



# **The Overcoming Disadvantage in Early Childhood Study**

**Evaluation of the Australian Literacy and Numeracy Foundation's Early Language and Literacy Program**

**Final Report**

**AUSTRALIAN  
LITERACY &  
NUMERACY  
FOUNDATION**

### **Acknowledgement of Country**

ALNF acknowledges the traditional custodians of the Country where we live, learn and work; and pay respect to them and their Elders – both past and present.

We are committed to improving the language, literacy and learning achievements of all children.

We acknowledge the contributions of Aboriginal and Torres Strait Islander people and non-Indigenous Australians in empowering all children to write their own futures.

### **Australian Literacy & Numeracy Foundation Ltd**

ALNF is a registered national charity dedicated to raising language, literacy and numeracy standards in Australia. ALNF's programs aim to create a more inclusive Australia, unlocking potential through the power of language, literacy and learning. It works in partnership with First Nations, refugee and other vulnerable Australians to address the language and literacy learning needs of individuals and communities, giving them a better chance to succeed.

### **Acknowledgements**

This seminal work was made possible with a major grant from the Eureka Benevolent Foundation and with the support of the CAGES Foundation. We thank them for their visionary support. They understand the vital role that oral language and early literacy enrichment in the years prior to school plays in setting children up for positive educational experiences, and the myriad of life opportunities which flow from that. They also know that disadvantaged children are the least likely to receive support in their language and pre-literacy development.

Australian Literacy & Numeracy Foundation Limited  
Level 1, Unit B, 110 McEvoy Street  
Alexandria NSW 2015  
Australia  
Phone: (02) 9362 3388  
ABN: 27 098 802 094

[www.alnf.org](http://www.alnf.org)

© The Australian Literacy and Numeracy Foundation 2023

# AUSTRALIAN LITERACY & NUMERACY FOUNDATION

August 2023

Between a child's first 1,000 days and the start of school, children engage in an intense period of curiosity and development. During these two years, children require adult-guided learning in order to effectively build oral language and early literacy skills in order to be ready for later formal classroom learning. This doesn't happen automatically.

Oral language and early literacy skills are the vital multipliers that underpin access to education, social, economic and other life opportunities. Australian and international research consistently attest to the importance of these domains.

To effectively address educational disparities, oral language and early literacy development must be a priority before children start school.

ALNF is pleased to present our contribution to national efforts for systemic educational change – with its Early Language and Literacy (**EL&L**) Program and the Early Language and Literacy Development Index (**ELLDI**).

We invite you to read the *Overcoming Disadvantage in Early Childhood Study: evaluation of the Australian Literacy and Numeracy Foundation's (ALNF) Early Language and Literacy Program Final Report (ODEC Report)*.

It confirms that the EL&L Program has closed the oral language and early literacy learning gap for children in marginalised communities in the year before school (p5).

It also introduces a world-first scale for 2- to 5-year-olds called the Early Language and Literacy Development Index (**ELLDI**). Uniquely, it can be deployed by local educators in early childhood education and care (**ECEC**) settings and in the early years of primary (it also extends to 8-year-olds) as a formative assessment to provide powerful insights to inform reflective practice.

The EL&L Program, accompanied by the ELLDI, constitutes evidence-informed practice in the early years. It promotes high quality language and pre-literacy engagement with children and can close the educational gap, particularly in communities where the vulnerabilities are most widespread.

The EL&L Program's ultimate goal is sustainable, community-driven quality language and literacy teaching and learning flourishing in Australia (with international scope) towards equitable opportunities for all.

In order to achieve sustained change, the EL&L Program incorporates a place-based and community-led approach, empowering local adults to break the cycle of entrenched oral language and early literacy inequity.

It generates an opportunity to connect services in health and education, and leverages existing infrastructure and workforces.

EL&L and the ELLDI, in the hands of local educators and other service providers, informs and strengthens their everyday work with children as well as conversations with parents/carers and families.

Local adults undertake evidence-informed training and mentoring<sup>1</sup> to learn strategies to support children's oral language and early literacy growth, equipped with robust information via the ELLDI feedback about each child's growth trajectory.

The ODEC Report shows that EL&L with the ELLDI is a powerful program that maximises children's oral language and early literacy growth.

Positive strategic change is achievable.

Yours sincerely,



Professor Tom Calma, AO FAA FASSA  
ALNF Co-Chair

FANZSOG FAA FASSA  
Hon DLitt CDU  
Hon DSc Curtin  
Hon DUniv Flin  
Hon DUniv UniSA  
Senior Australian of the Year 2023



Mary-Ruth Mendel CPSP  
ALNF Co-Founder, Co-Chair and Executive Director  
Speech and Language Pathologist

---

<sup>1</sup> Via the nationally accredited 11027NAT Certificate IV in Early Language and Literacy course.



## BRIEF OVERVIEW AND RESPONSE TO THE ODEC REPORT

### OBJECTIVES

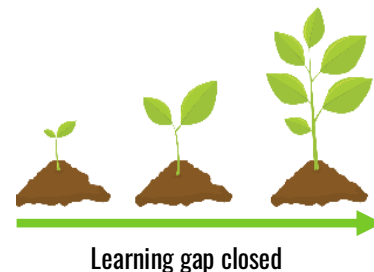
ALNF commissioned the Australian Council for Educational Research (ACER) to conduct the Overcoming Disadvantage in Early Childhood (ODEC) Study in order to achieve two core objectives (p4):

1. to determine the effect of the Early Language and Literacy (EL&L) Program on the development of language and literacy skills in preschool aged children; and
2. to develop a new formative measure of oral language and early literacy (the Early Language and Literacy Development Index (ELLDI)) which could be used by educators directly with children to inform the design of teaching and learning activities which are responsive to children's language and literacy growth needs. This did not previously exist.

### FINDINGS

The ODEC Final Report confirms:

1. **The EL&L Program demonstrates the capacity to close the learning gap in the twelve months before school** for children in highly disadvantaged populations. (pp5, 9, 117)
2. **The ELLDI:**
  - has been proven to be statistically reliable and valid (p4)
  - **robustly captures skills across 7 language and literacy strands** (p4).
  - **is internationally unique** as a formative assessment with a progression scale in oral language and early literacy in the hands of educators for 2- to 5-year-olds (p1) and further extends to 8-year-olds as a progression scale that gives visibility of children's growth from prior-to-school into the first three years of schooling.
  - provides valid and specific insights into individual children's learning growth with **meaningful, accessible recommendations for educational practice**
  - allows for the mapping of language and literacy components such as those identified in the science of reading, e.g. phonemic awareness, onto a scale of difficulty. This is the **essential developmental precursor to successful phonics instruction.**



## OPPORTUNITIES FOR SYSTEMIC IMPROVEMENTS TO AUSTRALIA'S EDUCATIONAL OUTCOMES

The ODEC Final Report highlights an urgent need for the EL&L and the ELLDI. Half of the sample in the Study would not meet the minimum achievement standard for the Foundation level of the Australian Curriculum by the time they enter Grade 1. (p 104)

This is consistent with the 2023 NAPLAN results which show one third of Australian children are not meeting expectations in literacy and numeracy, with one in ten requiring additional support. Results are even more dire in lower socioeconomic areas and non-urban areas (ACARA).

---

**“An expectation of preschool is for it to offer comprehensive programs ... [that] help close the gap between those from advantaged and disadvantaged backgrounds. The EL&L program did this.”**  
(p105)

---

The Report identifies how EL&L with the ELLDI can improve educational outcomes in Australia. These can be of benefit across a number of levels:

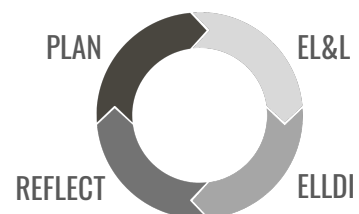
1. Early intervention by existing workforces at the local level
2. Regional and jurisdictional decision making to address educational disparities early
3. Improving Australia’s international rankings towards OECD Sustainable Development Goals (SDG) 4.1.1a and 4.2.1.
4. Translating the ELLDI into other languages, including First Nations’ languages to support children from EAL/D backgrounds.

### Early intervention by existing workforces at the local level

The EL&L strategies, implemented in response to children’s actual oral language and early literacy growth profiles from the ELLDI formative assessment, augment the skills and knowledge of existing local educational workforces to provide children with the everyday dosage of learning activities drawn from speech pathology and education that they require in order to be receptive to school literacy and classroom learning.

Recommendations from the ODEC Final Report speak to how EL&L with the ELLDI (pp6-7) achieves this:

- increasing children’s exposure to EL&L in the year before school (and earlier) and follow the child into school
- utilising the ELLDI feedback to support reflective practice to inform the selection of EL&L strategies – the cycle of planning, teaching, assessment and reflection.



In demonstrating that exposure to EL&L in the years prior to school closed achievement gaps, the ODEC Study provides evidence that local educators working in low socio-economic and other at-risk communities can be equipped to address educational language and literacy disparities early.

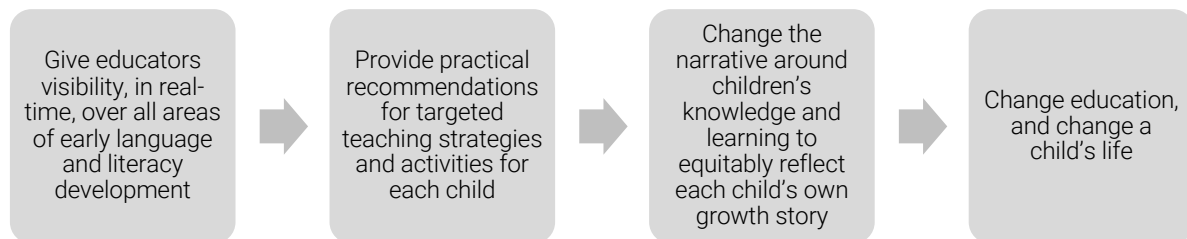
This is in spite of the fact that *“these children suffer multiple disadvantages: they receive the least preschool and other ECEC inputs and come from the most vulnerable families.”* (p105)

The Report also notes the opportunity for deepening educators' "...understanding of the reciprocal relationship between language and literacy, and how they develop. It could help educators rely less on repeatedly assessing children or teaching to the test and build their confidence in making judgements about how children's language and literacy skills are progressing." (p7).

It terms the ELLDI as a 'described scale', whereby "educators (and researchers and policy makers) can look at the items that lie above a child's current location to get a sense for what skills come next. ... This allows for targeted practice ... at the right level – to support children to consolidate what they are currently demonstrating and to be scaffolded to learn what comes next... [It] gives educators a common language...[that] allows for more productive transitions..." (p91)

The approach is an example of genuine place-based, community capacity building where local community members are breaking the cycle of intergenerational educational disadvantage.

#### **EL&L and the ELLDI:**



#### **Regional and jurisdictional decision making to address educational disparity early**

The ODEC Report notes that the ELLDI Scale can be used to contextualise other benchmarks, particularly at important Australian educational transition points such as at school entry. (pp102-103)

It also highlights the potential of the ELLDI to benefit regional and jurisdictional decision making, such as linking children's oral language and early literacy growth to child development data such as NAPLAN and other major assessments used in school to allow earlier intervention (p115); and the use of the ELLDI to support referral (for example, by maternal and child health services) into high quality preschool programs, including the EL&L program. (p114)

#### **Improving Australia's international rankings towards OECD Sustainable Development Goals 4.2.1 and 4.1.1a.**

The ODEC Study is "in line with the global efforts to increase opportunities to access highly effective ECEC programs (e.g., SDG 4.2) to contribute to economic growth because individuals are able to reach their potential and overcome the effects of inequity (OECD, 2011)." (p15).

The Report highlights that data on Australian children's ELLDI locations can be utilised to reference the external criteria of SDG 4.2.1 and 4.1.1(a). (pp101-102).

Note: up to 20% of children in the ODEC study did not meet ELLDI Level 4, indicating they have not consolidated the skills described in SDG 4.2.1. (p102)

## Translating the ELLDI into other languages

Translating the ELLDI and EL&L into a variety of languages, including Australian Indigenous Languages, provides the opportunity to extend equity of access for all children as they enter school with strong language and literacy skills in their home language and in English.

## FURTHER DEVELOPMENTS

ALNF has undertaken a number of initiatives building upon the ODEC Study to pave the way for systemic change addressing oral language and early literacy disparities.

### Digital ELLDI

Whilst the ODEC Study was underway, ALNF developed a digitised version of the ELLDI which automatically collates children's results to provide rapid feedback to educators on children's ELLDI Growth Profiles with recommended learning strategies for individuals and groups. Children engage with bespoke hard copy storybooks and illustrations, designed in the ODEC Study to mirror authentic adult-child interactions.

### Case Study of EL&L with the ELLDI in Central Australian primary school

The digital version of the ELLDI was deployed in a case study with a primary school with onsite preschool in Central Australia which provided a practical demonstration of how EL&L, the ELLDI and informed staff can maximise children's learning growth. In this project, EL&L with the ELLDI formative assessment feedback informed reflective practice to enable school staff to maximise children's curriculum learning whilst also effectively responding to speech pathology assessments and recommendations. This was presented to the ACER Education Research conference in 2022 and the paper is published at: <https://research.acer.edu.au/rc21-30/rc2022/rc2022/18/>

### 2023 research study on ECEC professional learning

In addition, ALNF is partnering with the Australian Education Research Organisation (AERO) in 2023 on a research study documenting a way to achieve efficient systemic change. Professional development delivery is provided to Educational Leaders and champions in early childhood education and care services (ECECs) with a view to sharing quality practices with all staff to support children's early language and literacy growth. This study leverages the role of Educational Leader stipulated by ACECQA, to explore a potential model for the scalable rollout of professional development to ECEC staff to support children's readiness for their transition to school.

### Towards Systemic Change

These initiatives pave the way for informed, evidence-based teaching and learning trajectories that follow the child from their early education settings into the first three years of primary school.

We are proud to present these elements that would constitute a sustainable and cost-effective national model to maximise oral language and early literacy growth for Australian children aged 2 to 8 years of age.



