 (0.39–2.17)**	1.31 (0.45–2.17)**	-0.2 (-1.29–0.88)	-0.18 (-1.32–0.97)
Routines only	-0.11 (-0.54–0.32)	-0.15 (-0.57–0.26)	0.58 (-0.05–1.22)	0.51 (-0.09–1.1)	-0.36 (-0.83–0.11)	-0.36 (-0.83–0.12)	-0.39 (-0.91–0.13)	-0.39 (-0.88–0.1)	-0.35 (-0.98–0.27)	-0.4 (-1.04–0.24)

Notes: P-value <0.05*, <0.01**, <0.001***. Unadj = unadjusted model; Adj = Adjusted model; ECERS-R = Early Childhood Environment Rating Scale – Revised; CLASS = Classroom Assessment Scoring System. Model fit to 50 imputed datasets. Coefficient and 95% Confidence Interval reported.

Summary of key findings: E4Kids

Extending the analytic strategy applied to the PLIDA dataset, the quality of ECEC was modelled to explain changes in child cognitive ability and achievement outcomes between 2 years to assess the 'value add' from participating in high quality ECEC for 1,969 children. Latent class analyses were also used to derive and evaluate quality types across CLASS and ECERS-R.

Key findings were:

- **Association between cognitive ability and achievement and CLASS and ECERS-R**
  - Features of process quality and structural quality were associated with gains in cognitive development, identifying 'value add' of ECEC. Children in services with higher Classroom Organization had greater improvement in WJ III outcomes (visual matching). Participation in services with higher ECERS-R activities (observed curriculum content) was also associated with greater improvement on the WJ III outcome understanding directions (listening ability, language development), and ECERS-R furnishings was associated with improved visual matching.
  - Some aspects of structural quality may not be beneficial for child outcomes. Children who participated in services with higher scores on ECERS-R routines (e.g., toileting, meals, sleep/rest practices) had lower gains on visual matching and understanding directions. This may be an artefact of the ECERS-R measurement.
- **Latent class analysis: CLASS and ECERS-R**
  - Latent classes of CLASS and ECERS-R mapped to service type. The latent class analysis delivered a 4-class solution. These quality profiles mapped closely to different types of provision (preschool, CBDC and FDC). The results likely reflect that different structural features (staffing, hours of operation) assessed through ECERS-R enable process quality, but may also map to selection effects into program type.

Overall implications

The results show that participation in higher-quality ECEC services, as assessed by CLASS and ECERS-R, was associated with improved gains in cognitive ability and achievement. This result is strengthened by adjusting for baseline outcomes and process and structural features of ECEC quality. These findings add to the body of international literature that supports investment in ECEC to deliver the highest possible quality. Important in these findings is the confirmation of the 'value add' of high quality ECEC for greater gains in cognitive achievement and ability.

Next steps

The E4Kids analyses of CLASS and ECERS-R quality and cognitive ability and achievement outcomes were strengthened by detailed, standardised assessments of ECEC service quality and child outcomes and the availability of baseline developmental data to allow a 'value add' model that examines the unique contribution of ECEC to children's development. However, future research could make use of multiple observations of quality to further reduce confounding or selection and examine the role of ECEC routines in child development in greater nuance.

References

Adams, R., & Wu, M. (Eds.). (2003). *Programme for international student assessment (PISA): PISA 2000 technical report*. OECD Publishing. https://doi.org/10.1787/9789264199521-en

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural equation modeling: A multidisciplinary Journal*, *21*(3), 329–341. https://doi.org/10.1080/10705511.2014.915181

Australian Bureau of Statistics. (2019). Person Level Integrated Data Asset (PLIDA) First five years: What makes a difference? ABS DataLab. Findings based on use of PLIDA data.

Australian Bureau of Statistics. (2021). *Data integration project register*. https://www.abs.gov.au/about/data-services/data-integration/data-integration-project-register

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Cloney, D., Nguyen, C., Adams, R. J., Tayler, C., Cleveland, G., & Thorpe, K. (2017). Psychometric properties of the Classroom Assessment Scoring System (Pre-K): Implications for measuring interaction quality in diverse early childhood settings. *Journal of Applied Measurement, 18*(3), 299–318.

Curby, T. W., Brock, L. L., & Hamre, B. K. (2013). Teachers' emotional support consistency predicts children's achievement gains and social skills. *Early Education & Development, 24*(3), 292–309. https://doi.org/10.1080/10409289.2012.665760

Green, K. M., & Stuart, E. A. (2014). Examining moderation analyses in propensity score methods: application to depression and substance use. *Journal of Consulting and Clinical Psychology, 82*(5), 773. https://doi.org/10.1037/a0036515

Greifer, N. (2023a). *WeightIt: Weighting for covariate balance in observational studies*. The Comprehensive R Archive Network. https://cran.r-project.org/package=WeightIt

Greifer, N. (2023b). *cobalt: Covariate balance tables and plots*. R package version 4.5.1. The Comprehensive R Archive Network. https://cran.r-project.org/package=cobalt

Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale*. Teachers College Press, Columbia University.

