Scoping report

Australian Education Research Organisation

Early childhood data in Australia

Published June 2022

The Australian Education Research Organisation is Australia's national education evidence body, working towards excellent and equitable outcomes for all children and young people.

Acknowledgements

The Australian Education Research Organisation (AERO) acknowledges the Traditional Owners and Custodians of Country throughout Australia and their continuing connection to land, waters and community. We pay our respects to them and their cultures, and Elders past and present.

AERO acknowledges that this publication was made possible by the joint funding it receives from Commonwealth, state and territory governments.

Contributors

This document was prepared by Bridget Healey, Jen Jackson, Sandra Ofei-Ferri and Bryn Lampe. Our thanks to the Project Advisory Group (PAG) and Vulnerable and Disadvantaged Children Subgroup (VDC) of Early Childhood Policy Group (ECPG) for valuable insight. This report is AERO's view, not an endorsed view of either group.

Contents

Why using early childhood data matters

Australia has made tremendous progress in building a better early childhood education and care (ECEC) system to serve all Australian children and families. According to the Australian Early Development Census (AEDC), the majority of children were developmentally on track in 2021. Yet too many children are still developmentally vulnerable when they start school. Developmental vulnerability in the AEDC has increased across 4 out of 5 domains, and inequalities in outcomes according to children's family backgrounds still persist. Too little is known about what kinds of ECEC experiences can lift learning and development outcomes. Evidence is needed about what works for all children, and how to 'shift the dial' for the most vulnerable. Clearly, we can do more as a system to turn these trajectories around.

Substantial data already exists about ECEC services in Australia, including access, participation and quality. Large-scale data sets are held by governments (state and federal), research organisations, and sometimes by ECEC providers themselves. Yet these data are not fully utilised to investigate what makes the most difference for children's learning and development.

Better insights from large-scale data can help providers, peak bodies, policymakers, teachers and educators tell a powerful story about what works for children, what impact quality ECEC is making, and what conditions need to be in place to deliver on the promise of a strong start in life for every child. Policymakers and sector leaders agree that the top priority is understanding how ECEC can better serve children who are vulnerable or disadvantaged. This includes answers to the questions below.

- How are ECEC experiences contributing to children's outcomes? What difference does ECEC quality make?
- How do children's ECEC experiences and outcomes vary by cohort? What can the data reveal about the strengths of different children, families and communities, that could inform how support is directed?
- What trends in outcomes can be observed over time and how might these relate to changing conditions (e.g. the COVID-19 pandemic, the introduction of the National Quality Standard (NQS)?
- What are the patterns of children's participation in ECEC?
- Who is missing out on access to quality services?

This scoping report provides an initial step towards answering these questions. It maps the existing ECEC data landscape in Australia, and identifies opportunities for data to be better used to support children's learning and development. It builds on a substantial body of work already underway across governments and research in improving ECEC data use, which is continuing to evolve as the data architecture becomes more sophisticated. This scoping report signals promising projects, as well as enduring gaps that need to be addressed.

International contexts show what may be possible if Australia can improve how it uses large-scale data to inform policy and practice. Much of the evidence used to inform ECEC policy in Australia is currently drawn from large-scale international studies that connect participation in quality ECEC services to children's short- and long-term learning and development outcomes. By making better use of ECEC data, Australia can build a stronger homegrown evidence base to inform the work of teachers and educators across our diverse services and communities.

This report aims to guide and catalyse further work towards this long-term vision, including by the Australian Education Research Organisation (AERO) and its partners across the ECEC sector, policy and research.

Priorities for better ECEC data use

There is strong appetite across ECEC stakeholders for Australia to do more with data. AERO's consultations with policymakers and sector leaders revealed a high level of agreement about the importance of storytelling through large-scale data, the value of greater collaboration between the sector and policymakers, and the need for support for the sector to analyse and apply data insights to improve service and system support for priority cohorts. Our specific findings are listed below.

- **Better insights from existing ECEC data assets need to be accessible,** for policymakers and sector leaders to use in their work. Sector leaders and policymakers alike invest considerable resources in accessing data and extracting insights, but are limited in their capacity to gain a comprehensive overview across the entire ECEC system. A better ECEC data architecture, with insights shared across all stakeholders, would improve efficiency and effectiveness in how the ECEC system works.