## ABOUT THE PROGRAMS

### EARLY LANGUAGE AND LITERACY (EL&L) PROGRAM

ALNF's Early Language and Literacy Program (EL&L) aims to address intergenerational disadvantage and inequity through education for all children. Over two decades, ALNF has successfully upskilled thousands of local educators (e.g., early childhood and primary teachers, Aboriginal Education Workers, Student Learning Support Officers and Teachers Aides), families and others in urban, regional and remote communities across Australia via its nationally accredited Certificate IV in Early Language and Literacy course and community workshops to achieve strong oral language and early literacy outcomes for children.

Combining speech and language pathology and education principles, EL&L includes practical, multisensory strategies which are easily integrated into play-based, curriculum and other pedagogical approaches to ensure that children gain a strong foundation in listening, speaking, reading, writing and comprehension so they are equipped to fully engage in education.

In 2018, ALNF commissioned an independent random control study by the Australian Council for Educational Research into the effectiveness of the EL&L Program on the NSW Mid-North Coast where it has been operating since 2008. This Report presents the findings of this work.

This place-based approach has a tiered impact in that it benefits children and also adult community members, who improve their own knowledge, skills and confidence in English and also First Languages. In this way, ALNF supports communities so that they can be the drivers of intergenerational education success into the future.

*"In over 25 years of teaching I have found [Early Language and Literacy] one of the most practical and beneficial experiences both for student and teacher learning."*

- Classroom Teacher, Mid-North Coast NSW

### EARLY LANGUAGE AND LITERACY DEVELOPMENT INDEX (ELLDI)

EL&L is complimented by the Early Language and Literacy Development Index (ELLDI). Up until now, there has not been an instrument available to educators that provides them with valid, reliable and timely information about children's oral language and literacy growth in the prior-to-school years. The ELLDI was designed to do exactly this, and it is a world-first.

The ELLDI includes a world-first progression scale in the 2-5 age bracket providing a new capacity for insight into children's language and literacy growth at this highly sensitive age across 7 subdomains: phonological awareness; vocabulary; oral language (including fluency, volume and clarity); pre-writing and writing; reading; print awareness and comprehension. The ELLDI Scale provides robust, statistically valid and reliable measures of children's progress to enable evidence-informed, reflective practice. It extends up to 8 years with numerous age-appropriate booklets allowing for repeated ELLDI Interviews to provide ongoing insight into children's growth needs.

It also enables a strengths-based approach to children's learning. Rather than measuring children against learning benchmarks, the ELLDI places children on a progression scale to provide rapid insight on children's growth.



# The Overcoming Disadvantage in Early Childhood Study

Evaluation of the Australian Literacy and Numeracy  
Foundation's Early Language and Literacy Program.

Final Report

Dan Cloney  
Kellie Picker  
David Jeffries  
Prue Anderson



# *The Overcoming Disadvantage in Early Childhood Study: Evaluation of the Australian Literacy and Numeracy Foundation's Early Language and Literacy Program. Final Report*

Dan Cloney, Kellie Picker, David Jeffries and Prue Anderson

Australian Council for Educational Research Ltd  
19 Prospect Hill Road,  
Camberwell, Victoria, 3124,  
Australia  
Phone: (03) 9277 5555  
ABN 19 004 398 145

[www.acer.org](http://www.acer.org)

ISBN 978-1-74286-710-6

© The Australian Council for Educational Research 2022



With the exception of any material protected by a trademark, and where otherwise noted, all material presented in this document is provided under a Creative Commons Attribution 4.0 International licence, <https://creativecommons.org/licenses/by/4.0>

## **Recommended citation**

Cloney, D., Picker, K., Jeffries, D., & Anderson, P. (2022). *The Overcoming Disadvantage in Early Childhood study: Evaluation of the Australian Literacy and Numeracy Foundation's Early Language and Literacy program. Final Report*. Australian Council for Educational Research.  
<https://doi.org/10.37517/978-1-74286-710-6>

## **Acknowledgments**

The Overcoming Disadvantage in Early Childhood Project was funded by donations from Eureka Benevolent Foundation and CAGES Foundation to the Australian Literacy and Numeracy Foundation's Early Language and Literacy program. ALNF is a registered national charity dedicated to raising language, literacy and numeracy standards in Australia. ALNF's programs aim to create a more inclusive Australia, unlocking potential through the power of language, literacy and learning. It works in partnership with First Nations, refugee and other vulnerable Australians to address the language and literacy learning needs of individuals and communities, giving them a better chance to succeed.

A heartfelt thanks to all the study participants. In particular, thanks to the local services and schools who welcomed us into their settings and to the children, families, and staff who participated in the research. Without you, this research could not have happened.

ACER acknowledges the support of the Overcoming Disadvantage in Early Childhood Study Steering Committee chaired by Associate Professor Delyse Hutchinson (Deakin University) and with members Rachel Kerry (CAGES Foundation), Mary-Ruth Mendel (ALNF), Eric Brace (ALNF), Prue Anderson (ACER), and Dr Dan Cloney (ACER). ACER also acknowledges the open and positive engagement of the ALNF team in the evaluation process.

Cover photo: Early Language and Literacy Program.  
Supplied by the Australian Literacy and Numeracy Foundation

Cover design: ACER Creative Services

# Executive summary

The Overcoming Disadvantage in Early Childhood (ODEC) study is a longitudinal evaluation of the Australian Literacy and Numeracy Foundation's (ALNF) Early Language and Literacy (EL&L) program. The study was designed to answer the research question: What is the effect of the EL&L program on the development of language and literacy skills in preschool aged children?

An integral component of this study was to develop a new measure of oral language and literacy – the Early Language and Literacy Developmental Index (ELLDI). The ELLDI was developed in recognition of the need for a best-practice measure of oral language and literacy that was fit for purpose. The measure needed to:

- a. capture within-child growth over many years (from before preschool to the end of the lower grades of primary school)
- b. be suitable for use in a research study asking about average learning trajectories over time
- c. be usable into the future by ALNF as a core part of their program, including being used as an adaptive formative assessment, linked to a described scale, and also allow for future development including the addition of new assessment material.

No such measure exists at the time of publication.

In particular, the ELLDI is internationally unique due to its progression scale in oral language and early literacy for 2–5-year-olds.

## Early Language & Literacy (EL&L) program

The EL&L program is a language and literacy intervention created by ALNF. It operates on a program logic whereby the teaching capacity of educators<sup>1</sup> is developed through a nationally accredited Certificate IV course including training and on-site mentoring, aimed at improving language and literacy knowledge, planning, integrated programming and curriculum implementation, and everyday pedagogical practices in education settings. Educators' increased capacity is applied in everyday practice and pedagogy to improve the learning outcomes of children.

The strength of the EL&L course is drawn from speech and language pathology and education, particularly research in the science of reading and effective practice of educator professional development (PD). The science of reading synthesises extensive empirical research to identify the language and literacy skills that underpin reading development. This research stresses the importance of the explicit teaching of six components of reading: oral language, phonological awareness, phonics, vocabulary, fluency, and comprehension. EL&L draws on this research to ensure educators have the content knowledge and skills to understand what and how to teach the components that lead to effective reading. Added to this, mentoring provides educators with opportunities to review their implementation of EL&L and reflect on how this is impacting children's learning.

---

<sup>1</sup> When discussing the EL&L program the term educator includes teachers, paraprofessionals, and non-qualified staff across learning contexts.



## Overall study design

This report documents the sum of work completed across the duration of the study from 2018-2022. The design of the study was based on the principles of a longitudinal outcome evaluation and had three components:

- a. Design of the outcome measure: the ELLDI
- b. Interim results and recommendations
- c. Final results, recommendations and reporting.

### Design of the ELLDI

The ELLDI was designed to meet the needs of good quality research *and* formative practice. To do this, it is necessary to bring together: (1) cutting edge psychometric and assessment development methods seen in the large-scale assessment and national assessment programs with (2) authentic interactions between adults and children in early childhood settings.

Such an assessment would not appear as a static test or a checklist, but rather as an interaction between an adult and a child: reading a picture book; telling stories; and responding by speaking and pointing (and later, when children are more advanced and typically in school, reading and writing). Children's responses are captured, encoded by an expert administrator (including by educators when used formatively) and scaled using item response theory to yield a linear (interval) measure. By choosing a set of items that are fit for a specific purpose, the assessment can be used for research, population measurement, or as a formative assessment in the classroom.

Information about the relative challenge of each item and task on the measure can provide insights into what children can do (and has the useful property of not requiring all children to answer the same items) – the assessment can be targeted to provide the right amount of challenge to children. This feature has a dual purpose. The first relates to the ability to describe the scale, to yield a progression of skills and abilities and a description of how these skills and abilities change. The second relates to measurement efficiency whereby better targeted testing minimises the uncertainty in estimates of children's abilities.

ACER test developers undertook a review of existing assessment frameworks and adapted the Longitudinal Literacy and Numeracy Study (LLANS) framework (Meiers et al., 2006), as the basis of the initial description of key skills and abilities within the domains. This study provided the blueprint for national *on-entry* assessments including The Early Start program in Queensland and the Best Start school-entry assessment in New South Wales. It also led to the LLANS: Transitions from Preschool to School (LLANS: TPS) study, that developed measures of children's literacy development at the end of preschool.

In this study, the conceptual framework described is lower, and a new set of assessment items and materials are developed to target earlier in children's learning and development, but from a well validated beginning point that is highly relevant to children's learning in preschool and the early grades of primary school.

Such an assessment design is a significant innovation for the early years sector – both for researchers and for educators. The construction of such an assessment can challenge the sector to advance towards using learning progressions to understand how language and literacy develops over time, while providing a common language of learning. This translates into educators being able to pinpoint what knowledge and skills children have in order to target their teaching at what comes next, while being able to carefully monitor and map progress over time.

### Evaluation design

The 571 children who participated in the study were recruited from 12 centres implementing the EL&L program (intervention centres) and 10 that did not (control centres), consisting of a mix of preschools and long day care centres located on the Mid North Coast of New South Wales. Intervention centres were selected based on staff members having completed the EL&L training, and the EL&L program being implemented in their classrooms.

During the study, the impact of the intervention on children's language and literacy development was measured at four timepoints using the ELLDI. That is, the longitudinal progress of children was observed. Along with this, a set of centre and family questionnaires were used to collect contextual and background information, and the quality of the teaching practice in settings was also observed.

The study is non-experimental. The EL&L program is already well established, having been in operation in the region for more than 7 years at the start of the study. The children who attend the intervention centres do so because of the availability of these programs to families. To address the risk of selection bias – where some external, extraneous factor causes a difference between intervention and control groups – two strategies are implemented. The first is the recruitment of a matched control group. The second is the collection of a rich set of covariates to control for other factors that also affect learning and development – relating to the child, family, home environment, local community, and centre. The control group is matched on a set of observed characteristics, including community SES, the National Quality Standard Rating of centres, the centre type and size.

## Key findings

### Design of the ELLDI

In each year of the study, ACER test developers created new assessment materials. This included new items/tasks, as well as resources including picture books, story prompts, and rubrics. After data was collected in the field, psychometric analysis was undertaken to validate the new materials and equate only the best content to the ELLDI Scale.

The measure developed is reliable and valid, capturing skills across the language and literacy strands of print conventions, phonological/phonemic awareness, phonics, vocabulary, reading comprehension, oral language and writing for children from age 2 to age 8 (Cloney et al., 2022; Cloney & Picker, 2021). The measure has interval properties and can be used to describe a progression in learning across major transition points. New assessment material can be developed and equated onto the measure and the tool can be used formatively in ECEC settings and in primary school classrooms as well as for research and population reporting.

The observed learning trajectories are curvilinear. That is, children grow fastest when they are youngest, and the rate of growth slows as they transition into school and across the first three years of school. This is a pattern typically seen in cognitive and academic development.

### Interim results

Children in the intervention group are significantly more disadvantaged than those in the control group. This is because ALNF operates its EL&L program in a particularly disadvantaged sub-population of the mid North Coast. Despite efforts to select a similar group of centres as intervention groups, there was not a perfect match. The intervention sites tend to operate in the lowest socio-economic status (SES) neighbourhoods and are generally rated as lower quality on the ACECQA National Quality Standard, both more likely to be *working towards* the standards, and less likely to be *exceeding* the standard. Reflecting this, the children who are attending these services are also more disadvantaged.

The intervention group comes from households with lower incomes that are more likely to be single carer households. Their families are more likely to experience unemployment and to have less education attainment. Children in the intervention group also receive less educational stimulation at home (e.g., being read to).

This is compounded by the general findings that many children in the region do not receive the universal entitlement to 15 hours of preschool education in the year before school. The reason for this is unknown, however, in this study, more than 30% of children entitled to universal access to preschool did not receive 15 hours of formal ECEC programming per week.

Taken together, it was predictable that the intervention group entered the study with lower oral language and literacy skills than their peers in the control group. Those children in the intervention group, exposed to the EL&L programs, were able to catch up to their peers, such that in 2019 there was no difference between the two groups after accounting for background factors. This represents a positive piece of evidence of the potential for the EL&L program to narrow and even close the gap.

### Final results

The gains made in the first two waves of the study are not maintained when contextual factors – disadvantage – are not controlled for. By the end of the study, children from the intervention group are significantly behind their peers, though not as much as at entry to the study. That is, some of the gains are maintained. When we control for contextual factors, in a linear growth model, the children from the intervention group are not significantly behind their peers. The achievement gap is explained by contextual factors.

To explore why the benefits of the program are outweighed by contextual factors, analysis of the treatment effect is undertaken. This analysis shows that very few children in the study were exposed to the intervention for an extended period: 47 (out of 558) children were exposed to low-intensity (or greater) for all four cycles, 15 (out of 558) children were exposed to continuous medium or high levels of intervention intensity beyond the first two years of the study. No children were exposed to continuous high levels of intervention intensity beyond the first two years of the study.

As children moved out of preschool settings and into school, they tended to move into sites that either did not offer the EL&L program, or who had previously been engaged but had since either disengaged completely (e.g., due to staff turnover) or reverted to low levels of fidelity of implementation.

In essence, due to the complexity and size of the school system around transition/entry to school, the intervention is in effect a 12-month intervention on preschool age learning. There are few children who receive intensive literacy intervention across preschool and the formative years of early primary school. Without continuous intervention, the observed effect is as expected: the intervention cohort, disadvantaged compared to the control group, do not maintain the early gains observed.