Holmes, T. H., & Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research, 11*(2), 213–218. https://doi.org/10.1016/0022-3999(67)90010-4

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. L., Walters, E. E., & Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine, 32*(6), 959–976. https://doi.org/10.1017/s0033291702006074

Li, F., & Li, F. (2019). Propensity score weighting for causal inference with multiple treatments. *The Annals of Applied Statistics, 13*(4), 2389–2415. https://doi.org/10.1214/19-aoas1282

Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software, 42*, 1–29. https://doi.org/10.18637/jss.v042.i10

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*(3), 767–778. https://doi.org/10.1093/biomet/88.3.767

Lumley, T. (2021). *svyVGAM: Design-based inference in vector generalised linear models*. The Comprehensive R Archive Network. https://cran.r-project.org/web/packages/svyVGAM/index.html

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008a). *Classroom Assessment Scoring System (CLASS): manual K–3.* Paul H. Brookes Publishing.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008b*). Classroom Assessment Scoring System (CLASS): manual Pre-K*. Paul H. Brookes Publishing.

Prior, M. R., Sanson, A. V., & Oberklaid, F. (1989). The Australian temperament project. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 537–554). John Wiley & Sons.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Sanson, A., Smart, D., Prior, M., Oberklaid, F., & Pedlow, R. (1994). The structure of temperament from age 3 to 7 years: Age, sex, and sociodemographic influences. *Merrill-Palmer Quarterly, 4*, 232–252.

Staton, S., Houen, S., Rankin, P., Beatton, T., Zheng, Z., Cooke, E., Irvine, S., & Thorpe, K. (2020). *Observing ECEC quality: Development of an observation strategy for authorised officers*. Queensland Government, Department of Education.

Tang, A., Rankin, P., Staton, S., & Thorpe, K. (2024). Access to high-quality early childhood education and care: Analysis of Australia's national integrated data. *Early Childhood Research Quarterly, 67*, 352–362. https://doi.org/10.1016/j.ecresq.2024.02.001

Tayler, C., Cloney, D., Adams, R., Ishimine, K., Thorpe, K., & Nguyen, T. K. C. (2016). Assessing the effectiveness of Australian early childhood education and care experiences: Study protocol. *BMC Public Health, 16*(1), 1–12. https://doi.org/10.1186/s12889-016-2985-1

Tayler, C., Ishimine, K., Cloney, D., Cleveland, G., & Thorpe, K. (2013). The quality of early childhood education and care services in Australia. *Australasian Journal of Early Childhood, 38*(2). https://doi.org/10.1177/183693911303800203

Thorpe, K., Rankin, P., Beatton, 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

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, 45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Riverside Publishing.

Appendix A: Tables

Table A1: Australian National Quality Standard for Early Education and Care – Quality Areas, standards and elements (2012 version)

Standard	Element	
Quality Area 1 – Educational Program and Practice		
1.1 An approved learning framework informs the development of a curriculum that enhances each child's learning and development.	1.1.1	Curriculum decision-making contributes to each child's learning and development outcomes in relation to their identity, connection with community, wellbeing, confidence as learners and effectiveness as communicators.
	1.1.2	Each child's current knowledge, ideas, culture, abilities and interests are the foundation of the program.
	1.1.3	The program, including routines, is organised in ways that maximise opportunities for each child's learning.
	1.1.4	The documentation about each child's program and progress is available to families.
	1.1.5	Every child is supported to participate in the program.
	1.1.6	Each child's agency is promoted, enabling them to make choices and decisions and influence events and their world.
1.2 Educators and co-ordinators are focused, active and reflective in designing and delivering the program for each child.	1.2.1	Each child's learning and development is assessed as part of an ongoing cycle of planning, documenting and evaluation.
	1.2.2	Educators respond to children's ideas and play and use intentional teaching to scaffold and extend each child's learning.
	1.2.3	Critical reflection on children's learning and development, both as individuals and in groups, is regularly used to implement the program.
Quality Area 2 – Children's Health and Safety		
2.1 Each child's health is promoted.	2.1.1	Each child's health needs are supported.
	2.1.2	Each child's comfort is provided for and there are appropriate opportunities to meet each child's need for sleep, rest and relaxation.
	2.1.3	Effective hygiene practices are promoted and implemented.
	2.1.4	Steps are taken to control the spread of infectious diseases and to manage injuries and illness, in accordance with recognised guidelines.