- **Insights from data must be informed by understanding of the sector.** Sector leaders in particular reported concern about the misuse of data in public debate, and called for broad dissemination of guidance on generating and applying credible data insights in policy discussions. This reflects the complexity of the ECEC sector, and the need for any insights from data to be accompanied by explanation of their context (for example, patterns in data may differ across different ECEC service types).
- **Sector leaders and policymakers have a shared vision for better data.** Shared aspirations for a better Australian ECEC data architecture include:
    - data on children's learning and development outcomes and the factors that enable their learning and development
    - linkage of data on children's ECEC experiences with data on health, wellbeing and family services
    - longitudinal data linking children's early experiences with later outcomes.
- **Greater collaboration is needed between governments and the sector.** Sector representatives expressed a strong interest in, and appreciation of, data efforts at government level (especially in relation to children who are vulnerable or disadvantaged), and greater collaboration to connect system insights with insights from practice. This could include opportunities for ECEC organisations and governments to co-create data collection instruments that are relevant and useful for the sector. It may also involve working with communities, to ensure ECEC data reflects their priorities.
- **Data collection and analysis must be aligned across ECEC initiatives.** Representatives from several jurisdictions indicated that any proposed improvements to early childhood data architecture should align with the Preschool Reform Funding Agreement and any related reforms. Sector leaders agreed that alignment between data initiatives is essential for reducing the data collection burden on the sector, and ensuring data are used well.

Current state of ECEC data use

To address the priorities above, it is first necessary to understand the current ECEC data landscape in Australia. This section canvasses the key large-scale data sets most relevant to ECEC, and existing work for linking and analysing them.

Existing large-scale data sets

The Australian ECEC data landscape includes a number of valuable large-scale, regularly updated data sets that can be used to explore the relationship between ECEC experiences, their quality, and

children's learning and development outcomes. AERO's audit of existing data sets identified the following as most relevant to this goal:

- AEDC
- Child Care Subsidy System (CCSS)
- National Early Childhood Education and Care Collection
- NQS data.

A large body of data also exists for other services that support children's learning and development, including allied health, child protection and community services. This data can provide a fuller picture of the experiences and services that make a difference in the crucial early years, especially for vulnerable children.

Research organisations also hold significant longitudinal data that contribute vital causal evidence about the experiences that contribute to children's outcomes long-term. These important studies include E4Kids, the Longitudinal Study of Australian Children, and the Longitudinal Study of Indigenous Children.

Data integration and infrastructure efforts

Analysis of a single large-scale data set can be further enhanced by data linkage. Children's early learning and development are shaped by many factors, including family and kin, communities, services, environments and policies. Linkage approaches integrate data from a range of sources, to better reflect the reality of children's experiences across ECEC and other contexts.

Promising early childhood data integration and linkage efforts include those below.

- The Australian Bureau of Statistics (ABS) **Multi-Agency Data Integration Project (MADIP)** is an important effort to make continually updated, accessible, linked data assets available to researchers. Within MADIP, the Australian Government Department of Education, Skills and Employment's (DESE) First Five Years enduring data asset aims to build a comprehensive picture of early childhood development in Australia, positioning better data about children as a key component of the long-term National Data Infrastructure Strategy.
- The Australian Government's **National Collaborative Research Infrastructure Strategy (NCRIS)** is funding the development of a national Humanities, Arts and Social Sciences Research Data Commons. The Commons will integrate existing data infrastructure and help institutions to govern, share and link data ethically, enhancing research capacity in education and a range of other fields.
- The **National Child Health and Development Atlas** ('The Atlas') is a proof-of-concept project (scheduled for completion in June 2023). It aims to establish a public, online interface or 'dashboard' presenting metrics on the health and development of children nationwide. The Atlas is intended for public use, including by communities, local governments, social services providers and nonprofits. The project is supported by a range of research and government agencies.
- At state and territory level, several jurisdictions have also pursued large-scale efforts to connect data on early childhood development, to better understand how well ECEC and related services are supporting children and families; such as the **South Australian Early Childhood Data Project.**
- Linkage of existing large-scale data is already yielding insights into the enablers of children's learning and development. For example, analysis linking longitudinal data from the Effective Early Educational Experiences (E4Kids) study with National Assessment Program for Literacy and Numeracy (NAPLAN) data found that emotional support for children in ECEC was associated with better reading, writing and numeracy skills at age 8 (Thorpe and Staton 2016).

Measurement and reporting frameworks

Measurement and reporting frameworks provide an alternative way to analyse early learning and development across multiple data sets, even if the data sets are not formally linked for analysis. This is especially important for monitoring progress for children identified as vulnerable or disadvantaged, whose circumstances are best understood by examining multiple factors simultaneously. They are also useful for monitoring change over time, recognising that change in outcomes may require coordinated change across multiple contributing indicators.

Over time, a number of projects have aimed to establish shared frameworks for tracking and reporting children's progress against a range of indicators nationally. These are outlined below.

- **Early Childhood Development (ECD) Outcomes Framework** (2011, 2014) was an indicator-based reporting framework that was developed under the National Early Childhood Development Strategy to measure progress toward the Strategy's 2020 goals.