# Recommendations

## Interim recommendations

1. Implement the use of the ELLDI assessment with the EL&L program as a formative assessment.
  - a. Construct a described scale to simplify the interpretation of the results of the ELLDI.
  - b. Construct supporting materials to support the interrogation of ELLDI results for individual children – particularly to quantify growth, and to highlight strengths and limitations.

The authors note that all interim recommendations have been implemented and documented by ALNF to support improved professional discussions, explicit planning and learning outcomes (Groom et al., 2022).

## Final recommendations

1. **Increase children’s exposure to the intervention in the preschool year** (and earlier) through intensifying the program expectations (hours per day and days per weeks). Such intensity is required time for consolidation and mastery of language and literacy skills when it is likely that most children will not get long-term exposure and support in the primary school years under the current scope of the program.
2. **Build further quality control into the EL&L program** to ensure fidelity of implementation and maximise engagement with the program. For example, ALNF mentors could use measures of pedagogical quality as a lens to focus and improve practice (Cloney & Hollingsworth, 2018). If one or more ALNF staff were trained to use the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2009), they could continue to support the implementation of the EL&L program while also supporting educators to improve the quality of their interactions with children within and beyond EL&L experiences. In this context, interactions refer specifically to the pedagogical strategies used by educators to model language, to give feedback and to introduce and connect concepts (Levickis et al., 2023).

3. In addition to this, ALNF may elect to **use more general measures of ECEC program quality**, including the MELQO MELE to monitor or track quality in its settings. This could be done referring to external benchmarks to ensure the quality in EL&L sites is high. An observation schedule was developed as part of the ODEC study, and this could be deployed or adapted as required.
4. **Continue to embed use of the ELLDI into the EL&L program.** The ELLDI becomes a part of the all-important plan, teach, assess and reflect cycle for effective teaching. This should be embedded in the EL&L program.
5. **Use the ELLDI Scale to describe a learning progression, and develop tools to orient educators to the learning being observed,** what comes next, and which EL&L materials and strategies are well-targeted to these needs. Such a learning progression has the potential to empower educators with a deep understanding of the reciprocal relationship between language and literacy, and how they develop. It could help educators rely less on repeatedly assessing children or teaching to the test and build their confidence in making judgements about how children's language and literacy skills are progressing.
6. **Use the ELLDI to link to the learning outcomes of children on NAPLAN.** For example, with children who are in the lowest performance group (those in NAPLAN bands 1 (below the national standard) and 2 (at the national standard)), the ELLDI can be used to describe the specific skills and abilities these children demonstrate. Importantly, the significant overlap in the ELLDI and NAPLAN Reading means that children's developmental trajectories on the ELLDI can be used to identify children requiring additional support to meet national minimum standards on NAPLAN well before they arrive at school. This may involve empirical work, including formal equating where possible.
7. **Expand the scope of the program to cover more school sites** or to implement intervention models that follow the child, no matter their location.
8. **Broaden the use of the ELLDI to support referral into high quality preschool programs,** including those offering the EL&L program. If, for example, maternal and child health nurses could identify children with, for example, persistent ear infections, they could administer the ELLDI and refer children into preschool programs where they may otherwise miss out.
9. **Translate and adapt the ELLDI into other languages, including Indigenous languages,** to support children who do not speak English at home. Such approaches can allow children to demonstrate that they are acquiring the language skills that will later help them to successfully transition to the language of instruction in school.
10. **Develop a numeracy measure that assesses mathematical literacy –** to analyse, reason, and communicate mathematical ideas and to solve problems in a wide range of contexts. Like literacy, numeracy is a general skill that is applied in many curriculum contexts and is essential for lifelong success.

## Conclusions

The EL&L program is demonstrated to have significant promise, including the potential to narrow gaps between children from disadvantaged backgrounds and their less disadvantaged peers. This is highly contingent on access to the program.

Most importantly, across the span of the study, ALNF acted upon the interim findings to strengthen the program and this has resulted in a program that is significantly more focused on teaching and learning activities at the right level: educators can use good quality assessment (i.e., ELLDI) to generate robust evidence of what children can do – their strength's and limitations – and use this to select EL&L materials and to plan at the right level so as to effectively scaffold children's learning. This also bolsters the capacity of educators by building a shared understanding of children's development – supporting professional discussions and improved planning – hallmarks of quality embedded within the National Quality Standard.

These findings are *not* reflected in the final results of this study – in order to preserve the fidelity of the research design, ALNF did not modify their programming for the cohorts of children participating in the study on the Mid North Coast of NSW. It is likely the EL&L program of 2023 is substantially stronger from the one initially observed during the study presented here. It is expected that this will result in stronger impact on children's learning.

There is a crisis of learning in the regions where ALNF is operating the EL&L program. Children's learning on the Mid North Coast of NSW, an area typified by significant disadvantage, is lower than we would expect to result in strong outcomes in school. This is true for all children in the study – intervention and control. This study estimates that only approximately half of students enter school ready to engage and achieve minimum expectations in the Foundation Curriculum. Further, up to 10 per cent of children do not meet the international minimum proficiency level established as part of the Sustainable Development Goals: a metric primarily focused on learning in developing contexts. This is coupled with an underinvestment in intervention in general.

Despite efforts to provide programs to children, the EL&L program operates in a small fraction of the schools and services in the region. When children move schools there is a high chance they slip out of the program. While it may be assumed that children could or should get the EL&L program for five (or more years) - from the year before preschool to grade 3 - most children get no more than 12 months of support (this is because they transition out of preschool into a school that is not providing EL&L or some EL&L sites have disengaged from the program, for example due to staff turnover).

This should be a stark insight to those wishing to successfully intervene in children's early learning through the school system. The quantum of investment needed to reach children in a sustained way must ensure the support can either:

1. follow the child (no matter which school they attend or move to) or,
2. ensure there is sufficient coverage of schools so that children do not fall through the cracks.

This requires an expansion of coverage beyond the scope of the current EL&L program.

The EL&L program demonstrates a capacity to close achievement gaps for children who access the program. This is notable given the highly disadvantaged population in the intervention group. These children are disadvantaged even relative to the local community that is almost entirely in the lowest 25 per cent of SES in Australia. The sample of children who enter the EL&L program in 2018 begin the study behind their peers. After 12 months of exposure, the learning gap is closed for children in the EL&L program. Over time, these gains are lost as children leave intervention EL&L programs and enter business-as-usual schooling.

This learning gap is wholly explained by child, family, and home contextual factors – a product of disadvantage. Holding these factors constant, we observe no difference between the intervention and control groups. This is consistent with findings that show that persistent intervention is needed over early years and lower primary to build a solid language, literacy and reading foundation.



# Table of contents

<b>Executive summary</b> .....	<b>1</b>
Early Language & Literacy (EL&L) program .....	1
Overall study design .....	2
Key findings .....	4
Recommendations .....	6
Conclusions .....	8
<b>Table of contents</b> .....	<b>10</b>
<b>List of tables</b> .....	<b>11</b>
<b>List of figures</b> .....	<b>13</b>
<b>Introduction</b> .....	<b>14</b>
Study rationale.....	15
EL&L intervention program .....	16
<b>Method</b> .....	<b>17</b>
Sampling.....	17
Instrumentation .....	22
Questionnaires.....	31
Fieldwork .....	33
Data processing and analysis.....	36
Analytic approach .....	39
<b>Results</b> .....	<b>46</b>
Descriptive analysis .....	46
Item response theory (IRT) analysis.....	65
Population model.....	74
Longitudinal analysis .....	82
Intervention intensity.....	88
Contextualising findings.....	91
<b>Discussion and recommendations</b> .....	<b>105</b>
The Early Language & Literacy (EL&L) program.....	105
The Early Language and Literacy Developmental Index (ELLDI).....	111
Concluding remarks .....	116
<b>References</b> .....	<b>118</b>
<b>Appendices</b> .....	<b>125</b>
Appendix 1: ELLDI Assessment Booklets .....	126
Note: Appendices 2-16 are located in a separate document.	

## List of tables

Table 1. Design parameters for study.....	17
Table 2. SES and centre size for simulated samples drawn from three possible geographic locations .....	19
Table 3. NQS quality ratings for simulated samples drawn from three possible geographic locations. .....	20
Table 4. Centre types for simulated samples drawn from three possible geographic locations.....	20
Table 5. SES and centre size characteristics of achieved sample.....	21
Table 6. NQS quality ratings for the achieved sample.....	21
Table 7. Centre types of achieved sample. ....	22
Table 8. Summary of the achieved sample in the ODEC study.....	22
Table 9. Measures used in the study.....	23
Table 10. Type of visual stimulus used to prompt an oral language response during each assessment cycle. ....	25
Table 11. Summary of sub-domains and items of the assessment of language and literacy in the first cycle of data collection.....	26
Table 12. Summary of sub-domains and items of the assessment of language and literacy in the last cycle of data collection.....	27
Table 13. Items included in the measure of quality from the MELQO MELE.....	29
Table 14. The three CLASS domains and associated dimensions.....	30
Table 15. Yearly ODEC fieldwork staff breakdown.....	34
Table 16. Children interviewed during assessment cycles.....	34
Table 17. CLASS observations across control and intervention centres.....	36
Table 18. List examples of variables in child assessment booklets that contained keystroke errors (with count of errors and examples). ....	38
Table 19. Variables included as regressors in the population model.....	40
Table 20. Characteristics of the children in the study.....	46
Table 21. Characteristics of the main caregivers in the study and their households.....	49
Table 22. Characteristics of directors and services.....	50
Table 23. Status of Director Training in EL&L.....	50
Table 24. Directors' participation in EL&L professional development activities in past 12 months ..	51
Table 25. Total months staff have delivered EL&L program in Centre.....	51
Table 26. Characteristics of educators and study rooms.....	52
Table 27. Completion of training in ALNF's EL&L program.....	53
Table 28. Year completed ALNF's EL&L program.....	53
Table 29. Total months of delivery of EL&L program in Centre.....	54
Table 30. Frequency of visits from ALNF EL&L mentors.....	54
Table 31. Frequency with which participants are explicitly taught to... ..	55
Table 32. Frequency with which participants use... ..	56
Table 33. Participation in professional development activities in past 12 months - type.....	56
Table 34. Participation in professional development activities in past 12 months - content.....	57
Table 35. Frequencies of key contextual categorical variables (N (%)).....	60
Table 36. Descriptive statistics of key contextual continuous variables.....	63
Table 37. Item recodes.....	66
Table 38. Bundled items.....	69
Table 39. Unconditional correlation/covariance/variance matrix.....	77
Table 40. Average oral language and literacy (ELLDI) achievement of children with ear infection/glue ear compared with the rest, by year.....	77
Table 41. Regression estimates from population model.....	80
Table 42. Mean abilities of children across cycles (logits).....	83
Table 43. Mean abilities of children by intervention and control group and cycle.....	83
Table 44. Comparing growth models.....	84
Table 45. Comparing actual means with model-implied means by cycle.....	84
Table 46. Regression estimates for conditional latent growth model – Intervention effect.....	86

Table 47. Regression estimates for conditional latent growth model – Intervention effect + Additional child and family effects .....	87
Table 48. Criteria for rating intervention intensity .....	88
Table 49. Descriptive statistics for ELLDI by continuity of intervention intensity – ELLContInt1: Low, Medium and High.....	89
Table 50. Descriptive statistics for ELLDI by continuity of intervention intensity variables – ELLContInt2: Medium and High.....	90
Table 51. Descriptive statistics for ELLDI by continuity of intervention intensity variables – ELLContInt3: High .....	90
Table 52. ELLDI Scale.....	93
Table 53. Proportions of children in each level on the ELLDI, by year .....	98

## List of figures

Figure 1. Distribution of children's exposure to formal ECEC programs (including home-based) by intervention group .....	48
Figure 2. Boxplots of the distribution of ECEC quality .....	59
Figure 3. Iteration history for Xsi (Item difficulty) .....	70
Figure 4. Adequately discriminating dichotomous item .....	72
Figure 5. Adequately discriminating polytomous item .....	72
Figure 6. Over-discriminating item.....	73
Figure 7. Under-discriminating item.....	73
Figure 8. Item-person map for 2018 .....	75
Figure 9. Item-person map for 2019 .....	75
Figure 10. Item-person map for 2021 .....	76
Figure 11. Item-person map for 2022 .....	76
Figure 12. Density plot of abilities by cycle.....	82
Figure 13. Cubic growth curve (blue) plotted over individual child abilities (piecemeal growth, in grey) over time .....	85
Figure 14. Item-person map for the ELLDI Scale, including RP65 thresholds from all cycles (2018 sample only) .....	99
Figure 15. "Predicted percentage effects on adult earnings of early childhood programmes, based on test scores versus adult outcomes" Reproduced from OECD (Shuey & Kankaraš, 2018) .....	117

# Introduction

Early Childhood Education and Care (ECEC) programs have been recognised internationally as an effective public-policy intervention to promote growth and reduce inequality. They represent the most cost-effective intervention on human capital and skill formation and increase capacity in both the short-term and across the lifespan. Children from low-income families in Australia are at greater risk of poorer developmental outcomes than their more socially advantaged peers (Cloney, 2016) and this pattern is further well-established in research from the United States (Duncan & Sojourner, 2013), Europe (Dearing et al., 2018) and Australia (Nicholson et al., 2012). There is considerable evidence that ECEC programs are a potent intervention, as showcased in randomised controlled trials in the United States (Campbell & Ramey, 1994; Schweinhart, 2005). These studies show that ECEC programs have the potential to ensure that children enter school on a more equal developmental footing relative to their peers with higher socio-economic status (SES) backgrounds (Pungello et al., 2010; Schweinhart, 2005). ECEC potentially has the net effect of reducing the influence that SES has on lifelong outcomes and the intergenerational transmission of poverty and disadvantage.

A strong predictor of achievement and an important area of disparity between children from high and low SES backgrounds is early language and literacy development, where children from low SES homes have been identified as having lower language skills, apparent even in the early months of life (Hoff & Ribot, 2015). This disadvantage can be countered with quality ECEC intervention programs shown to enhance early language and literacy skills and promote school readiness (Chambers et al., 2016; Hattie, 2009). These quality programs provide children with experiences to simultaneously learn to use language while also learning through language (Paatsch & Nolan, 2020). Such programs provide a comprehensive curriculum which balances skills-focused and developmental learning experiences, while setting out clear learning goals supported with the use of appropriate measures, so that educators understand the impact of the program and children's progress (Chambers et al., 2016). The accompanying learning experiences are rich with interactions where more effective language users model language use and provide children with opportunities to use and grow oral language skills, such as phonology, vocabulary, semantics and syntactics, which are linked to later reading success (Foorman et al., 2015; Hart & Risley, 1995; Konza, 2014; Morrow & Rand, 1991; Morrow et al., 1998).