Standard	Element	
2.2 Healthy eating and physical activity are embedded in the program for children.	2.2.1	Healthy eating is promoted and food and drinks provided by the service are nutritious and appropriate for each child.
	2.2.2	Physical activity is promoted through planned and spontaneous experiences and is appropriate for each child.
2.3 Each child is protected.	2.3.1	Children are adequately supervised at all times.
	2.3.2	Every reasonable precaution is taken to protect children from harm and any hazard likely to cause injury.
	2.3.3	Plans to effectively manage incidents and emergencies are developed in consultation with relevant authorities, practised and implemented.
	2.3.4	Educators, co-ordinators and staff members are aware of their roles and responsibilities to respond to every child at risk of abuse or neglect.
Quality Area 3 – Physical Environment		
3.1 The design and location of the premises is appropriate for the operation of a service.	3.1.1	Outdoor and indoor spaces, buildings, furniture, equipment, facilities and resources are suitable for their purpose.
	3.1.2	Premises, furniture and equipment are safe, clean and well maintained.
	3.1.3	Facilities are designed or adapted to ensure access and participation by every child in the service and to allow flexible use, and interaction between indoor and outdoor space.
3.2 The environment is inclusive, promotes competence, independent exploration and learning through play.	3.2.1	Outdoor and indoor spaces are designed and organised to engage every child in quality experiences in both built and natural environments.
	3.2.2	Resources, materials and equipment are sufficient in number, organised in ways that ensure appropriate and effective implementation of the program and allow for multiple uses.
3.3 The service takes an active role in caring for its environment and contributes to a sustainable future.	3.3.1	Sustainable practices are embedded in service operations.
	3.3.2	Children are supported to become environmentally responsible and show respect for the environment.

Standard	Element	
Quality Area 4 – Staffing Arrangements		
4.1 Staffing arrangements enhance children's learning and development and ensure their safety and wellbeing.	4.1.1	Educator-to-child ratios and qualification requirements are maintained at all times.
4.2 Educators, co-ordinators and staff members are respectful and ethical.	4.2.1	Professional standards guide practice, interactions and relationships.
	4.2.2	Educators, co-ordinators and staff members work collaboratively and affirm, challenge, support and learn from each other to further develop their skills, to improve practice and relationships.
	4.2.3	Interactions convey mutual respect, equity and recognition of each other's strengths and skills.
Quality Area 5 – Relationships with children		
5.1 Respectful and equitable relationships are developed and maintained with each child.	5.1.1	Interactions with each child are warm, responsive and build trusting relationships.
	5.1.2	Every child is able to engage with educators in meaningful, open interactions that support the acquisition of skills for life and learning.
	5.1.3	Each child is supported to feel secure, confident and included.
5.2 Each child is supported to build and maintain sensitive and responsive relationships with other children and adults.	5.2.1	Each child is supported to work with, learn from and help others through collaborative learning opportunities.
	5.2.2	Each child is supported to manage their own behaviour, respond appropriately to the behaviour of others and communicate effectively to resolve conflicts.
	5.2.3	The dignity and rights of every child are maintained at all times.
Quality Area 6 – Collaborative partnerships with families and communities		
6.1 Respectful supportive relationships with families are developed and maintained.	6.1.1	There is an effective enrolment and orientation process for families.
	6.1.2	Families have opportunities to be involved in the service and contribute to service decisions.
	6.1.3	Current information about the service is available to families.

Standard	Element	
6.2 Families are supported in their parenting role and their values and beliefs about child rearing are respected.	6.2.1	The expertise of families is recognised and they share in decision-making about their child's learning and wellbeing.
	6.2.2	Current information is available to families about community services and resources to support parenting and family wellbeing.
6.3 The service collaborates with other organisations and service providers to enhance children's learning and wellbeing.	6.3.1	Links with relevant community and support agencies are established and maintained.
	6.3.2	Continuity of learning and transitions for each child are supported by sharing relevant information and clarifying responsibilities.
	6.3.3	Access to inclusion and support assistance is facilitated.
	6.3.4	The service builds relationships and engages with their local community.
Quality Area 7 – Leadership and service management		
7.1 Effective leadership promotes a positive organisational culture and builds a professional learning community.	7.1.1	Appropriate governance arrangements are in place to manage the service.
	7.1.2	The induction of educators, co-ordinators and staff members is comprehensive.
	7.1.3	Every effort is made to promote continuity of educators and co-ordinators at the service.
	7.1.4	Provision is made to ensure a suitably qualified and experienced educator or co-ordinator leads the development of the curriculum and ensures the establishment of clear goals and expectations for teaching and learning.
	7.1.5	Adults working with children and those engaged in management of the service or residing on the premises are fit and proper.
7.2 There is a commitment to continuous improvement.	7.2.1	A statement of philosophy is developed and guides all aspects of the service's operations.
	7.2.2	The performance of educators, co-ordinators and staff members is evaluated and individual development plans are in place to support performance improvement.
	7.2.3	An effective self-assessment and quality improvement process is in place.

Standard	Element	
7.3 Administrative systems enable the effective management of a quality service.	7.3.1	Records and information are stored appropriately to ensure confidentiality, are available from the service and are maintained in accordance with legislative requirements.
	7.3.2	Administrative systems are established and maintained to ensure the effective operation of the service.
	7.3.3	The Regulatory Authority is notified of any relevant changes to the operation of the service, of serious incidents and any complaints which allege a breach of legislation.
	7.3.4	Processes are in place to ensure that all grievances and complaints are addressed, investigated fairly and documented in a timely manner.
	7.3.5	Service practices are based on effectively documented policies and procedures that are available at the service and reviewed regularly.