- The Nest is a framework for children and young people aged 0-24 years that uses six interconnected domains of wellbeing, and underpins the Report Card: The Wellbeing of Young Australians (last updated 2018).
- Enhancing measurement of child wellbeing (AIHW 2019) is a discussion paper exploring opportunities for improving Australia's national reporting of children's wellbeing.
- Australia's children is a report series (last updated 2022) that uses the seven domains of the Australian Institute for Health and Welfare (AIHW) 'people-centred data model' to examine the latest data related to the Children's Headline Indicators, a set of 19 indicators associated with children's health, development and wellbeing.
- Australian Children's Wellbeing Index provides a snapshot of how children are faring according to the domains of The Nest and UNICEF Australia's five goals for children.

More can be done to improve data integration and infrastructure

Despite the significant efforts above, more can be done to ensure that ECEC data are collected, shared and analysed as effectively as possible. AERO's consultations with ECEC sector leaders, policymakers and researchers found that many basic questions still cannot be answered.

The scenarios below illustrate the importance of linked data in understanding how ECEC and other services can support children and families, especially those who are vulnerable or disadvantaged. They are based on real-life scenarios identified in AERO's consultations (adapted for confidentiality reasons).

Scenarios like these clearly signal the potential for ECEC data to be better connected, and to be used more purposefully to explore the interrelated factors that affect children's learning and development. Barriers to data-sharing and analysis can best be overcome by a focus on all stakeholders' shared

Case study: Scenarios for better data use

An outer suburb of Melbourne has very low numbers of children participating in preschool. Data suggest that there are many refugee families in this area who have young children, but it is difficult to know how many children live in the area, and whether those who are not attending preschool are participating in other types of formal ECEC, or are missing out entirely. Analysis of linked ECEC data sets could help providers and governments understand whether children are missing out on quality ECEC, and work with the community to respond.

A remote community in NSW has seen a decline in the participation of Aboriginal children in the local preschool. Meanwhile, an Aboriginal Children's Service in a nearby community is experiencing growth in attendance, due to the expansion of its culturally responsive program for families. Because data are not linked between the two services, policymakers cannot tell whether the two trends are related.

A low-socioeconomic (SES) community in Tasmania wants to know whether local ECEC services are effectively supporting children's learning and development. They have no way of comparing data for children attending school-based preschools to data for children in other ECEC services, or for linking non-school data to school outcomes. Although they have NQS ratings for all ECEC services, they have limited insight into the qualifications of teachers and educators, or whether workforce supply issues are limiting services' ability to deliver quality programs.

interest in answering these questions, and commitment to doing better for all children.

The remainder of this report explores specific topics that could be explored through better-connected ECEC data, and more purposeful analysis driven by priorities of and for children, families and communities. Coordinating efforts across these topic areas will help shed light on the scenarios above, and many others.

Opportunities for data analysis

This section explores opportunities for using large-scale data to answer research questions where existing evidence in the Australian context remains limited. It provides examples of how answering these questions could support policymakers' and providers' goals; outlines existing data sets that could provide insights; and discusses associated challenges and opportunities for improving our understanding of what works for children in terms of ECEC provision.

01 Understanding what ECEC experiences make a difference to children's outcomes

Understanding the relationships between children's experiences in ECEC settings and their learning and development is at the core of ensuring that the ECEC sector works for all children. We need to get better at linking children's experiences in ECEC and their subsequent outcomes in nuanced, contextualised ways that take account of the many factors that influence children's development. This will help us understand how well the system is serving children (especially those identified as vulnerable and disadvantaged), the conditions under which children do well, and where to direct further support. It is also key to demonstrating the value of public investment in ECEC.

Key questions

- What is the relationship between children's participation in ECEC programs and their learning and development outcomes?
- What is the relationship between aspects of ECEC quality and children's learning and development outcomes?
- What trends in outcomes can be observed over time and how might these relate to changing conditions (e.g. the COVID-19 pandemic, the introduction of the NQS)?

Existing data

The MADIP First Five Years enduring linked data asset is an excellent resource for exploring these questions because it links a range of data sets, including information on participation in ECEC and other services, with the AEDC, which provides an outcome variable that is internationally recognised as a valid and reliable measure of children's development in the first year of school. As a continually updated data set, First Five Years also allows the examination of trends over time.