It is expected that ECEC programs focus on teaching language and literacy skills to support the shift from spoken language to written language. Hence, for children to understand and learn the concepts and principles of print, which are foundational to reading, they require the systematic teaching of the components that lead to proficient reading. These components are drawn from extensive empirical research known as the 'science of reading', which advocates for the teaching of oral language, phonological awareness particularly phonemic awareness, phonics, vocabulary, comprehension (listening and reading) and fluency (NICHD, 2000; Rose, 2006; Rowe, 2005). There is considerable evidence that ECEC programs teaching pre-literacy skills and the components from the science of reading (sometimes called the Big Six), impact language and literacy development that go beyond preschool and into kindergarten. This

impact extends to school readiness, as well as reducing the influence of SES (Chambers et al., 2016).

Recent research into child development and educator professional development (PD) established that improvements to pedagogical quality were associated with improved child development (Egert et al., 2018; Werner et al., 2016). Effective PD programs bring about such changes in educator's practice by moving delivery beyond dissemination of new ideas to providing them with experiences requiring the connection of these ideas to their actual classroom practice. According to Siraj et al. (2020) these programs include the critical components of the 'who': participant, provider and the context: 'what': the content and outcomes of the PD and 'how': duration, delivery approaches and formats. High-quality PD programs have been shown to include essential characteristics such as, *situated practices* that connect 'what' is being taught to the context for implementation, *active learning experiences* link to the 'how' by positioning educators as learners experiencing new instructional strategies in order to reflect on their effectiveness from a student's perspective, as well as *involving educators in sustainable activities* that are cyclical and involve opportunities for enactment and reflection of practice (Borko et al., 2010). In addition, it is important that educational leaders support teachers' engagement and participation to learn through the PD programs (Timperley et al., 2007). Finally, to increase the impact of PD, it has been shown that the combination of in-service PD programs and educator mentoring improves teacher instruction, with mentoring thought to be more effective when teachers receive PD aimed at developing their knowledge, skills and practices, followed by targeted mentoring (Kraft et al., 2018).

## Study rationale

The Overcoming Disadvantage in Early Childhood (ODEC) study contributes to the need for the identification of specific intervention programs that mitigate the negative effects of poverty and promote equitable opportunities for children. Evidence of such programs is essential to encourage public investment in ECEC programs to the levels where significant public benefits (e.g., greater employment, greater tax income, fewer arrests and crimes, fewer negative or risky behaviours and fewer welfare and public costs (Schweinhart, 2005)) are accrued. Such research is in line with the global efforts to increase opportunities to access highly effective ECEC programs (e.g., SDG 4.2) to contribute to economic growth because individuals are able to reach their potential and overcome the effects of inequity (OECD, 2011).

The aim of the study is to measure the impact of the Early Language and Literacy (EL&L) intervention on children's early language and literacy, and answer the research question: What is the effect of the Early Language and Literacy (EL&L) program on the development of language and literacy skills in preschool aged children? This will be achieved by looking at children's growth longitudinally, in the short term (the preschool year) as well as the medium term (after 2, 3 and 4 years).



## EL&L intervention program

The EL&L Certificate IV course is designed to combine speech pathology and education theory and practice to improve participants' ability to implement quality and targeted language and literacy learning experiences in the early years (0-8-years-old). The course involves the completion of six core units, that integrate theory, with practical classroom-based experiences and assessments. The Certificate IV focuses on improving the educator's ability to deliver language and literacy content through the use of learning experiences aimed at enhancing children's foundational language, literacy and pre-reading skills. These learning experiences are designed to be multisensory and multimodal to enhance engagement and explicitly target children's oral language, phonological/phonemic awareness, print awareness, vocabulary, early writing and listening comprehension skills to improve their language and literacy development. The program provides ongoing mentoring support through site visits and continual professional learning (for example, trainers modelling and coaching EL&L activities with educators on site). The program also provides materials to support the explicit activities (e.g., picture books, props to count syllables and segment words).

The program draws from the science of reading empirical research identifying key components of reading to support children in becoming proficient readers. The components of reading and other important emerging language and literacy skills are targeted throughout the EL&L program to provide educators with a range of developmentally appropriate, carefully planned, learning techniques to meet the children's developing language and literacy needs. Added to this, EL&L training and implementation is supported with mentoring to provide educators with opportunities to review their use of the program and have their children assessed using the School Entry Alphabetic and Phonological Awareness Reading Test (SEAPART) to understand how the program impacts children's learning. It is recommended in the EL&L program that preschool children receive approximately 2 hours of integrated EL&L practice, as a part of their day at preschool and primary school children receive a minimum of 2 hours a day.

# Method

## Sampling

The target sample was 20 preschools<sup>2</sup>: 10 EL&L services, and 10 control services drawn at random. The control sites were stratified to closely resemble the population of EL&L sites (matched-controls) and replacement sites were identified for each sampled control service so that any refusals to participate could be replaced with a like-service.

Up to two classrooms working with 3- and 4-year-old children (e.g., sessional groups) at each selected service were chosen at random to participate. Every child in the selected classrooms was invited to participate. A simulated power study indicated that this would produce a sample of adequate power to observe small and medium effects in statistical models (based on assumptions of an achieved sample of 285 children with complete data and balanced for intervention and control groups). The broad design is summarised in Table 1.

One parent or caregiver for each child was invited to participate, as well as the lead educator in each classroom or sessional group, and one centre director (and school principals when the study moved into schools) from the service.

*Table 1. Design parameters for study*

	Intervention (EL&L)		Control		Comments
	min	max	min	max	
ECEC services	.	10	.	10	
Classrooms (k)	.	20	.	20	Two preschool rooms/groups per ECEC setting
Teachers	.	20	.	20	Lead teacher in each room/group
Children (n)	200	225	200	225	~ 10 per room

## Target population

The study focused on children participating in formal ECEC programs in the year before school. The target population was therefore all children, participating in preschool programs where ALNF was operating its EL&L program on the Mid North Coast of New South Wales.

---

<sup>2</sup> Preschool is the name in New South Wales for pre-primary education, available nominally to all children in the year before starting school (approximately age 4). Preschool can operate in long day care, stand-alone preschool, and school co-located settings.

## Achieved population

A sampling frame was established by using the NSW Department of Education Early Childhood Education Directorate's list of registered Early Childhood Education and Care (ECEC) services (2017)<sup>3</sup>. ALNF identified 14 services that were participating in the EL&L program across the Mid North Coast region of NSW (in the Nambucca, Kempsey and the former Greater Taree local government areas (LGAs)).

## Target sample

The target sample included all of the ALNF EL&L sites (anticipating that some would refuse and the recruited number would approximate the required 10 sites) plus a control group of 10 ECEC services closely matched on observable characteristics on the Mid North Coast of New South Wales.

Three definitions of the target sample were considered: (1) ECEC services in Kempsey and Nambucca LGAs, (2) ECEC services in Kempsey and Nambucca LGAs and the proximal LGA of Port Macquarie-Hastings, and (3) ECEC services in Kempsey, Nambucca LGAs, Port Macquarie-Hastings, and the former<sup>4</sup> LGA of Greater Taree. To assess the three options, samples were drawn and their characteristics compared on observable traits. To draw the sample in each case, the following steps were taken:

1. Three sampling frames were created, including only the LGAs listed in each option.
2. Services that were *not* long day care (LDC), stand-alone or school co-located preschool (PRE), or mixed (e.g., hubs offering a mix of services including preschool) services were excluded. This is predominately home-based services (family day care).
3. Each sampling frame was ordered by:
  - a. Program type
  - b. National Quality Standard (NQS) overall rating
  - c. Australian Bureau of Statistics (ABS) Socioeconomic Index for Areas (SEIFA) Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD) (Pink, 2008) postal area centile within NSW<sup>5</sup>
  - d. Centre licensed capacity
4. The 14 ALNF sites were identified through information provided by ALNF. Control group sites were selected by choosing the services listed below each ALNF site in each sampling frame. Where two ALNF sites were listed one after each other, the two sites listed below both were selected as control sites. For each control site, up to two replacement sites were selected – the centre above and below the selected control service that was not an ALNF EL&L site.

---

<sup>3</sup> See file "1List of ECE services in NSW for ACER 20171108.xlsx"

<sup>4</sup> now subsumed into Mid-Coast Council. ALNF had previously identified the LGA of Greater Taree as an area of interest given its socio-demographic characteristics and had commenced work in the area.

<sup>5</sup> ABS 2033.0.55.001

Small variations in the above process were made due to the size of the sampling frames. In the case of the first sampling frame (Kempsey and Nambucca LGAs) there were only nine eligible control sites after the ALNF sites were identified and all these were selected as control sites. In the remaining two sampling frames, there were varying numbers of eligible control sites and this resulted in slightly different numbers of control and replacement sites being selected (see tables below) as the number of eligible sites was exhausted.

Table 2 summarises the three achieved samples (one for each option listed above) in terms of licensed capacity (measure of size) and SES. Selecting sites only in Kempsey and Nambucca produced good balance in terms of SES (SEIFA<sup>6</sup>), although the control sites were slightly larger (higher capacity) and there were no available replacement sites (the nine control services were all the remaining services in the region). Selecting control sites from Port Macquarie-Hastings provided more viable sites (both more control and replacement sites), however, the services were much higher in SES (the median centile of (within-state) SEIFA was 33rd versus 5th). Selecting control sites from Kempsey, Nambucca, Port Macquarie-Hastings, and Greater Taree produced a good balance in terms of size and SES and there were sufficient sites to have replacements for most control sites.

*Table 2. SES and centre size for simulated samples drawn from three possible geographic locations*

Sample	Sample group	n	Mean measure of size	Median SEIFA IRSAD	Median SIEFA State Centile
Kempsey LGA, Nambucca LGA only	EL&L	14	42	863	5
	Control	9	47	908	15
Kempsey LGA, Nambucca LGA, and Port Macquarie-Hastings LGA	EL&L	14	42	863	5
	Control	14	45	945	33
Kempsey LGA, Nambucca LGA, Port Macquarie-Hastings, and Greater Taree LGA	EL&L	14	42	863	5
	Control	11	39	913	17

Table 3 summarises the three sampling options in terms of quality as measured by the NQS. Sampling sites in only Kempsey and Nambucca produced a very unbalanced sample with zero control sites exceeding the NQS. Good balance was achieved in both alternative options.

<sup>6</sup> SEIFA national mean (SD) is 1000 (100), ABS 2033.0.55.001.

Table 3. NQS quality ratings for simulated samples drawn from three possible geographic locations.

Sample	Sample group	n	n exceeding	n meeting	n working towards
Kempsey LGA, Nambucca LGA only	EL&L	14	4	7	3
	Control	9	0	4	5
Kempsey LGA, Nambucca LGA, and Port Macquarie-Hastings LGA	EL&L	14	4	7	3
	Control	14	3	7	4
Kempsey LGA, Nambucca LGA, Port Macquarie-Hastings, and Greater Taree LGA	EL&L	14	4	7	3
	Control	11	4	5	2

Table 4 summarises the three sampling options in terms of mix of services selected. Selecting control sites in Kempsey and Nambucca created an imbalance in the mix of services with many more LDC services being recruited to the control group – reflecting that the majority of stand-alone preschools in the region were participating in the EL&L program. The two alternative sampling approaches produced good balance in the mix of service types.

Table 4. Centre types for simulated samples drawn from three possible geographic locations

Sample	Sample group	n	n LDC	n mixed	n PRE
Kempsey LGA, Nambucca LGA only	EL&L	14	3	2	9
	Control	9	6	1	2
Kempsey LGA, Nambucca LGA, and Port Macquarie-Hastings LGA	EL&L	14	3	2	9
	Control	14	3	4	7
Kempsey LGA, Nambucca LGA, Port Macquarie-Hastings LGA and Greater Taree LGA	EL&L	14	3	2	9
	Control	11	3	2	6

Note: Long day care (LDC), stand-alone or school co-located preschool (PRE) and hubs offering a mix of services including preschool (mixed) services.

On the basis that selecting the remaining (non-EL&L) sites in Kempsey and Nambucca created an unbalanced design (control sites were lower quality and tended to be LDC rather than PRE services) and there were no replacement services should a control service refuse to participate, and because of the importance of matching on SES<sup>7</sup> between the EL&L and control sample, the option of drawing the sample from Kempsey, Nambucca, Port Macquarie-Hastings, and Greater Taree LGA's was selected.

<sup>7</sup> Both family SES and neighbourhood SES are strongly related to ECEC program quality and children's learning and development. Even though it is not reflected in the NQS ratings, it is likely the process quality in Port Macquarie is higher than in Taree. See, for example Cloney et al. (2016).

## Achieved sample

Initial community consultations were undertaken to introduce the study to local services and schools, educators, and regional government staff. In the week of 27 November 2017, ACER and ALNF visited Kempsey and Nambucca. The visit included six site visits to services and schools and a meeting with the NSW Department of Education Regional Manager for Schools. Information sheets were also handed out at sites describing the study and providing contact details for both ACER and ALNF.

All sampled sites were invited to participate. Invitation was undertaken by simultaneous email and hard-copy mail-out of a letter of invitation, plain language statement, and consent form. Follow up phone calls to centres commenced approximately two to three days after the mail-out. Of the 25 sites sampled, 22 were recruited, two intervention (EL&L) sites refused to participate (one due to recent staff bereavement and one due to perceived disruption to regular planning), and there were no replacements available as all intervention sites were invited to participate. Two control sites refused to participate (one due to participating in another study at the time, one where no reason was given), one site was not contactable, and one site was excluded due to having fewer than five children enrolled who were eligible for the study. Three of these sites were replaced with the nearest neighbour in the sampling frame (i.e., the *first* replacement). The fourth site that refused was not replaced, as the two nearest neighbours in the sampling frame (i.e., the *first* and *second* replacements) also refused to participate. Recruitment was closed as the sample quota had been met.