Source: Education and Care ServicesNational Regulations (2011 SI 653) [Historical version for 1 September 2013 to 31 May 2014] by NSW Government, used under a CC BY 4.0 licence.

Table A2: Sample size for each AEDC domain for the 2016 and 2017 analyses examining NQS quality and AEDC domains

AEDC domain	Number of children 2016		Number of children 2017	
	Complete	Imputed	Complete	Imputed
Communication Skills and General Knowledge	89,971	125,605	82,339	116,335
Language and Cognitive Skills (school-based)	89,933	125,531	82,308	116,269
Emotional Maturity	89,689	125,170	82,066	115,922
Social Competence	89,977	125,611	82,349	116,342
Physical Health and Wellbeing	89,979	125,613	82,349	116,347

Table A3: Descriptive statistics for covariates in the 2016 sample examining NQS quality and AEDC domains

Variable	Number of children 2016		
	Complete	Imputed sample	Percentage missing imputed
Child gender			
Male	45,938 (51.1%)	64,084 (51%)	0%
Female	44,041 (48.9%)	61,529 (49%)	0%
Child language			
English	72,307 (80.4%)	99,549 (79.3%)	0%
Other than English	17,672 (19.6%)	26,064 (20.7%)	0%
Child age at first entry in childcare management system			
Older than 3 years	8,910 (9.9%)	14,176 (11.3%)	0%
Between 2 and 3 years	17,555 (19.5%)	26,046 (20.7%)	0%
Between 1 and 2 years	32,238 (35.8%)	44,657 (35.6%)	0%
1 or younger	31,276 (34.8%)	40,734 (32.4%)	0%
Aboriginal and/or Torres Strait Islander Order			
No	87,348 (97.1%)	119,821 (95.4%)	0%
Yes	2,631 (2.9%)	5,792 (4.6%)	0%
Tenure type of the house child was in for 2016 census			
Owned outright	6,091 (6.8%)	7,790 (7.3%)	15.20%
Own mortgage	56,680 (63%)	62,668 (58.9%)	15.20%
Other/shared equity/life tenure	275 (0.3%)	377 (0.4%)	15.20%
Renting	26,178 (29.1%)	3,4379 (32.3%)	15.20%
Rent free	755 (0.8%)	923 (0.9%)	15.20%
Not applicable	–	344 (0.3%)	15.20%
Combined carer income percentiles			
Greater than 80th ($101,855)	20,509 (22.8%)	22,364 (20%)	10.90%
60th ($72,186) to 80th ($101,855)	19,791 (22%)	22,367 (20%)	10.90%
40th ($47,888) to 60th ($72,186)	18,132 (20.2%)	22,385 (20%)	10.90%
20th ($27,487) to 40th ($47,888)	16,335 (18.2%)	22,357 (20%)	10.90%

Variable	Number of children 2016		
	Complete	Imputed sample	Percentage missing imputed
2nd ($2,088) to 20th ($27,487)	13,676 (15.2%)	20,170 (18%)	10.90%
Less than 2nd ($2,088)	1,536 (1.7%)	2,239 (2%)	10.90%
Amount of hours child participated in CBDC or FDC			
300 hours	4,925 (5.5%)	8,067 (6.4%)	0%
301 to 600 hours	8,522 (9.5%)	12,567 (10%)	0%
601 to 1200 hours	26,957 (30%)	37,553 (29.9%)	0%
1201 to 1800 hours	25,943 (28.8%)	35,056 (27.9%)	0%
1801 to 2400 hours	14,168 (15.7%)	19,244 (15.3%)	0%
2401 to 3000 hours	8,118 (9%)	11,278 (9%)	0%
Above 3000 hours	1,346 (1.5%)	1,848 (1.5%)	0%
Carer personal exertion income			
No	77,727 (86.4%)	94,031 (84%)	10.90%
Yes	12,252 (13.6%)	12,252 (13.6%)	10.90%
Carer age when child was born			
27 years or under	42,979 (47.8%)	58,422 (46.5%)	0%
27 to 35 years	9,969 (11.1%)	19,521 (15.5%)	0%
Over 35 years	37,031 (41.2%)	47,665 (37.9%)	0%
Carer using health services for chronic health condition			
No	74,398 (82.7%)	100,603 (82.6%)	3%
Yes	15,581 (17.3%)	21,243 (17.4%)	3%
Carer education			
Postgraduate degree	18,586 (20.7%)	21,325 (18.9%)	10.10%
Undergraduate degree	28,856 (32.1%)	33,471 (29.6%)	10.10%
Diploma	12,475 (13.9%)	15,361 (13.6%)	10.10%
Year 12	13,068 (14.5%)	13,068 (14.5%)	10.10%
Certification III or IV	13,068 (14.5%)	16,988 (15%)	10.10%
Year 11 or below	4,052 (4.5%)	8,492 (7.5%)	10.10%