Challenges and opportunities

- The First Five Years data asset does not include information on preschool in stand-alone and school-based settings, meaning it lacks a complete picture of children's participation in ECEC. This presents problems for examining data in states and territories where preschool programs are primarily delivered in those settings (e.g. Tasmania, Western Australia). There is an opportunity for collaboration between governments and preschool providers to address this gap.
- MADIP First Five Years includes NQS ratings, but NQS and AEDC data are collected at differing time intervals, so relationships between the quality of the service in which children participated and their subsequent outcomes may be difficult to establish. There is an opportunity for researchers to explore statistical workarounds or to seek data that can more reliably establish the quality-to-outcomes relationship.
- Other measures of environmental and pedagogical quality (such as the Classroom Assessment and Scoring System, and the Early Childhood Environmental Rating Scale) can be useful complements to NQS data, to understand the subtle features of ECEC quality that contribute to changes in children's outcomes. Typically these quality rating instruments are resource intensive and are therefore not routinely collected in large-scale data sets. However, there may be value in replicating or building on past studies.
- Available data can reveal relationships and trends but for some research questions will not establish causation, meaning that the factors that impact most on learning remain unclear. There is an opportunity for researchers, providers and policymakers to establish clearer causal relationships between evidence-based ECEC practices and better outcomes for children and families.

02 Understanding how experiences of ECEC vary by cohort

Building on the questions above, there is an opportunity to disaggregate the data to explore what is working well for priority cohorts, in a way that focuses on their strengths and is closely connected to communities' own priorities. The ECPG's Vulnerable and Disadvantaged Children (VDC) Subgroup, chaired by DESE and comprising AERO, ACECQA and representatives from all state and territory departments of education, is working together to build the evidence base for high-impact programs and initiatives for vulnerable cohorts and those from disadvantaged backgrounds.

Key questions

- How does the relationship between ECEC experiences and learning and development outcomes vary according to child and family background (e.g. SES, language or cultural background, remoteness)?
- What are the strengths and capabilities of children and families from diverse backgrounds, that provide foundations for learning and development?

Existing data

The AEDC Multiple Strengths Indicator (MSI) (Gregory and Brinkman 2016) within the AEDC data set can be used to understand the strengths found among groups of children whom the data also identify as vulnerable. Using the MADIP First Five Years dataset, there is also an opportunity to explore commonalities in these children's ECEC experiences such as participation, service type and quality, to understand which factors enable positive outcomes and can be influenced by policy.

Challenges and opportunities

- The limitations of the MADIP First Five Years data asset, in excluding stand-alone and school-based preschools, is a particular barrier to understanding the experiences of vulnerable children and those from disadvantaged backgrounds, as some cohorts are less likely to access the Child Care Subsidy (CCS)-eligible services currently included in MADIP data.
- The equity dimensions of Australia's ECEC system are not well understood. The relationship between the SES of communities, the quality and availability of ECEC services, and outcomes for children warrants further investigation to address equity issues before they intensify.
- The current data available lack important contextual information about the home learning environment, and the strengths and interests of families and communities surrounding children in priority cohorts. This limits what inferences can be made about protective factors and enablers of learning and development for children identified as vulnerable or disadvantaged. There is an opportunity to elevate the visibility of these factors in discussions about data.

03 Understanding patterns of participation

Understanding patterns of children's ECEC participation is essential for exploring the relationship

between ECEC and children's outcomes. Longitudinal research has found stronger cognitive gains among Australian children who attended ECEC programs from infant/toddler age compared with their peers who attended preschool only. The study also found that greater hours of ECEC per week was also associated with better outcomes (Coley et al. 2015). Further work is required to understand how variations in ECEC quality, service type and hours/weeks of participation affect these relationships. It is also important to understand how these relationships might vary for different age groups, cohorts and communities, and what is required to ensure that all children can access ECEC services.

Key questions

- What proportion of Australian children are enrolled in preschool in the year before school (building on current national data collections)?
- How often do enrolled children actually attend preschool in the year before school? How do attendance patterns vary for priority cohorts?
- What patterns of ECEC participation do children follow before they start school, and how do these patterns vary by geographical location?
- How many families use multiple ECEC services, either for one child, or multiple children, and how does this relate to families' participation in paid work?

Existing data

Some data to answer these questions can be found in the National Early Childhood Education and Care Collection, which captures information on parental work, study and training activity for children attending CCS-eligible services.

Challenges and opportunities

- Preschool participation data have ongoing limitations. Due to difficulties in calculation, the proportion of children enrolled in preschool exceeds 100% in some jurisdictions; and the Preschool Education data set does not record attendance outside the reference week. Efforts underway to improve the quality and consistency of preschool data (including the Preschool Reform Agenda) provide an opportunity for a fuller picture of ECEC participation across service types and jurisdictions.
- MADIP and other systemic data sets do not include data about Aboriginal children's services or family support programs, which are important contributors to learning and development for vulnerable groups. There is an opportunity for researchers and policymakers to explore possible connections.