The final achieved sample is described in Table 5, along with the average size of the services and the SES characteristics of the location of the centre. There was good agreement in the average size of the centres, and the replacement process had produced a slight increase in the SES of the control group. The SES of the location of the control group was higher than the intervention group, although both were inside the first quartile of SES for the state of NSW (low SES).

Table 5. SES and centre size characteristics of achieved sample

Sample group	n	Mean measure of size	Median SEIFA IRSAD	Median SIEFA State Centile
EL&L	12	42	863	5
Control	10	41	929	25

The mix of NQS ratings remained similar to the distribution seen in the target sample, where proportionally more control sites were *exceeding* the NQS and more EL&L sites were *working towards* (Table 6). The replacement process had not changed this pattern.

Table 6. NQS quality ratings for the achieved sample

Sample group	n	n exceeding	n meeting	n working towards
EL&L	12	3	6	3
Control	10	5	4	1

The mix of service types was well matched between intervention and control groups. The majority of services were stand-alone preschool services (PRE). This is described in Table 7.

*Table 7. Centre types of achieved sample.*

Sample group	n	n LDC	n mixed	n PRE
EL&L	12	2	2	8
Control	10	2	2	6

Note: Long day care (LDC), stand-alone or school co-located preschool (PRE) and hubs offering a mix of services including preschool (mixed) services.

Overall, the sampling process yielded an acceptable sample of services operating in the Mid North Coast of New South Wales. The characteristics of the services were in line with expectations and met the specification of the design.

In each sampled centre, a short screener interview was conducted with the director to find out about the current enrolment, use of classrooms or groupings, and the number of educators. All classrooms or groupings of children were recruited to the study where they included at least five children (or all the children in the group, whichever was smaller) of preschool age (e.g., who would be eligible to attend school in 2019, whether their parents intend them to or not). Within each sampled room or group, all children and their main caregiver (a caregiver who lives with the child, and usually the person who drops them off to the centre, or the person who knows the most about the child and their daily routine) were invited to participate.

Table 8 summarises the final numbers of participants in the study. It is important to note that this is the achieved sample, not the count of data records. For example, not all families would complete a questionnaire, or not all children would be present on the day/s when fieldworkers visited the centres to conduct assessments. The final numbers of completed assessments, and other data records is given in the results section.

*Table 8. Summary of the achieved sample in the ODEC study*

Sample group	n centres	n directors	n rooms	n educators	n children	n caregivers
EL&L	12	12	14	14	263	263
Control	10	10	13	13	308	307
Total	22	22	27	27	571	570

## Instrumentation

This section describes the various instruments and questionnaires used to measure children's language and literacy development and the broad domains covered in each. These instruments, combined with administrative data (e.g., the sampling frame provided by the New South Wales Government), represent coverage of the key domains of measurement. Table 9 summarises the major constructs and measures used in the study. Under each domain, a number of items or sub-domains were measured, and this detail is provided in the psychometric analysis.



Table 9. Measures used in the study

Level	Domain	Example Measure
Child outcomes	Language and literacy	ELLDI*
Child	Gender	Questionnaire
	Age	Questionnaire
	Child ATSI& status	Questionnaire
	Child migration background	Questionnaire
	Child language background	Questionnaire
	Social and emotional problems	SDQ#
	Child physical or learning delay	Questionnaire
	Temperament	STSC±
Family	Family Socio Economic Status	Questionnaire
	Household income relative to need	Questionnaire
	Household composition	Questionnaire
	Main caregiver ATSI& status	Questionnaire
	Main caregiver depression	K6^
	Main caregiver migration background	Questionnaire
	Main caregiver language background	Questionnaire
Home	Indirect Learning	HLE@
	Direct Teaching	HLE@
ECEC	Pedagogy curriculum and materials	MELQO MELE+
	Hours per week observed in formal ECEC^! program	Questionnaire
	Structural quality	Questionnaire
	Hours per week in Informal Care	Questionnaire
	Attendance at EL&L Program	Questionnaire
	Classroom interactions	CLASS
Community	SES	SEIFA IRSAD~
	ECEC availability	Licensing data

\* Early Language and Literacy Development Index

# Strengths and Difficulties Questionnaire

± Short Temperament Scale for Children

^ Kessler Psychological Distress Scale (short)

@ Home Learning Environment, Niklas et al. (2016)

+ Measuring Early Learning Quality and Outcomes Measure of Early Learning Environments

& Aboriginal and/or Torres Strait Islander

! Early Childhood Education and Care

~ Socioeconomic Index for Areas (SEIFA) Index of Relative Advantage and Disadvantage (IRSAD)

## Child literacy and language assessment

A new assessment was developed to measure oral language and literacy. The assessment framework of the ACER Longitudinal Literacy and Numeracy Study (LLANS) (Khoo & Meiers, 2006, p. 265) was expanded by test developers at ACER. The new tool, the Early Language and Literacy Development Index (ELLDI) is the bespoke measure of early language and literacy skills developed for this research. It is designed to have useful properties not seen in most measures of early learning – it is developed using cutting edge psychometric methods (modern measurement theory) and is embedded in authentic interactions including picture book reading and storytelling. The measure can accurately measure individual child growth and captures a wide breadth of skills across a long development range. It is designed so children engage with new material at each assessment timepoint, well targeted to their changing level of capability.

Following the first assessment in 2018 ACER developed six new assessment booklets, two each year of the study, which included picture prompts, stories and other supporting resources to capture children’s language and literacy development over the duration of the study. The new assessment materials and resources were developed post data collection and analysis periods and were designed to build on and extend the constructs assessed in the previous ELLDI materials.

The ELLDI was designed to assess key oral language and literacy skills and knowledge, present in the big six and essential to language and literacy development. The assessment is administered using one-to-one interactions between an adult and child, using authentic resources and experiences, beginning with three practice items and completed in a reasonable timeframe (approximately 20-40 minutes). The language and literacy sub-domains initially assessed in the first year of the study are summarized in Table 11.

In subsequent years of the study each set of ELLDI assessment booklets followed a similar format. They were created using a set of new and unique items and marking guides, a set of items and marking guides taken from the most recently administered assessment booklets, as well as a set of newly developed items and marking guides that were the same across the two new assessment booklets. These sets of items and marking guides connected and progressed the oral language and literacy constructs being measured to capture children’s development. The breakdown of the language and literacy sub-domains in the final ELLDI booklets can be seen in Table 12. Each new ELLDI assessment was classified by the year it was used in the study and titled with a number and a name based on the narrative being used for the listening comprehension task in that particular assessment booklet, e.g. 2019 Book 1 Wombat Trouble (see Appendix 1: ELLDI assessment booklets for a copy of the ELLDI assessment booklets).

### *Oral language coding rubric*

Along with the set of items and marking guides that were used to assess each child's response during the one-to-one interview, each ELLDI assessment included items to assess children's oral language development. These items were audio recorded and used visual stimulus, such as picture scenes, picture books or a sequence of images to prompt children to look at and talk about aspects of the image or to tell a story. Table 10 provides an overview of the visual stimulus used during each assessment cycle. To analyse the children's expressive language responses to the visual stimulus ACER developed a set of oral language coding rubrics (for a copy of the oral language coding rubrics see Appendix 2: Oral language coding rubrics).

The oral language coding rubrics included a number of skills used by children to express their thoughts and ideas in response to the stimulus (e.g., clarity of extended speech and, structure and detail). The rubrics were designed to code and score how well each skill was being executed. The skills were broken into a set of indicators outlining a progression from limited or incomprehensible to refined use of the oral language skill. Within a skill each indicator was assigned a level, where zero was the lowest level indicator and refined use of the skill received the highest level. The number of levels across the different skills varied depending on the complexity of the skill being described (e.g., clarity of extended speech consisted of five levels, and structure and detail consisted of six levels) and the child's score represented how well they used that skill to express themselves orally. For example, a child at level three for clarity of extended speech, for the most part spoke clearly but there were a few words that were difficult to understand, such that it took some effort to interpret what they were saying.

*Table 10. Type of visual stimulus used to prompt an oral language response during each assessment cycle.*

Year	Scene	Picture book	Story sequence
2018	X	X	
2019	X	X	
2021	X		X
2022	X		

Table 11. Summary of sub-domains and items of the assessment of language and literacy in the first cycle of data collection

Sub-domain	Example item	Approx. number of items	Breadth of tasks
Expressive vocabulary	Tell me the name of this body part (point to eye).	7	Naming body parts, animals, and things you eat, talking about a picture
Phonics	This letter makes the sound /m/. Point to the letter T in Tricky. What sound does this letter make?	15	Letter naming, letter sounds
Phonemic awareness	This is Monkey's tail (point to Monkey's tail on page 5). What is the first sound in tail?	6	First, last and medial sound identification
Phonological awareness	Now we're going to talk about rhyming words. <i>Mouse</i> rhymes with <i>house</i> .	4	Rhyme
Listening comprehension	Ask: What do you think is going to happen next?	11	Responding to aural and visual stimulus
Oral language	Now please hold this picture and look at it carefully. Then I will ask you some questions about it. There are lots of people doing things. Tell me something that is happening in the picture.	6	Expressiveness, clarity, utterance, vocabulary, coherence, clarity
Print conventions	Let's look at the writing. Can you show me a word?	7	Book and word/sentence orientation
Reading	Now I'm going to read you the story. Before I start, is there anything on this page that you can read to me?	1	Decoding, word reading
Writing	Write your name at the top of this piece of paper.	4	Copying, fine motor, writing

*Table 12. Summary of sub-domains and items of the assessment of language and literacy in the last cycle of data collection*

Sub-domain	Example item	Approx. number of items	Breadth of tasks
Expressive vocabulary	There are lots of things that can take us to different places...Tell me the names of things that can take us to different places.	1	Naming different types of transport
Phonics	Look at the picture and then shade the bubble next to the word that matches the picture	8	Spelling, decoding
Phonemic awareness	Say 'can'. Now tell me the sounds in 'can'.	17	Segmenting phoneme in words, deleting and swapping phonemes
Listening comprehension	'Ella's face exploded into the biggest smile.' What does the word 'exploded' mean in this sentence?	10	Responding to literal and inferential comprehension questions
Reading comprehension	Read the short story and then shade the bubble that matches the story.	10	Read and match word to picture and sentence, read sentence/question and select the correct answer
Oral Language	Tell me about a time you did something you can see in the picture.	1	Expressiveness, clarity, extended speech, fluency, structure, relevance, articulation
Reading	Here is a picture and sentence about koalas. Read the sentence for me.	1	Decoding
Writing	I will read you the sentence and you write it on this piece of paper	1	Spelling, punctuation, handwriting

## ECEC quality

### *Room Record*

A *Room Record form* captured information during visits to early childhood day care centres and preschools. It included details of the visit (including the fieldwork team and visit times) and basic information about the room (including the number of children and educators present towards the start and end of the observation and the age of the youngest and oldest children present). In addition to this information, the Room Record form was used to record an observation of each study room. The Room Record used items selected and/or adapted from the classroom observation from the MELQO MODEL (See Appendix 3: Room Record form).

A subset of items from the MELQO Measuring Early Learning Environment (MELE) Classroom Observation Tool formed part of the Room Record form. The MELE is designed to capture key practices and activities (interactions) as well as the materials available to the children. Most items are scored on a 1-to-4 scale, with higher scores reflecting higher levels of quality. The items observed measured the aspects outlined in Table 13.

The items were based on well validated measures of both ECEC interaction quality and general environment quality. These include the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2008), the Teacher Instructional Practices and Processes System (TIPPS) (Seidman et al., 2014), and the Early Childhood Environment Rating Scale-Extension (ECERS-E) (Sylva et al., 2010).



*Table 13. Items included in the measure of quality from the MELQO MELE*

Sub-domain	MELE topic areas	MELE items
Classroom interactions and approaches to learning	Child–Teacher interactions and classroom management (all 7 items)	Item #24 Teacher engagement Item #25 Positive discipline Item #26 Negative interactions Item #27 Child behaviour Item #28 Child engagement and productivity Item #29 Grouping Item #30 Use of themes
	Individualised and inclusive environment (2 out of 4 Items)	Item #31 Individualised instruction Item #33 Gender inclusion
Classroom arrangement, space, and materials	Classroom setup and furniture (all 4 items)	Item #35 Classroom space Item #36 Classroom furniture Item #37 Interest centres Item #38 Wall displays
	Materials (all 6 items)	Item #39 Writing utensils Item #40 Art Item #41 Dramatic Play and Role Play Item #42 Manipulatives/Maths Materials Item #43 Storybooks Item #44 Number of complete books in the room (count)

## *Classroom Assessment Scoring System*

The Classroom Assessment Scoring System (CLASS) was developed to identify observable differences in the quality of interactions between teachers and the children in their class, along with the interactions between the children who play and learn together in these classes. The CLASS theoretical framework is based on developmental theory and research showing teacher-child interactions are foundational to children's learning (Pianta et al., 2008). The CLASS allows trained observers to actively analyse the quality of teacher-child interactions across ten dimensions that sit under three broad domains: Emotional Support, Classroom Organisation, and Instructional Support, using a 7-point scale (1-2 low, 3-5 medium, 6-7 high).

The Emotional Support domain observes how social-emotional relationships are formed during classroom interactions, focusing on how the teacher addresses the student's individual and collective learning needs, while supporting their desire to be heard and belong. The Classroom Organisation domain examines the orchestration of the classroom and the various learning experiences, and how this optimizes opportunities for student learning. The Instructional Support domain observes classroom interactions looking for those that provide students with deeper conceptual learning experiences. It can be seen in Table 14 the Emotional Support domain is underpinned by four related dimensions. Whereas the Classroom Organisation and Instructional Support are each underpinned by three related dimensions.

*Table 14. The three CLASS domains and associated dimensions*

Emotional Support	Classroom Organisation	Instructional Support
Positive climate	Behaviour management	Concept development
Negative climate	Productivity	Quality of feedback
Teacher sensitivity	Instructional learning formats	Language modelling
Regard for student perspective		

CLASS has been validated (Cloney et al., 2017; Hamre et al., 2013) and used extensively in research to analyse the quality of early childhood education. Here, studies show that greater emotional, organisation and instructional support can be associated with student achievement and social performance (Cadima et al., 2010; Curby et al., 2009; La Paro et al., 2004; Ponitz et al., 2009; Rimm-Kaufman et al., 2009). However, observational research indicates the quality of instructional support in ECEC classes can be considered to be low (Cadima et al. 2014), which is a common finding in CLASS research (Hamre & Pianta, 2005; Soliday et al., 2021). Nevertheless, Curby et al. (2009) established that greater levels of emotional support were associated with greater development in children's phonological awareness (Curby et al., 2009), and Picker (2022) found a relationship between the quality of teacher-student interactions in all CLASS domains and the teachers' level of reading content knowledge.