Variable	Number of children 2016		
	Complete	Imputed sample	Percentage missing imputed
English proficiency			
English only	70,056 (77.9%)	88,267 (76.8%)	8.50%
Speaks English well or lower	19,923 (22.1%)	26,665 (23.2%)	8.50%
Carer used health services for mental health			
No	55,753 (62%)	77,165 (61.6%)	0.30%
Yes	34,226 (38%)	48,132 (38.4%)	0.30%
Carer migration			
Third-plus-generation migrant	53,352 (59.3%)	65,169 (57.6%)	10%
Second-generation migrant	14,129 (15.7%)	18,212 (16.1%)	10%
First-generation migrant	22,498 (25%)	29,710 (26.3%)	10%
Carer income support			
No	70,526 (78.4%)	88,481 (70.4%)	0%
Yes	19,453 (21.6%)	37,132 (29.6%)	0%
Number of adults and persons in household			
Only 1 adult, 2 persons in total	2,177 (2.4%)	3,032 (2.9%)	16.30%
Only 1 adult, 3 persons in total	3,018 (3.4%)	4,472 (4.3%)	16.30%
Only 1 adult, 4 persons in total	1,229 (1.4%)	2,167 (2.1%)	16.30%
Only 1 adult, 5 persons in total	433 (0.5%)	908 (0.9%)	16.30%
Only 1 adult, 6 or 7 persons in total	180 (0.2%)	417 (0.4%)	16.30%
More than 1 adult, 3 persons in total	12,398 (13.8%)	14,159 (13.5%)	16.30%
More than 1 adult, 4 persons in total	43,573 (48.4%)	48,045 (45.7%)	16.30%
More than 1 adult, 5 persons in total	17,762 (19.7%)	20,357 (19.4%)	16.30%
More than 1 adult, 6 persons in total	6,091 (6.8%)	7,396 (7%)	16.30%
More than 1 adult, 7 persons in total	1,810 (2%)	2,303 (2.2%)	16.30%
More than 1 adult, 8 persons in total	1,308 (1.5%)	1,847 (1.8%)	16.30%

Variable	Number of children 2016		
	Complete	Imputed sample	Percentage missing imputed
Service remoteness area classification			
Major cities	69,140 (76.8%)	95,168 (75.8%)	0%
Inner region	14,669 (16.3%)	21,172 (16.9%)	0%
Outer region	5,400 (6%)	8,089 (6.4%)	0%
Remote or very remote	770 (0.9%)	1,184 (0.9%)	0%
NQS rating year			
2012	4,838 (5.4%)	6,651 (5.3%)	0%
2013	25,002 (27.8%)	34,833 (27.7%)	0%
2014	20,494 (22.8%)	28,560 (22.7%)	0%
2015	24,036 (26.7%)	33,698 (26.8%)	0%
2016	15,609 (17.3%)	21,871 (17.4%)	0%
Overall quality rating			
Exceeding NQS/Excellent	30,230 (33.6%)	40,601 (32.3%)	0%
Meeting NQS	35,498 (39.5%)	49,903 (39.7%)	0%
Working Towards NQS	24,251 (27%)	35,109 (28%)	0%
Quality Area 1			
Exceeding NQS/Excellent	23,704 (26.3%)	31,823 (25.3%)	0%
Meeting NQS	48,331 (53.7%)	67,434 (53.7%)	0%
Working Towards NQS	17,944 (19.9%)	26,356 (21%)	0%
Quality Area 2			
Exceeding NQS/Excellent	24,437 (27.2%)	32,930 (26.2%)	0%
Meeting NQS	50,595 (56.2%)	70,782 (56.3%)	0%
Working Towards NQS	14,947 (16.6%)	21,901 (17.4%)	0%
Quality Area 3			
Exceeding NQS/Excellent	24,281 (27%)	24,281 (27%)	0%
Meeting NQS	24,281 (27%)	71,675 (57.1%)	0%
Working Towards NQS	14,454 (16.1%	21,496 (17.1%)	0%

Variable	Number of children 2016		
	Complete	Imputed sample	Percentage missing imputed
Quality Area 4			
Exceeding NQS/Excellent	31,060 (34.5%)	31,060 (34.5%)	0%
Meeting NQS	52,833 (58.7%)	74,458 (59.3%)	0%
Working Towards NQS	6,086 (6.8%)	9,202 (7.3%)	0%
Quality Area 5			
Exceeding NQS/Excellent	34,495 (38.3%)	46,585 (37.1%)	0%
Meeting NQS	48,104 (53.5%)	68,075 (54.2%)	0%
Working Towards NQS	7,380 (8.2%)	10,953 (8.7%)	0%
Quality Area 6			
Exceeding NQS/Excellent	39,128 (43.5%)	53,204 (42.4%)	0%
Meeting NQS	44,571 (49.5%)	62,863 (50%)	0%
Working Towards NQS	6,280 (7%)	9,546 (7.6%)	0%
Quality Area 7			
Exceeding NQS/Excellent	34,546 (38.4%)	46,669 (37.2%)	0%
Meeting NQS	42,850 (47.6%)	60,334 (48%)	0%
Working Towards NQS	12,583 (14%)	18,610 (14.8%)	0%
Sub-service type used for NQS quality rating			
CBDC	82,427 (91.6%)	114,192 (90.9%)	0%
FDC	7,552 (8.4%)	11,421 (9.1%)	0%

Notes: Sample size for Physical Health and Wellbeing domain presented as it had the largest sample size and sample size varies slightly for each AEDC domain (see Appendix Table A2).