04 Understanding workforce supply, demand and development

Workforce development is the heart of quality ECEC, as well as the ongoing viability and accessibility of ECEC services. The National Children's Education and Care Workforce Strategy (ACECQA 2021) will mobilise stakeholders to advance a range of actions toward professional recognition, attraction and retention, leadership capability, wellbeing, qualifications and career pathways. These activities must be informed by data and evidence, both to achieve long-term goals for workforce development, and to ensure timely, focused responses to the immediate workforce challenges affecting ECEC services.

Key questions

- Are enough early childhood educators and teachers being educated and trained at all qualification levels, to keep up with demand for ECEC?
- What are the qualifications of educators and teachers in all types of ECEC services, and how do these relate to the quality of services?
- Among which communities, locations and service types are the shortages of qualified educators and teachers most acute?
- What are the contextual factors that affect recruitment, retention and development of teachers and educators, at service or community level?

Existing data

The qualifications of current teachers and educators are collected via the DESE National Workforce Census around every three years. The National Skills Commission collects data on vacancies and future growth in ECEC related occupations and skills. The National Quality Agenda IT System, administered and maintained by ACECQA, includes data on staffing waivers granted to approved providers by state and territory regulatory authorities, as well as whether approved providers and services are meeting regulatory requirements, including the NQS. ACECQA data confirm staffing waivers and whether minimum qualification requirements are met for services. Self-reported income is available through the ABS national census, although categorisation of educators is very unreliable. DESE collects information about qualification enrolments and completions for vocational

education and training (VET) and Higher Education levels. ECEC providers also hold more detailed data about their workforce, but there is not currently a forum for sharing this to gain system-wide insight.

Challenges and opportunities

- Data on workforce supply and demand are held by many organisations (e.g. training institutions, governments, providers), making it difficult for policymakers and sector leaders to gain line-of-sight across the sector. Further, existing data lack sufficient detail to identify the communities, locations, and service types where shortages of qualified educators are most acute. The National Children's Education and Care Workforce Strategy (2022-2031) commits to enhancing national data collection, analysis and strategic discussion as part of Focus Area 6 (Data and evidence). Actions include refreshing the existing National Workforce Census collection, exploring options for contemporary and comprehensive data collections, such as a live workforce database, and regular national workforce forums to facilitate ongoing monitoring and evaluation of the strategy and its actions. These activities may present an opportunity to advance access to more detailed data.
- Data on completions of early childhood qualifications in universities and VET are not currently systematically collated in a way that supports workforce planning.
- Workforce supply issues need to be closely connected to quality to ensure that ECEC workforce supply initiatives also help to lift quality and advance opportunities for professional growth. Exploring how to better analyse workforce and quality data could help to support this connection. There is an opportunity to consider these issues as the National Workforce Strategy moves to implementation and evaluation.

05 Understanding provision, access and quality

If children are to benefit from ECEC, both access and quality need to improve in tandem. Exploring the relationship between the geographic spread of ECEC services and SES will shed further light on equity in ECEC provision and children's experiences. Such analysis could also support providers and governments to understand where unmet demand may exist for quality services, and where services could collaborate to meet the needs of their communities.

Key questions

- Where are ECEC services located, and how do service types and quality levels vary by location (e.g. remoteness, community SES)?
- What is the anticipated demand for quality ECEC services in future, and how well is the sector positioned to fulfil that demand?
- Who, in broad terms, are the children and families missing out on access to quality ECEC services and why?

Existing data

The National Quality Agenda IT System, includes data on service location and type, maximum number of approved places, and quality as measured by the NQS. A range of sources – including the Census, birth registry, CCSS and Preschool Education – provide location data for children/families' place of residence. These data have been linked in MADIP and other large-scale projects. Researcher access to data at an adequate level of granularity may depend on the discretion of the data custodian. Given the right data, it might be possible to estimate whether service provision is meeting demand within a region.

Challenges and opportunities

- To estimate whether supply is truly meeting demand might require a significant number of assumptions about the choices a given family is likely to make. There is nevertheless an opportunity to explore how well existing data can answer the question – including through simple descriptive analysis – as a potential starting point for future modelling.
- Estimates of who is 'missing out' would be improved with linkages between the relative number of children within a geographic location and data on the services in which those families participate. Evidence suggests that most families travel short distances to ECEC (median 2.9 km) (Cloney 2016). There is an opportunity to make these linkages more visible to all stakeholders involved in making service provision decisions.

06 Understanding how services can stack to deliver better outcomes

Achieving better outcomes for children cannot be done by ECEC alone; it requires the collaboration of multiple professionals and a range of services and supports around the child and their family.

Key questions

- What combination of ECEC and wraparound services and programs produces optimal outcomes for children and families, especially priority cohorts?
- How well do data and insights from integrated approaches to early childhood development in specific communities translate to scale?