# Questionnaires

## Director questionnaire

The director questionnaire was designed to be completed by centre directors and school principals either online or as a paper-based questionnaire. The same questionnaire was provided to directors from control and intervention sites.

The questionnaire, comprising 19 items in total, covered the following topics:

- a. Demographics (gender, age and whether they identify as Aboriginal or Torres Strait Islander);
- b. The education and work experience of the director (including their highest level of education in early childhood and other fields and their hours working at the centre);
- c. The centre (including how long the centre had been open, the type of centre funding)
- d. The children at the centre (including the number of children enrolled and on the waiting list);
- e. The staff at the centre (including the number of part-time, full-time staff); and
- f. The EL&L program (including whether the director had undertaken the EL&L training and whether the program was delivered in the centre).

The purpose of most of the items was to provide context about the ECEC setting, and the leadership within it. Some additional items were also included specifically about the EL&L program (for all directors). This served two purposes, the first was to ensure the control sites had not been exposed to the intervention and to account for within-intervention variations (for example, some sites had been implementing the program for longer and had been exposed to more mentoring and training). The questionnaire primarily comprised multiple-choice items. Some open-response items were provided for directors to write numerical responses. The director questionnaire is provided in Appendix 4: Director questionnaire.

## Educator questionnaire

The educator questionnaire was an online or paper-based questionnaire designed to be completed by the lead educator of the study or classroom. In some cases, the lead educator was also the centre director, and therefore, both the educator and director questionnaires were completed. The same questionnaire was provided to educators from control and intervention sites.

The questionnaire, comprising 22 items in total, covered the following topics:

- a. Demographics (gender, age and whether they identified as Aboriginal or Torres Strait Islander);
- b. The education and work experience of the educator (including their highest level of education in early childhood and other fields and their hours working at the centre);
- c. The children in the study room (including the number of children in the room who were of non-English speaking background and the number of children that had a diagnosed developmental delay or disability);
- d. The EL&L program (including whether the educator had undertaken the EL&L training, whether the program was delivered in the centre and the frequency of using different EL&L resources); and
- e. The educator's PD (including the type, quantity and topics of any PD undertaken in the past year).

The purpose of most of the items was to provide context about the ECEC setting, and the educators within them. Some additional items specifically about the EL&L program (for all educators) were also included. This served two purposes, the first is to ensure the control sites had not been exposed to the intervention (e.g., educator may have been previously employed at an intervention site) and to account for within-intervention variations (for example, some sites had been implementing the program for longer and had been exposed to more mentoring and training). The questionnaire was primarily comprised of multiple-choice items. Some open-response items were provided for educators to write numerical responses. The educator questionnaire is provided in Appendix 5: Educator questionnaire.

### Main parent/guardian questionnaire

The parent/guardian questionnaire was a multi-modal (online, paper-based, interview, phone/supported interview) questionnaire designed to be completed by the main parent/guardian of the study child. The main parent/guardian was defined as “the person who lives with the child and knows about his/her behaviour, personality and daily care arrangements”. Where two people considered themselves the main parent/caregiver, they were asked to select one of them to participate. The same questionnaire was provided to parents/guardians of children attending control and intervention sites.

In order to maximise response rates to key items, the parent/guardian questionnaire was divided into two parts which could be completed at different times. Part one included seven key items and part two included 22 items. The same parent/guardian was asked to complete both parts of the questionnaire.

Part one of the questionnaire covered the following topics:

- a. The child's demographics (including whether their child was of Aboriginal or Torres Strait Islander origin, whether or not English was the main language spoken at home); and
- b. The child's childcare arrangements (including the different types of care and frequency used in 2018, 2017 and 2016).

Some of the items included within part two of the questionnaire were from existing validated scales and these are indicated below. Part two of the questionnaire covered the following topics:

- a. The child's, parent's/caregiver's and partner's demographics (including whether they were born in Australia);
- b. The child's home learning environment (including the child's experiences in the past week (direct and indirect teaching (Niklas et al., 2016)) and the number of books in the home);
- c. The child's temperament and behaviour (including the child's behaviour over the past 6 months, measured by the SDQ (Goodman et al., 2010) and their temperament/nature (Prior et al., 2000; Sanson et al., 1994));
- d. The child's health, learning or behavioural difficulties experienced (including whether they have had vision, hearing or other difficulties);
- e. Household characteristics (including the number of adults and children in the household);
- f. The psychological distress of the parent/guardian (the Kessler Psychological Distress Scale, K6 (Kessler et al., 2003)); and
- g. The family SES, measured by the education, employment status, and occupation of the parent/guardian and their partner (including the highest level of education completed, whether they were employed and on what basis, the occupation title), and the household income. These indicators are consistently used in large scale research (Willms & Shields, 1996).

The questionnaire primarily comprised multiple-choice items. Some open-response items were provided for parents/guardians to write short responses. The parent/guardian questionnaire is in Appendix 6: Parent/Caregiver questionnaire.

## Fieldwork

The ODEC fieldwork took place on the Mid North Coast of New South Wales in the Taree, Port Macquarie, Kempsey, and Nambucca Heads regions. The fieldwork was conducted in discrete stages across the duration of the study from 2018-2022<sup>8</sup>, with data collected during four assessment cycles and one classroom observation phase. In the first year of the study the ELLDI assessment cycle and CLASS observation phase were conducted in preschool and childcare centres. In the second year the assessment cycle took place in preschools, childcare centres, and schools. By the third year of the study, all assessments took place in schools.

---

<sup>8</sup> Due to the COVID-19 pandemic in 2020, there was no fieldwork during this year.

## ELLDI training

A mix of ACER and local staff living on the Mid North Coast with experience and qualifications involving working with children, were recruited for the fieldwork (see Table 15 for a breakdown of ODEC fieldwork team). Fieldwork cycles began with ELLDI training, either via video conference or in-person. The workshop introduced the study, its aims, the *ODEC Fieldwork Operation Manual* (Appendix 8: ODEC Field Operation Manual) and a discussion about how to use the *Room Record sheet* (Appendix 3: Room Record form). This was followed by a comprehensive dive into the ELLDI, its research base, the language and literacy concepts it assesses, along with a close look at each item, its conceptual focus and how it is scored. The fieldwork staff practised administering the ELLDI with each other and the subsequent discussion was used to moderate scoring and reinforce the importance of following the script. ELLDI training also covered logistics and concluded with a visit to an early childhood centre or school, where the fieldwork staff practiced administering the ELLDI to preschool or school aged children.

Table 15. Yearly ODEC fieldwork staff breakdown

Fieldworkers	2018	2019	2021	2022
ACER staff	4	7	2	5
Local staff	2	1	4	1

## Assessment cycle

The four assessment cycles involved fieldwork staff visiting study children while they were attending an early childhood service or school (referred to as centres), to administer the ELLDI during a one-to-one interview. Centre visits were organised in advance to confirm a day and time to visit, to check the enrolment of the children listed on the *Centre Information sheet* (see Appendix 7: Centre information sheet) and their attendance on the day of the assessment. Centres were informed of the number of fieldworkers visiting and suggestions were provided about the type and size of space needed to conduct the assessments. While assessing study children across different centres, fieldworkers adhered to centre rules and the study procedures outlined in the ODEC field operations manual. This included fieldworkers being able to adapt to the interview spaces available at centres, where in some instances due to limited space, assessments were conducted outside, in storerooms, or on the floor while another fieldworker conducted their interviews at a desk in the same confined room.

In instances where children were absent or they only attended a centre on specific days, arrangements were made to visit the centre on a second occasion in an attempt to interview as many children as possible. All assessment cycles generally took place during term one and/or term two of the school year and Table 16 provides details about the number of children interviewed during each cycle.

Table 16. Children interviewed during assessment cycles

	2018	2019	2021	2022	Total
Assessments completed	379	445	485	446	1755

During assessment visits in 2018-2019 to preschool and childcare centres the Room Record forms was used on the first visit to capture environmental information about the children, educator-child interactions, learning experiences and the availability of class resources for the children to use (see Appendix 3: Room Record form). Where two rooms from one centre participated in the study, one Room Record form was completed for each room. Observations were conducted after the fieldwork team had been in the room for a minimum of an hour. Fieldworkers in the team moderated their scores before finalising the Room Record form. The full schedule is described in the MELQO manual and in the ODEC field operations manual (see Appendix 8: ODEC Field Operation Manual).

### ELLDI moderation

To ensure quality and consistent administration of the ELLDI and reliable judgements of student performance and item scoring, moderation took place during the training, as well as across the assessment cycle in the field. The ELLDI training moderation was a collaborative process that was facilitated by a literacy and ELLDI expert. During this process the fieldwork staff were provided with child responses to each question in the ELLDI assessment booklet, which they used to score the item and then openly discuss their thinking behind the scoring judgement. Following a discussion about the scoring, fieldwork staff were provided with feedback about the assessment question, the construct being assessed and justification of the scoring. The process of moderating as a collaborative experience extended to the field, where the fieldwork staff regularly shared and discussed children's responses in order to build a common understanding about the task, variation in answers and to ensure consistent scoring of ELLDI items.

### Observation phase

Three ACER CLASS trained fieldworkers completed 24 room observations across 20 of the 22 ECEC centres from the achieved sample in December 2018. Table 17 provides a breakdown of the observations across the control and intervention centres, noting that CLASS observations were completed in two participating rooms of two intervention centres. The CLASS trained observers were randomly assigned to centres and were not privy to the centre's study status. Centre visits were scheduled over a two-week period, so the CLASS trained observer could visit two centres or rooms on one day to observe the 'regular' program for that day. The aim of the visit was to analyse the quality of the interactions between the children and the educator, as well as the interactions between the children to understand if and how the pedagogies taught during the EL&L training influenced and integrated in with other regular day-to-day experiences. During each observation the CLASS trained observer watched the classroom activities as a non-participant observer, taking notes about the instructional interactions related to each of the 10 CLASS dimensions. Details about the start and end time of the observation, the number of adults and children involved in the observation, the content focus and the grouping of the lesson were also recorded as a part of the observation. Each observation occurred in cycles of approximately 20-minutes, for four cycles. At the conclusion of a cycle the CLASS trained observer used their notes to judge the quality of the interactions and allocate a score (1-2: low; 3-5: medium; 6-7 high) to each dimension.



Table 17. CLASS observations across control and intervention centres

	Control	Intervention
CLASS trained observer 1	3	5
CLASS trained observer 2	1	8
CLASS trained observer 3	6	1

## Data processing and analysis

All data was stored in a relational database management system (RDBMS) in order to model the relationships between the various entities that make up the study (for example, relationships between parents, children, classrooms, and services). This was essential to ensure the study could track children’s variable exposure to programs over time – the purpose of this longitudinal study. To model these relationships a table object was created for each entity (e.g., “children”, “centres”, “assessment data”) in the study, and linking table objects were created to represent the relationships between objects. Each entity table object had one unique row and ID (primary key) that represented one entity in the study (e.g., one child who is participating in the study, one assessment booklet that has been filled out) and columns within the table object described the attributes of that entity (e.g., a date of birth column). Each linking table object had one row and ID (primary key) that represented one relationship between two entities. It was possible for one entity to have many relationships to other entities. For example, a child in the study might be enrolled in several rooms, or more than one parent may have signed a consent form or completed a questionnaire. For each of these relationships there was one row in the linking table object and the columns described the relationship by including the primary keys from the entity table objects (called foreign keys, in the linking table object). The linking table object also had additional columns to describe the attributes of the link (e.g., the date the link was created, whether the link was currently active). This relationship between primary keys and foreign keys within linking table objects<sup>9</sup> (and the constraints placed on them within the RDBMS) was the way referential integrity was maintained and the data modelled.

To achieve the data model, each entity was defined and a physical record associated with it. For example, children and parents were defined by the consent form they signed, and an assessment was defined by the physical assessment booklet. Each physical record had an ID (primary key – an arbitrary number if not otherwise defined in the fieldwork manual) written on it. When returned from the field, one researcher was responsible for creating one row in a table and entering sufficient data to describe the entity (e.g., centre, room). At the same time a link was made in a linking table (e.g., a consent form had information about children and one parent or multiple rooms were linked to a single centre), and this was entered in two entity table objects and each primary key was also added to one linking table object. Referential integrity was then tested by linking related entities and confirming the right number of unique entities and

<sup>9</sup> Note: where there was a strict one-to-one relationship (e.g., one assessment booklet can only be associated with one child), linking tables were not used. Instead, the child’s ID (e.g., the primary key in the child entity table object) existed in the assessment entity table object as a foreign key. This maintained the referential integrity but imposed less flexibility in the relationships that could be modelled.

relationships was reported and that correct attributes were returned (e.g., class lists of children were linked to the correct centre and rooms within that centre).

## Data entry

The process of data entry required the researcher to retrieve one physical record (e.g., an assessment book), find the primary key in the pre-filled data-entry pro forma and enter the data. This built in a check that all data was entered. Double entry of 5% of assessment booklets and scoring of the oral language analysis was undertaken as per the Data Entry Manual and checks for the quality of each process were completed. The data entry process is documented in Appendix 9: Data entry manual.

The quality of the data entry was high. For example, in the child assessment booklet, the double entry process showed that:

- a. The total rate of keystroke errors was 0.48%;
- b. The total rate of keystroke error was low-risk and overstated because most of the errors occurred in verbatim transcribing of large chunks of text and the actual error was the mis-typing of punctuation or mis-reading poor letter formation;
- c. An example of a low-risk keystroke error was for the variable “Notes”: “Interview conducted in two parts due to lock down procedure” VS “lockdown procedure”;
- d. This type of error was classified as low risk because it was not used in statistical analysis;
- e. When low-risk errors were omitted, the total rate of keystroke errors was 0.41%. This was well-below the 1% criterion;
- f. Moderate risk keystroke errors (i.e. those that are not low-risk) appeared in only 6.3% of columns reflecting that the majority of variables included no errors; and
- g. Examples of eleven variables (of 122) that contained errors is provided in Table 18

The data and double entry process was followed by data validation, whereby the variables were checked for correct value ranges, that there were no missing values and unique identifiers were unique. Data validation took place prior to any statistical analysis.