Table A4: Descriptive statistics for covariates in the 2017 sample examining NQS quality and AEDC domains

Variable	Number of children 2017		
	Complete	Imputed sample	Percentage missing imputed
Child gender			
Male	42,121 (51.1%)	59,280 (51%)	0%
Female	40,228 (48.9%)	57,067 (49%)	0%
Child language			
English	65,121 (79.1%)	89,874 (77.2%)	0%
Other than English	17,228 (20.9%)	26,473 (22.8%)	0%
Child age at first entry in childcare management system			
Older than 3 years	9,855 (12%)	16,396 (14.1%)	0%
Between 2 and 3 years	14,916 (18.1%)	22,092 (19%)	0%
Between 1 and 2 years	28,570 (34.7%)	39,829 (34.2%)	0%
1 or younger	29,008 (35.2%)	38,030 (32.7%)	0%
Aboriginal and/or Torres Strait Islander Order			
No	79,932 (97.1%)	111,077 (95.5%)	0%
Yes	2,417 (2.9%)	5,270 (4.5%)	0%
Tenure type of the house child was in for 2016 census			
Owned outright	5,431 (6.6%)	6,996 (7.1%)	15.90%
Own mortgage	50,998 (61.9%)	56,480 (57.7%)	15.90%
Other/shared equity/life tenure	243 (0.3%)	343 (0.4%)	15.90%
Renting	24,988 (30.3%)	3,2935 (33.6%)	15.90%
Rent free	689 (0.8%)	839 (0.9%)	15.90%
Not applicable	–	290 (0.3%)	15.90%
Combined carer income percentiles			
Greater than 80th ($105,341)	18,855 (22.9%)	20,641 (20%)	11.40%
60th ($74,863) to 80th ($105,341)	18,044 (21.9%)	20,586 (20%)	11.40%
40th ($50,188) to 60th ($74,863)	16,545 (20.1%)	20,625 (20%)	11.40%
20th ($29,815) to 40th ($50,188)	14,973 (18.2%)	20,626 (20%)	11.40%

Variable	Number of children 2017		
	Complete	Imputed sample	Percentage missing imputed
2nd ($2,757) to 20th ($29,815)	12,544 (15.2%)	18,547 (18%)	11.40%
Less than 2nd ($2,757)	1,388 (1.7%)	2,059 (2%)	11.40%
Hours at service, presence of preschool flag in AEDC and attending preschool			
0 to 300 hours and flagged as attending preschool	9,210 (11.2%)	12,927 (11.1%)	0%
0 to 600 hours and flagged as not attending preschool or preschool flag was missing	1,499 (1.8%)	2,832 (2.4%)	0%
301 to 600 hours and flagged as attending preschool	6,212 (7.5%)	8,742 (7.5%)	0%
601 to 1,200 hours	20,841 (25.3%)	29,717 (25.5%)	0%
1,201 to 1,800 hours	23,036 (28%)	31,662 (27.2%)	0%
1,801 to 2,400 hours	13,238 (16.1%)	18,487 (15.9%)	0%
2,401 to 3,000 hours	7,193 (8.7%)	10,419 (9%)	0%
Above 3,000 hours	1,120 (1.4%)	1,561 (1.3%)	0%
Carer personal exertion income			
No	72,238 (87.7%)	88,361 (85.7%)	11.40%
Yes	10,111 (12.3%)	14,723 (14.3%)	11.40%
Carer age when child was born			
27 years or under	39,421 (47.9%)	54,184 (46.6%)	0%
27 to 35 years	9,336 (11.3%)	18,280 (15.7%)	0%
Over 35 years	33,592 (40.8%)	43,879 (37.7%)	0%
Carer using health services for chronic health condition			
No	65,332 (79.3%)	89,194 (79.2%)	2.30%
Yes	17,017 (20.7%)	23,402 (20.8%)	2.30%
Carer education			
Postgraduate degree	17,114 (20.8%)	19,661 (18.9%)	10.60%
Undergraduate degree	25,933 (31.5%)	30,268 (29.1%)	10.60%