Existing data

A number of major data reporting initiatives, such as the People-Centred Data Model (AIHW 2020), bring together the range of services that affect outcomes in early childhood, including health, housing, justice and safety. Research projects such as Restacking The Odds at the Centre for Community Child Health have also begun exploring how interventions can be 'stacked' to deliver better outcomes for children and families (Molloy et al. 2019). Many place-based integrated services also demonstrate how ECEC and other services for children and families can be connected at the local level through co-location and collaboration. Data can be a driving force for coalescing such collaborations around shared goals (see, for example, Harris 2018).

Challenges and opportunities

- Data linkage is a complex process that requires sophisticated data architecture and negotiation among the parties involved. The South Australia Early Childhood Data Project found that the success of their multi-agency data linkage – and the translation of its lessons to policy – depended on data custodians' willingness to commit extensive time to the collaboration. The ABS's MADIP initiative has succeeded in linking a range of child-centred data sets held by Australian Government departments, and it is possible to commission new linked data assets with the permission and cooperation of the relevant data custodians. Adding data on service provision collected by states and territories could shed new light on the ways in which children are benefiting from access to multiple government programs.
- Several projects exploring the integration of data across jurisdictions and service systems may provide useful demonstrations of the value of large-scale data linkage for revealing how interventions interact (e.g., The Atlas, NCRIS). Many of the relevant data sets on children's participation in ECEC and other programs are held at different levels of government. While the data infrastructure required for these initiatives presents technical challenges, the negotiation of data sharing, usage and governance presents the greater challenge. For these projects to work, they must navigate the interests of multiple levels of government and service systems, and ensure that the data are governed and safeguarded by a trusted, independent party.

Working together to create change

The early years are a critical period for a child's learning and development, during which the right support can shape their long-term life trajectory. Seizing that precious opportunity requires information on which children have access to quality ECEC, and how well services are supporting children across differences in ability, language, culture, location and wealth.

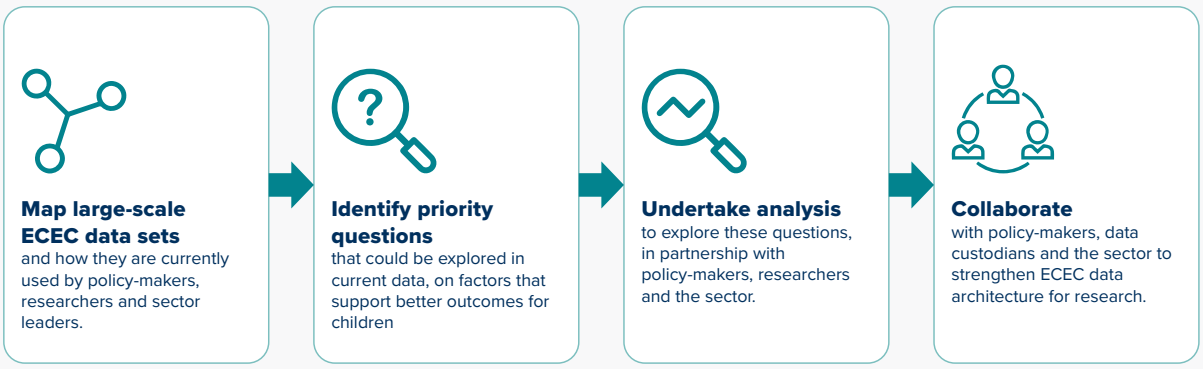
This scoping report has described the role that large-scale integrated data can play in ensuring that Australia's ECEC system delivers the best possible support to all children, especially those identified as vulnerable or disadvantaged. The enthusiasm for this task from the ECEC sector and policymakers is evidence of its value. At the same time, the breadth of issues that require exploration, and the breadth of data available to explore them, is also evidence of its challenge and complexity.

AERO is uniquely positioned to work in partnership with ECEC stakeholders to connect and catalyse improvement in how Australian ECEC data are used. The Productivity Commission's 2016 report on the National Education Evidence Base detailed the challenges in Australia's early childhood data architecture; including the quality, comparability, granularity and localisation of early childhood data. The Commission recommended a new national evidence institute to ensure that both school and ECEC data are accessible and used to inform policy and practice (Productivity Commission 2016). AERO was established to fulfil this role.

Partnerships are key to better use of ECEC data, to support and amplify data linkage and analysis efforts that are already underway. Mapping these efforts in this report, and identifying priority questions for analysis, is the first stage of AERO's 'early childhood data project' (see Figure 1). Next steps will involve analysing the conditions that boost early learning and development, sharing that analysis with sector and policy stakeholders, and collaborating to strengthen the ECEC data architecture.