*Table 18. List examples of variables in child assessment booklets that contained keystroke errors (with count of errors and examples).*

Variable label	Total keystroke errors (count)	Example of keystroke errors
C04d	1	0 vs 99
C34a	1	11 vs 1
C07a_s	1	0 vs 1
C07j_s	1	0 vs 1
C34e	1	1 vs 0
C34m	1	1 vs 0
C33x	1	2 vs 0
C15f	1	1 vs 2
C22k	1	0 vs 1
C22m	1	1 vs 0
VOCC22t	1	990 vs 99

### Oral language analysis

The children’s oral language was analysed using their responses to the visual stimulus, which were audio recorded during the one-to-one interview and then uploaded to the cloud by fieldwork staff for secure storage. The audio files were analysed and scored at the conclusion of the fieldwork, using an ACER developed oral language rubric (see Appendix 2: Oral language coding rubrics) to understand the child’s expressive language use.

#### *Staff training*

ACER staff who conducted the oral language analysis completed three hours of training before scoring the children’s oral language audio files. In the first 2-hour session, ACER staff were trained to understand the tasks, the focus of each language skill being assessed on the corresponding rubric and the differences between scoring levels of each skill. This was followed by the coding of audio files as a team, which included discussions about any discrepancy in scoring and moderation of the teams understanding of each skill to ensure it was consistent with the indicator on the scoring rubric and the rest of the teams’ analysis. The oral language analysis training then progressed to individuals scoring of a number of audio recordings, the sharing of scores and further moderation where needed. At the conclusion of the training ACER staff were provided with access to a series of practice audio recordings to be scored prior to the second session for final moderation before the analysis. During the analysis and scoring ACER staff shared audio files to seek advice from the team about the scoring of skills that did not easily fit a coding level. This acted as further moderation. The oral language scoring rubric considered the child’s expressive language by assessing skills such as the structure and detail, relevance to the prompt and articulation.

## Analytic approach

The analysis was conducted in four distinct phases: descriptive analysis, calibration of the ELLDI measure, scaling of the ELLDI measure (conditioned on observed contextual and background data), secondary analysis (including longitudinal analysis).

### Calibration model

The EL&L outcome measure was scaled using IRT in ACER ConQuest version 5 (Adams et al., 2020). ACER ConQuest version 5 is a statistical program for fitting both unidimensional and multidimensional item response and latent regression models. It provides data analysis based on a comprehensive and flexible range of item response models (IRM), allowing examination of the properties of performance assessments, traditional assessments and rating scales. ConQuest offers analysis procedures based on the most up-to-date psychometric methods of multifaceted item response models, multidimensional item response models, latent regression models and drawing plausible values (PVs) (Adams et al., 1997).

A one parameter logistic (1PL) IRM – the many facets model (facet model) (Linacre, 1994) – an extension of the Partial Credit Model (PCM) (Masters, 1982) was fitted to the data. This model allowed the responses (at all time points) to all the items, to be decomposed into an item difficulty component (the location of the items on the oral language continuum) and some average deviation from that difficulty at each time point. That is, a concurrent calibration of all the items with multiple response vectors from each participating child (i.e. a ‘long’ format dataset).

Child responses to items were integer scored from 0 (most incorrect) to  $m$  (most correct) at each time point  $t$ . If we denote the latent ability of child  $n$  as  $\theta_n$ , and the difficulty of each item,  $i$ , is made up by the item category boundaries (e.g. the boundary between scoring 1 rather than 0, and between scoring 2 rather than 1, and so on), which have two components,  $\delta_i$  (the ‘average’ difficulty of the item), plus  $\tau_{ik}$  (the deviation from the average difficulty for this category boundary), plus the average time effect (the shift given a year of time has passed)  $\alpha_t$ , then the probability of child  $n$  scoring  $x$  on item  $i$  at time  $t$  is given by:

*Equation 1 - Probability model of the many facets item response model*

$$p(X_{nit} = x) = \frac{\exp \sum_{j=0}^x (\theta - (\delta_i + \tau_{ik} + \alpha_t))}{\sum_{k=0}^m \exp (\sum_{j=0}^k (\theta - (\delta_i + \tau_{ik} + \alpha_t)))}$$

The continuing product of the probabilities for child  $n$ 's responses to many items represents a likelihood, given their response vector and conditional on the latent variable  $\theta$ . To estimate the model by marginal maximum likelihood, the abilities of specific children,  $\theta_n$ , are integrated-out, replaced by the density  $f(\theta; \gamma, \Sigma)$ , where  $\gamma$ , and  $\Sigma$  are parameters of the multivariate normal distribution. The unknown parameters  $\gamma$ ,  $\Sigma$ ,  $\delta_i$ ,  $\tau_{ik}$  and  $\alpha_t$  are then estimated. At each stage of the study, calibration models were estimated, and the psychometric properties of the items were examined to determine which items should be retained and whether new items were needed to be added to reflect the longitudinal nature of the study. Several psychometric properties were examined to help with this process, including item facility, item fit, item-rest correlations, item characteristic curves (ICCs), and differential item functioning (DIF) (e.g., gender DIF and cycle DIF). The results of the calibration model estimated on completion of the

entire study, which includes the adequately functioning items across all cycles of the study, are presented in this report.

### Scaling model

The item parameters ( $\delta_i, \tau_{ik}$ ) from the calibration stage were taken as fixed and used as anchors in subsequent models. A four-dimensional 1PL item response model (one dimension for each time point) was estimated. That is, the time facet was removed from the specification and each time point was modelled as a separate dimension<sup>10</sup>. The removal of the time facet allows the average growth in ability to be expressed as a fixed effect in the latent abilities. As such this model captures growth over time, and the full variance-covariance matrix explaining the relationships amongst growth over time.

### Population model

Given that there was a need to generate statistics for sub-groups of the population (e.g., intervention intensity, gender, age), these group structures need to be taken into account when producing ability estimates (Wu, 2005) by including conditioning variables in the scaling model. This is done by adding a latent regression to the IRM, which includes several key regressors including child, parent, educator/teacher and director/principal variables, resulting in a population model. In order to build a population model a number of steps were taken:

1. An unconditional scaling model was estimated to produce a set of fifteen PVs.
2. The 1<sup>st</sup> vector of the set of PVs produced in step 1 were merged with the dataset and used to impute missing contextual data using a fully conditional multiple imputation model (van Buuren & Groothuis-Oudshoorn, 2011).
3. This imputed dataset was then used to estimate the population model (combine item response and latent regression model).
4. A set of five plausible values for the model estimated in step 3 were then drawn and used in secondary analysis using appropriate pooling techniques.

The contextual variables included in the population model are found in Table 19.

*Table 19. Variables included as regressors in the population model*

Stakeholder	Variable	Details
Child	ELLIntensity_2018 (1)	Description: Intervention intensity 2018 Reference: 2018 control site Contrast(s): 2018 low intensity EL&L site, 2018 medium intensity EL&L site, 2018 high intensity EL&L site
	ELLIntensity_2019 (2)	Description: Intervention intensity 2019 Reference: 2019 control site Contrast(s): 2019 low intensity EL&L site, 2019 medium intensity EL&L site, 2019 high intensity EL&L site

<sup>10</sup>This approach results in unbiased population estimates of the means, variances and covariances at each of the different time points.

Stakeholder	Variable	Details
	ELLIntensity_2021 (3)	Description: Intervention intensity 2021 Reference: 2021 control site Contrast(s): 2021 low intensity EL&L site, 2021 medium intensity EL&L site, 2021 high intensity EL&L site
	ELLIntensity_2022 (4)	Description: Intervention intensity 2022 Reference: 2022 control site Contrast(s): 2022 low intensity EL&L site, 2022 medium intensity EL&L site, 2022 high intensity EL&L site
	Gender	Description: Gender Reference: Female Contrast(s): Male
	ageAtAssessment_2018 (1)	Description: Age (in months) at the time of 2018 assessment
	ageAtAssessment_2019 (2)	Description: Age (in months) at the time of 2019 assessment
	ageAtAssessment_2021 (3)	Description: Age (in months) at the time of 2021 assessment
	ageAtAssessment_2022 (4)	Description: Age (in months) at the time of 2022 assessment
Parent	P01_2018	Description: Child's ATSI status Reference: Child is not ATSI Contrast(s): Child is ATSI
	P02_2018	Description: Child's ESL status Reference: Child has ESL Contrast(s): Child does not have ESL
	P09a_2018	Description: Parent reads to their child from a book Reference: 0 - 3 days per week Contrast(s): 1= 4 - 6 days per week, 2 = everyday
	P12_2018	Description: Books in the child's home Reference: Less than 50 Contrast(s): 1 = 50 to 100, 2 = more than 100
	P13a_2018	Description: Child's birth weight Reference: Low birth weight Contrast(s): Normal birth weight
	P13b_2018	Description: Child hearing difficulties in the past Reference: Has had hearing difficulties Contrast(s): Has not had hearing difficulties
	P15_2018	Description: Child ear infection/glue ear in the past 12 months Reference: Has had ear infection/glue ear Contrast(s): Has not has ear infection/glue ear
	P20_2018	Description: Parents highest level of education Reference: Year 12 at most Contrast(s): 1 = VET, 2 = Uni
	P21_2018	Description: Parental employment status Reference: Not employed Contrast(s): Employed

Stakeholder	Variable	Details
	P24_2018	Description: Parental relationship status Reference: Parent not living with partner or no partner Contrast(s): Parent living with partner
	P29_2018	Description: Parental household income Reference: Less than 50k Contrast(s): 1 = 50k to 100k, 2 = More than 100k
	P10_SDQ_Ext_2018	Description: Strengths and difficulties questionnaire (SDQ) - Externalising
	P10_SDQ_Int_2018	Description: SDQ - Internalising
	P11_STSC_Social_2018	Description: Short temperament scale for children (STSC) - Sociability
	P11_STSC_Flex_2018	Description: STSC - Flexibility
	P11_STSC_Persist_2018	Description: STSC - Persistence
	P19_Kessler6_2018	Description: Kessler K6 non-specific distress scale
Educator	E04_2018 (1)	Description: Educators highest level of education Reference: Year 12 at most Contrast(s): 1 = VET, 2 = Uni
	E08_2018 (1)	Description: Number of children in the class Reference: Less than 20 Contrast(s): 1 = 20 to 30, 2 = More than 30
	E13_2018 (1)	Description: Educator's awareness of ALNF Reference: Educator has not heard of ALNF Contrast(s): Educator has heard of ALNF
	E15_2018 (1)	Description: Educator's level of EL&L training Reference: Never Contrast(s): 1 = currently doing, 2 = completed
Director	D04_2018 (1)	Description: Directors highest level of education Reference: Vet Contrast(s): 1 = Bachelor, 2 = Postgrad
	D06b_2018 (1)	Description: Director's level of experience Reference: Less than 5 years' experience as a director Contrast(s): 5 or more years' experience as a director
	D15_2018 (1)	Description: Director's awareness of ALNF Reference: Director has not heard of ALNF Contrast(s): Director has heard of ALNF

*Note: Values in brackets denote that the variable was only included as a regressor for that particular timepoint. Reference and contrast categories specified for categorical variables.*

If the contextual variables that were to be used in secondary analysis were not included in the population model, it would have resulted in model mis-specification and an under estimation of the true regression coefficients (Marsman et al., 2016; Monseur & Adams, 2009; Wu, 2005).



Several raw child items were bundled to capture the age each individual child was at the time they undertook an assessment. Each child's date of birth and the date they undertook each assessment was used to produce a continuous variable for each cycle (i.e., `ageAtAssessment_2018`, `ageAtAssessment_2019`, `ageAtAssessment_2021` and `ageAtAssessment_2022`).

Several items from the parent questionnaires (relevant to a specific individual child) were bundled to create aggregate indicators of different contextual covariates included in the population model. Emotional, peer, behavioural and hyperactivity items from the Strengths and Difficulty Questionnaire (SDQ) were bundled. A sum scored 'internalising' item (`P10_SDQ_Int`) which includes the five emotional items and five peer items was calculated with scores ranging from 0-20. A sum scored 'externalising' item (`P10_SDQ_Ext`) which includes the five behavioural items and five hyperactivity items was also calculated with scores ranging from 0-20. 12 items from the Short Temperament Scale for Children (STSC) were also bundled. Using these items, three different indicators (mean scores) were produced (with four raw items in each) to represent Sociability (`P11_STSC_Social_2018`), Flexibility (`P11_STSC_Flex_2018`) and Persistence (`P11_STSC_Persist_2018`). Finally, the six raw items from the Kessler 6 non-specific distress scale were bundled using a sum score method to create a distress covariate scored 0-24. All of these bundled items were treated as continuous variables.

Items relating to the location and intensity of intervention activities that were created by ALNF were also bundled. A variable which indicates whether a centre/school was an intervention site was bundled with a variable which indicates the intensity of the intervention related activities with a centre/school. This resulted in a variable for each cycle that indicates both group membership and the intensity of activities that happened in each centre/school for that particular year (i.e., `ELLIntensity_2018`, `ELLIntensity_2019`, `ELLIntensity_2021` and `ELLIntensity_2022`). This variable was considered as categorical as outlined in Table 19.

Many more raw items included in child assessments, parent surveys, educator surveys and director surveys were recoded and entered into the population model as categorical variables. Details of these recodes can be seen in Table 19.

All secondary analysis was then run on these data sets and pooled to yield parameter estimates with unbiased (for missingness) standard errors.

## Secondary analysis

Data manipulation and integration with ACER ConQuest is undertaken in R (R Core Team, 2022) using the library `conquestr` (Cloney & Adams, 2021). Linear mixed models (LMMs) were fit to the data to estimate the growth trajectory of each child, and the influence of key demographic and contextual variables on both their initial ability and growth over time. These types of models were used to account for the complex residual variance-covariance structure in the estimation of data with repeated observations within children (O'Connell et al. 2017). Running these models for each plausible value separately and applying appropriate pooling techniques using the Multivariate Imputation by Chained Equations (MICE) package (van Buuren & Groothuis-Oudshoorn, 2022) results in unbiased population estimates of the means, variances and covariances at each of the different time points. Models were estimated using `lmer`, a function in the `lme4` library (Bates et al. 2019).