Variable	Number of children 2017		
	Complete	Imputed sample	Percentage missing imputed
Diploma	11,386 (13.8%)	14,116 (13.6%)	10.60%
Year 12	12,219 (14.8%)	16,237 (15.6%)	10.60%
Certification III or IV	11,756 (14.3%)	15,501 (14.9%)	10.60%
Year 11 or below	3,941 (4.8%)	8,250 (7.9%)	10.60%
English proficiency			
English only	62,938 (76.4%)	79,318 (74.8%)	8.90%
Speaks English well or lower	19,411 (23.6%)	26,660 (25.2%)	8.90%
Carer used health services for mental health			
No	49,614 (60.2%)	70,025 (60.4%)	0.30%
Yes	32,735 (39.8%)	45,975 (39.6%)	0.30%
Carer migration			
Third-plus-generation migrant	47,738 (58%)	58,194 (55.8%)	10.40%
Second-generation migrant	12,893 (15.7%)	16,782 (16.1%)	10.40%
First-generation migrant	21,718 (26.4%)	29,278 (28.1%)	10.40%
Carer income support			
No	64,077 (77.8%)	80,951 (69.6%)	0%
Yes	18,272 (22.2%)	35,396 (30.4%)	0%
Number of adults and persons in household			
Only 1 adult, 2 persons in total	2,114 (2.6%)	2,883 (3%)	16.90%
Only 1 adult, 3 persons in total	2,855 (3.5%)	4,114 (4.3%)	16.90%
Only 1 adult, 4 persons in total	1,176 (1.4%)	2,030 (2.1%)	16.90%
Only 1 adult, 5 persons in total	410 (0.5%)	833 (0.9%)	16.90%
Only 1 adult, 6 or 7 persons in total	174 (0.2%)	413 (0.4%)	16.90%
More than 1 adult, 3 persons in total	11,745 (14.3%)	13,437 (13.9%)	16.90%
More than 1 adult, 4 persons in total	39,509 (48%)	43,778 (45.3%)	16.90%
More than 1 adult, 5 persons in total	15,662 (19%)	18,048 (18.7%)	16.90%
More than 1 adult, 6 persons in total	5,624 (6.8%)	6,906 (7.1%)	16.90%

Variable	Number of children 2017		
	Complete	Imputed sample	Percentage missing imputed
More than 1 adult, 7 persons in total	1,736 (2.1%)	2,285 (2.4%)	16.90%
More than 1 adult, 8 persons in total	1,344 (1.6%)	1,980 (2%)	16.90%
Service remoteness area classification			
Major cities	64,077 (77.8%)	89,878 (77.2%)	0%
Inner region	12,901 (15.7%)	18,500 (15.9%)	0%
Outer region	4,668 (5.7%)	6,934 (6%)	0%
Remote or very remote	703 (0.9%)	1,035 (0.9%)	0%
NQS rating year			
2012	1,975 (2.4%)	2,731 (2.3%)	0%
2013	17,143 (20.8%)	23,964 (20.6%)	0%
2014	15,398 (18.7%)	21,590 (18.6%)	0%
2015	19,284 (23.4%)	27,242 (23.4%)	0%
2016	17,692 (21.5%)	25,143 (21.6%)	0%
2017	10,857 (13.2%)	15,677 (13.5%)	0%
Overall quality rating			
Exceeding NQS/Excellent	30,006 (36.4%)	40,585 (34.9%)	0%
Meeting NQS	33,800 (41%)	47,862 (41.1%)	0%
Working Towards NQS	18,543 (22.5%)	27,900 (24%)	0%
Quality Area 1			
Exceeding NQS/Excellent	23,931 (29.1%)	32,359 (27.8%)	0%
Meeting NQS	44,389 (53.9%)	62,517 (53.7%)	0%
Working Towards NQS	14,029 (17%)	21,471 (18.5%)	0%
Quality Area 2			
Exceeding NQS/Excellent	23,556 (28.6%)	31,887 (27.4%)	0%
Meeting NQS	46,763 (56.8%)	66,015 (56.7%)	0%
Working Towards NQS	12,030 (14.6%)	18,445 (15.9%)	0%

Variable	Number of children 2017		
	Complete	Imputed sample	Percentage missing imputed
Quality Area 3			
Exceeding NQS/Excellent	24,307 (29.5%)	32,678 (28.1%)	0%
Meeting NQS	46,763 (56.8%)	65,992 (56.7%)	0%
Working Towards NQS	11,279 (13.7%)	17,677 (15.2%)	0%
Quality Area 4			
Exceeding NQS/Excellent	29,912 (36.3%)	40,612 (34.9%)	0%
Meeting NQS	47,216 (57.3%)	67,336 (57.9%)	0%
Working Towards NQS	5,221 (6.3%)	8,399 (7.2%)	0%
Quality Area 5			
Exceeding NQS/Excellent	33,052 (40.1%)	44,907 (38.6%)	0%
Meeting NQS	43,744 (53.1%)	62,802 (54%)	0%
Working Towards NQS	5,553 (6.7%)	8,638 (7.4%)	0%
Quality Area 6			
Exceeding NQS/Excellent	38,024 (46.2%)	51,942 (44.6%)	0%
Meeting NQS	39,102 (47.5%)	55,828 (48%)	0%
Working Towards NQS	5,223 (6.3%)	8,577 (7.4%)	0%
Quality Area 7			
Exceeding NQS/Excellent	33,099 (40.2%)	44,950 (38.6%)	0%
Meeting NQS	39,097 (47.5%)	55,564 (47.8%)	0%
Working Towards NQS	10,153 (12.3%)	15,833 (13.6%)	0%
Sub-service type used for NQS quality rating			
CBDC	75,901 (92.2%)	10,5853 (91%)	0%
FDC	6,448 (7.8%)	10,494 (9%)	0%

Notes: Sample size for Physical Health and Wellbeing domain presented as it had the largest sample size and sample size varies slightly for each AEDC domain (see Appendix Table A2).