This report also aims to inspire others to pursue opportunities to use ECEC data, either by accessing insights from existing initiatives, or undertaking their own analysis against priority topics. Coordinating ECEC data use around this shared research agenda will help insights and discoveries to be better connected and shared, building a stronger evidence base to benefit everyone in the sector.

Figure 1: 'Early childhood data project' plan



References

ACECQA (The Australian Children's Education & Care Quality Authority) (2021) 'Shaping Our Future: A ten-year strategy to ensure a sustainable, high-quality children's education and care workforce 2022-2031. National Children's Education and Care Workforce Strategy', Education Services Australia, Accessed May 2022.

AIHW (Australian Institute of Health and Welfare) (2020) 'Australia's Children,' catalogue number CWS 69, AIHW, Australian Government, accessed 5 May 2022.

Cloney D (2016) 'Variations in the Availability and Quality of Early Childhood Education and Care by Socioeconomic Status of Neighborhoods', Early education and development (1040-9289), 27(3): 384.

Coley R L, Lombardi C M and Sims J (2015) 'Long-term implications of early education and care programs for Australian children', Journal of Educational Psychology, 107(1):284–299, doi.org/10.1037/a0037456.

Gregory T and Brinkman S (2016) 'Exploring two new indices for the Australian Early Development Census (AEDC) programme: The Multiple Challenge and Multiple Strength Indicators', Telethon Kids Institute, Adelaide.

Halle T, Hair E, Wandner L, McNamara M and Chien N (2012) 'Predictors and outcomes of early vs. later English language proficiency among English language learners', Early Childhood Research Quarterly, 27(1):1–20, doi:10.1016/j.ecresq.2011.07.004.

Harris G (2018) 'Communities for Children Logan. AEDC National Conference Presentation', accessed 27 April 2022.

Ho C (2019) 'Ethnic divides in schooling', In a Class of Their Own Part 4, Centre for Policy Development, accessed 16 February 2022.

Molloy C, O'Connor M, Guo S, Lin C, Harrop C, Perini N and Goldfeld S (2019) 'Potential of 'stacking' early childhood interventions to reduce inequities in learning outcomes', Journal of Epidemiology and Community Health, 73(12):1078-1086, doi:10.1136/jech-2019-212282.

Productivity Commission (2016) 'National Education Evidence Base, Report No. 80', Productivity Commission, Canberra.

Thorpe K and Staton S (2016) Relationship quality in ECEC predicts NAPLAN outcomes, accessed 27 April 2022. https://issr.uq.edu.au/article/2018/04/relationship-quality-ecec-predicts-naplan-outcomes

Appendix 1. Summary of consultations

The success of the 'early childhood data project' depends on collaboration from a broad coalition. While stakeholders expressed a diversity of views and experiences, there was clear consensus that AERO's project is responding to an urgent need for analysis of the experiences and outcomes of vulnerable children and those from disadvantaged backgrounds, within the service system, and enablers of learning and development that can inform evidence-based responses.

We will collaborate at each stage of the project with ECEC sector leaders, policymakers and researchers to ensure we are investigating questions that are relevant to our stakeholders and address the most pressing needs. We will seek to communicate findings in ways that are timely and readily accessible and join up efforts aimed at building a stronger ECEC data architecture, working in ways that are complementary and accelerate progress towards the application of evidence in policy and practice.

Below is a summary of themes AERO heard from policy and sector representatives, and how AERO will respond to ensure the 'early childhood data project' adds practical value to their work to advance better outcomes for children through quality ECEC.

Feedback	AERO's project...
Supporting vulnerable or disadvantaged children to thrive was a strong priority of all stakeholders. The VDC will conduct a stocktake on available evidence to identify gaps in existing research and evidence in the context of the VDC cohort.	AERO will review Australian and international research evidence related to enablers of learning and development and protective factors for children identified as vulnerable or disadvantaged, to inform the Subgroup's stocktake of existing evidence/gaps identified across jurisdictions for what works for the VDC cohort.
Priorities regarding research questions centre on the nature of the relationship between ECEC experiences and outcomes for children, and that relationship varies by cohort (with a focus on priority cohorts), by ECEC dosage, and by ECEC quality.	AERO will review and prioritise research questions with consideration for their: • relevance of the questions to AERO's research agenda and to stakeholder interests • accessibility of data for the analysis • feasibility of the analysis within project timelines • uniqueness of the analysis, prioritising analysis not underway elsewhere • strategic value of the analysis.
Timely availability of credible data insights. Some peaks and providers fund their own data analysis, but greater access to up-to-date data insights would be welcome. As well as guidance for the sector on applying credible data insights.	AERO will test messaging for any public release of our analysis with sector representatives on the Project Advisory Group, and seek their feedback on what kind of sector guidance would be most useful to mitigate risk of misinterpretation.
Linked, child-centred data across setting types, wrap-around services, and into school.	AERO will work with jurisdictions and the ABS to explore how the full National Early Childhood Education and Care Collection could be included in a new linked data asset through MADIP to allow analysis of children's experiences in a wider range of approved early childhood education and care settings.

Appendix 2. Enablers of learning and development for children with a language background other than English

This appendix outlines the first analysis that AERO will undertake, in partnership with the First Five Years project, to explore factors affecting learning and development for children in a priority cohort. It represents the first of a series of data analyses to be undertaken as part of AERO's 'early childhood data project'.

What is the problem?

Early years development has a significant impact on later life outcomes. Analysis of Australian Early Development Census (AEDC) and NAPLAN data illustrates that the AEDC predicts NAPLAN results nine years later. This highlights the importance of focusing on areas of concern in early childhood education and care (ECEC) to prevent early issues compounding in later life.

The AEDC 2021 National Report identified children from a language background other than English (LBOTE) who are not proficient in English as a cohort of concern because they were found to be developmentally vulnerable across all five AEDC domains (physical health and wellbeing, social competence, emotional maturity, language and cognitive skills, communication skills and general knowledge). Specifically, findings from the report show:

- LBOTE children who are not proficient in English are more likely to be developmentally vulnerable by the time they start school than LBOTE children who are proficient in English
- the developmental vulnerability gaps between LBOTE children who are proficient in English and those who are not have been increasing since 2009
- significant and widening gaps in vulnerability between LBOTE children proficient in English and those not proficient in English is not limited to language skills domain. For example:
    - in 2021, 24.6% of LBOTE children not proficient in English were developmentally vulnerable in the emotional maturity domain
    - this was nearly four times higher than the 5.8% of LBOTE students who were proficient in English
    - Of 7,586 LBOTE children who were reported as not proficient in English, 4,590 (60.5%) were vulnerable in two or more domains, up from 3,830 (58%) in 2012.

Research suggests that the home environment and proficiency in home language influence whether LBOTE children who are not proficient in English catch up to their LBOTE and non-LBOTE peers in later years (see for example, Halle et al. 2012; Ho 2019). However, the AEDC does not capture information about the home environment.

AERO is looking to understand what can be done in the years before formal schooling to enable learning and development for LBOTE children, including those who are not proficient in English. We are especially interested in the relationship between ECEC experiences and learning and development outcomes for this diverse cohort of children.

Why does it matter?

Understanding the problem may help identify potential solutions

A better understanding of the problem and new insights from data may help us to identify what to do about it and inform ECEC policy and practice decision-making.

Investigating demographic differences within the LBOTE cohort, and how these differences relate to vulnerability, could help explain changes in reported vulnerability.

Uncovering risk factors or protective factors for the LBOTE cohort, such as home learning environment, SES, or participation in ECEC, may help identify ways policy, planning and practice could better support to these children and their families.

The evidence-base for developmental vulnerability in the LBOTE cohort is limited

There is currently no research on why developmental vulnerability may be growing among LBOTE children not proficient in English, or how demographic characteristics and schooling outcomes differ for LBOTE children who are and are not proficient in

English upon entry to school. It is also unclear whether proficiency in English for the group of developmentally vulnerable LBOTE students is associated with better future outcomes.

The analysis will add to the limited evidence base on the enablers of learning and development for LBOTE children in Australia who are not proficient in English and investigate the widening developmental vulnerability gap within the LBOTE cohort.

What will AERO do?

AERO will analyse the MADIP First Five Years data asset

The MADIP First Five Years enduring data asset integrates the Australian Early Development Census (years 2015 and 2018) with data from:

- Child Care Subsidy System data
- ACECQA's National Quality Standard data
- the 2016 Census
- Personal Income Tax submissions
- the Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS)
- other rich administrative data.

These data sets augment the detail of the AEDC with important context about children's home environments.

The analysis will explore ECEC experiences for children from LBOTE, including those who are not proficient in English, and their relationship to learning and development outcomes.

Underpinning research questions include the below.

- What factors influence learning and development in LBOTE children (including those who are not proficient in English)?
- What distinguishes LBOTE children with high levels of vulnerability, and those without?
- To what extent does ECEC (including service type) influence the learning and development of LBOTE children?

Findings will be shared with sector leaders and policymakers

Outputs from this work will aim to inform the ECEC sector and policy and practice decision-making. Key ECEC sector leaders and policymakers will be invited to provide feedback throughout the project.

How to cite

When referencing this publication, please attribute it as:

Australian Education Research Organisation (AERO) Ltd. (2022), Early childhood data in Australia, edresearch.edu.au

Further information

AERO produces resources to support the use of high-quality research. Explore these at: edresearch.edu.au/resources

Australian Education Research Organisation

edresearch.edu.au