Table A5: Sample characteristics of the rooms (n = 242) in the E4Kids latent class analysis

Variable	n	Mean or percentage	SD	Min	Max	Percentage missing
CLASS: Emotional Support	242	5.13	0.93	2.44	6.9	0%
CLASS: Instructional Support	242	2.37	0.98	1	5.73	0%
CLASS: Classroom Organization	242	4.59	0.92	1.92	6.33	0%
ECERS-R: Furnishings average	242	4.13	1.03	1.25	6.5	0%
ECERS-R: Routines average	242	2.89	1.29	1	7	0%
ECERS-R: Activities average	242	3.46	0.9	1.2	6.11	0%
Service type: Preschool	242	27.7%	–	–	–	0%
Service type: Preschool (3-year-old)	242	13.2%	–	–	–	0%
Service type: CBDC	242	40.5%	–	–	–	0%
Service type: FDC	242	18.6%	–	–	–	0%

Notes: ECERS-R averages presented, but individual items included in latent class analysis.

Table A6: Outcomes, covariates and sample characteristics of the children (n = 1,969) and caregivers in the E4Kids outcomes analysis

Variable	n	Mean or percentage	SD	Percentage missing
Child 2010: Verbal comprehension	1,742	451.3	13.6	11.5%
Child 2011: Verbal comprehension	1,655	460.8	14.1	15.9%
Child 2010: Concept formation	1,748	446.4	17.8	11.2%
Child 2011: Concept formation	1,675	457.8	19.4	14.9%
Child 2010: Visual matching	1,753	429.4	15	11%
Child 2011: Visual matching	1,697	445.4	16.3	13.8%
Child 2010: Understanding directions	1,750	454.3	16.1	11.1%
Child 2011: Understanding directions	1,675	464.7	14.4	14.9%
Child 2010: Applied problems	1,733	396.5	26.1	12%
Child 2011: Applied problems	1,685	415.6	21.8	14.4%

Variable	n	Mean or percentage	SD	Percentage missing
Child: Age years at assessment 2010	1,775	4.01	0.58	9.9%
Child: Age between 2010 and 2011	1,521	0.87	0.17	22.8%
Child: Female	1,969	48%	–	0%
Child: First Nations	1,956	2.7%	–	0.7%
Child: Born outside Australia	1,965	6.7%	–	0.2%
Child: Difficult temperament	1,633	2.02	0.66	17.1%
Child: Developmental delay	1,614	6.2%	–	18%
Child: English primary language (ref)	1,964	82.9%	–	0.3%
Child: English primary, speaks another language	1,964	11.9%	–	0.3%
Child: English not primary language	1,964	5.2%	–	0.3%
Stressful life event that seriously affected child	1,626	10.6%	–	17.4%
Caregiver Indigenous	1,930	1.4%	–	2%
Caregiver born outside Australia	1,936	23.5%	–	1.7%
Caregiver has health care benefits card	1,926	24.4%	–	2.2%
Caregiver education: Postgraduate (ref)	1,931	17.6%	–	1.9%
Caregiver: Bachelor degree	1,931	28.9%	–	1.9%
Caregiver: Diploma	1,931	13.2%	–	1.9%
Caregiver: Year 12 or TAFE	1,931	30.1%	–	1.9%
Caregiver: Year 10 or lower	1,931	10.3%	–	1.9%
Caregiver: Psychological distress	1,615	14.66	4.52	18%
Home learning environment	1,618	3.24	1.16	17.8%
Service type: Preschool	1,900	31.1%	–	3.5%
Service type: Preschool (3-year-old)	1,900	16.6%	–	3.5%
Service type: CBDC	1,900	47.3%	–	3.5%
Service type: FDC	1,900	3.2%	–	3.5%
Service type: After hours care	1,900	1.9%	–	3.5%

Table A7: Sample characteristics of the rooms (n = 249) in the E4Kids outcomes analysis

Variable	n	Mean or percentage	SD	Min	Max	Percentage missing
CLASS: Emotional Support	249	5.15	0.92	2.44	6.9	0%
CLASS: Instructional Support	249	2.38	0.97	1	5.73	0%
CLASS: Classroom Organization	249	4.6	0.93	1.92	6.5	0%
ECERS-R: Furnishings average	244	4.18	0.99	1.25	6.5	2%
ECERS-R: Routines average	244	2.92	1.3	1	7	2%
ECERS-R: Activities average	244	3.49	0.86	1.5	6.11	2%
Service type: Preschool	239	28%	–	–	–	4%
Service type: Preschool (3-year-old)	239	13.8%	–	–	–	4%
Service type: CBDC	239	41%	–	–	–	4%
Service type: FDC	239	15.9%	–	–	–	4%
Service type: After hours care	239	1.3%	–	–	–	4%

For more information visit
edresearch.edu.au

040624.CH.AA.V1