Several models were fit:

- Empty model:  $ELLDI \sim 1 + (1 | ChildID)$  to assess the suitability of random effects
- Random intercept fixed slope models:
  - Linear growth:  $ELLDI \sim Cycle + (1 | ChildID)$
  - Quadratic growth:  $ELLDI \sim (Cycle + I(Cycle^2)) + (1 | ChildID)$
  - Cubic growth:  $ELLDI \sim (Cycle + I(Cycle^2) + I(Cycle^3)) + (1 | ChildID)$
- Random intercept random slope model with cubic terms:
  - $ELLDI \sim (Cycle + I(Cycle^2) + I(Cycle^3)) + (Cycle | ChildID)$  to assess the parsimony of assumed fixed slopes
- Conditional latent growth models<sup>11</sup>:
  - Intervention model
  - Multivariate model, Intervention plus Gender, Age, ATSI, SDQ - Externalising behaviour, Parent/Caregiver Education, Parent/Caregiver Employment

To illustrate the Cubic linear growth model with effects for the EL&L intervention ("Intervention model", above):

$$Y_{ti} = \beta_{0i} + \beta_{1i}Cycle_{ti} + \beta_{2i}Cycle_{ti}^2 + \beta_{3i}Cycle_{ti}^3 + e_{ti}$$

where

$$\beta_{0i} = \gamma_{00} + \gamma_{01}Intervention_i + \mu_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}Intervention_i + \mu_{1i}$$

$$\beta_{2i} = \gamma_{20} + \gamma_{21}Intervention_i + \mu_{2i}$$

$$\beta_{3i} = \gamma_{30} + \gamma_{31}Intervention_i + \mu_{3i}$$

And

$$e_{it} \sim N(0, \sigma_e^2)$$

$$\begin{bmatrix} u_{0i} \\ u_{1i} \\ u_{2i} \\ u_{3i} \end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} & \sigma_{03} \\ \sigma_{10} & \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{20} & \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{30} & \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \right)$$

$Y_{ti}$  is the predicted ability of the child  $i$  at time  $t$  (coded 0 for Cycle 4 (2022), -1 for Cycle 3 (2021), -3 for Cycle 2 (2019) and -4 for Cycle 1 (2018) – in this way the intercept is the ability of children at the end of the study),  $\beta_{0i}$  is the random intercept made up of two

---

<sup>11</sup> Note: all conditional models are fit on the preferred growth model: random intercepts and fixed slopes.

components, the grand mean of children in the control group in Cycle 4  $\gamma_{00}$ , plus some effect for the intervention  $\gamma_{01}$  (the mean difference between intervention and control at the end of the study), and some random deviation from that mean for each child  $\mu_{0i}$ .  $\beta_{1i}$  was the linear change (random slope) in ability of child  $i$ , at time  $t$  (note also the quadratic and cubic terms).  $\beta_{1i}$  is also made up of three components, the mean growth of children in the control group  $\gamma_{10}$ , and some deviation in growth for the intervention group  $\gamma_{11}$ , (the average change in steepness in linear growth for the intervention group), and some random deviation from that mean for each child  $\mu_{1i}$ .  $e_{ij}$  is the residual term.

This approach had the benefit that the coefficients were highly interpretable. The coefficients reflected the average ending point for the two groups (Intervention and Control) and their average growth between cycles. As additional fixed parameters were added to the model, the conditional effect of these covariates are added (i.e.,  $\beta_{01}$  becomes the mean ability of children at the end of the study for the control group when all covariates are held constant at zero, and similarly  $\beta_{1i}$  is the average growth of children in the control group when all covariates are held constant at zero).

# Results

## Descriptive analysis

The following section is a descriptive analysis of the key variables collected during the fieldwork. This section excludes variables that were measured using latent measurement approaches (i.e., children's oral language and literacy abilities). All descriptive analysis including tabulations, cross tabulations, frequencies, and descriptions of univariate means and distributions was done with the cleaned data without the use of sampling weights or any handling of missing data. The classification of children, educators and directors to EL&L and control groups was based on the sampling frame<sup>12</sup>.

Please note that if values in tables do not add up to the total number stated then it is likely that some information is missing for those particular variables.

### Child characteristics

From the section on the Achieved sample, it was shown that the number of children who were drawn to participate in the study was 571. Of these, 559 completed at least one assessment during the life of the project. Background information about the children was gained from parent questionnaires in 2018. Table 20 provides a breakdown of characteristics of the children, disaggregated by EL&L and Control groups.

It can be seen that a total of 364 parent/caregiver questionnaires were returned and processed. This was a response rate of 65 per cent. A lower response rate was seen for the EL&L cohort and this warranted further follow up as planned with a physical visit to sites to conduct supported interviews in November 2018. Overall the samples were quite similar, with a slightly higher representation of children identified as Aboriginal and/or Torres Strait Islander in the EL&L group.

*Table 20. Characteristics of the children in the study*

Child characteristics	EL&L	Control
Total child participants	296	263
Returned parent/caregiver questionnaires (N)	172	192
Returned parent/caregiver questionnaires (%)	58	73
Gender (female %)	49	46
Age at 1 May 2018 (months)	53	53
Minimum age at 1 May 2018 (months)	31	26

<sup>12</sup> In later multivariate analyses, the concept of fidelity of implementation/intensity of the EL&L intervention is considered based on qualitative judgements and ALNF records. This results in, for example, one intervention site being classified as having zero intensity, and two control sites being classified as also having zero intensity, but also as having been exposed to the program but never implementing it.

Child characteristics	EL&L	Control
Maximum age at 1 May 2018 (months)	65	64
Identified as Aboriginal or Torres Strait Islander (%)	24	15
Speaks a language other than English at home	9	14
Attendance at formal ECEC (mean hours per week)	21	19
Informal care (mean hours per week)	11	4
Ear infection in last 12 months (%)	12	13
Low birthweight (% <2500g)	7	7
Social and emotional difficulties (%)	18	16

In 2018, children were, on average, exposed to more than 15 hours of formal ECEC programs per week. However, this included any attendance at family day care. It is also important to note that the distribution was skewed, meaning a significant number of children did not receive the statutory entitlement of 15 hours of pre-primary program in the year before school. In this study 38% and 30% in the control and EL&L groups respectively received less than 15 hours. This pattern is summarised in Figure 1, where a dashed vertical line indicates the threshold exposure of 15 hours per week. This is in part because of the inclusion of children younger than 4 (e.g., children in 3 to 5-year-old long day care rooms) in this study; 12% of children were not old enough to attend preschool.<sup>13</sup> Subtracting those children not eligible for preschool indicates more than 30% of children eligible for preschool in NSW did not receive 15 hours of formal ECEC program per week.

<sup>13</sup> “Children can enrol from the beginning of the school year if they turn four years of age on or before 31 July that year.” <https://education.nsw.gov.au/teaching-and-learning/curriculum/preschool/enrolment>

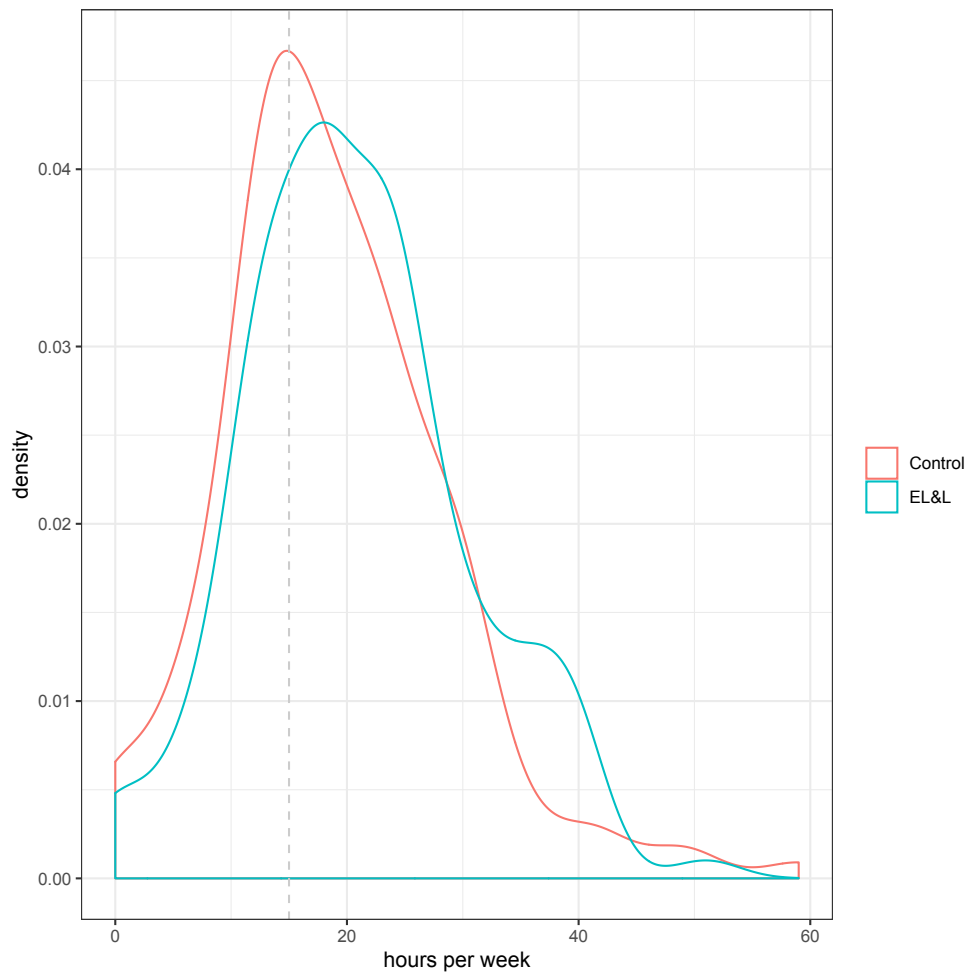


Figure 1. Distribution of children's exposure to formal ECEC programs (including home-based) by intervention group

## Family characteristics and the home learning environment

As with the children in the study, the families that participated were mostly similar (Table 21). This was especially true for education, and household size. The Control sample did appear to have higher proportion of families with many children's books (more than 50) and more main caregivers in any employment. Conversely, the EL&L sample had a higher proportion of single parent households (22% v 12%).

*Table 21. Characteristics of the main caregivers in the study and their households*

Main caregiver and family characteristics	EL&L	Control
Total participants	296	263
Returned parent/caregiver questionnaires (N)	172	192
Returned parent/caregiver questionnaires (%)	58	73
Born in Australia (%)	85	88
Less than upper-secondary education (%)	20	18
Currently employed (%)	52	55
More than 50 children's books at home (%)	61	64
Single parent household (%)	22	12
People in household (mean)	4	5

## ECEC centre and director characteristics

The response rate for directors in 2018 was very high at 100 per cent. Table 22 disaggregates the characteristics of the study centres by EL&L and control sites. The proportions should be interpreted with caution, as there was only ten and twelve control and EL&L sites respectively, and so small differences appear as large per cent differences. Overall, there was good balance between EL&L and control sites.

As expected, EL&L sites had more exposure to the EL&L intervention. Approximately 80% (compared to 0% in control sites) ran the intervention for two years or more. Nearly 80% of EL&L sites (compared to 10% in control sites) had a director working towards or completed their Certificate IV in EL&L, and a similar proportion of centres had been visited by an EL&L mentor in the last two years. This is an important demonstration of the slight blurring between the intervention and control groups – some intervention sites had exposure to the intervention. As a result, additional information was collected at the end of the study to gain a better understanding of the intensity of the intervention for each of the sites involved in the study. ALNF assigned intervention intensity ratings to each EL&L site for each year of the study (None, Low, Medium and High). Further, continuity of this intensity was also explored to try and unpack the impact that having a more intense focus on EL&L in a site for longer continuous periods might have on children's growth.



*Table 22. Characteristics of directors and services*

Director and centre characteristics	EL&L	Control
Total services	12	10
Returned director's questionnaires (N)	12	10
Returned director's questionnaires (%)	100	100
Director identifies as Aboriginal or Torres Strait Islander (%)	18	0
Director has a bachelor's degree or higher (%)	82	100
Director has more than 5 years of experience (%)	70	89
Director completed or working towards Cert IV EL&L (%)	80	10
Centre has been running EL&L for more than 24 months (%)	80	0
Centre visited by EL&L mentor in 2017 or 2018 (%)	70	10

There was good local familiarity of the EL&L program in the region, with all intervention groups and 60% of control group directors having heard of EL&L. Of these directors slightly more than 50% of the intervention sites had a director who was working towards or completed the Certificate IV whereas two directors in the control group had completed some EL&L training (none had completed the Certificate IV): see Table 23.

*Table 23. Status of Director Training in EL&L*

Group	Completed Cert. IV	Working towards Cert IV	Training other than Cert. IV	No training
Intervention	5	2	1	2
Control	0	1	1	8

In general, Directors had completed their EL&L training before 2017, with only two EL&L directors working towards their Certificate IV (Table 24) in 2018.

*Table 24. Directors' participation in EL&L professional development activities in past 12 months*

Response option	EL&L (n)	Control (n)
I have never been trained to deliver the EL&L program	2	9
I am currently being trained to deliver the EL&L program	2	1
2018	0	0
2017	1	0
2016	1	0
2015	2	0
2014	0	0
2013	0	0
2012	2	0
2011	0	0
2010 or earlier	0	0

This was also reflected in the duration that the centres had been delivering the EL&L program (Table 25). The majority of intervention sites delivered the program for more than 24 months, with one intervention centre indicating that they did not deliver the program and one control site indicating that they had delivered the program for the last 6 months. These were key variations to explore during longitudinal follow up.

*Table 25. Total months staff have delivered EL&L program in Centre.*

Response option	EL&L (n)	Control (n)
We do not deliver the EL&L program	1	9
1-6 months	1	1
7-12 months	0	0
13-18 months	0	0
19-24 months	0	0
More than 24 months	8	0

Most (70%) intervention sites indicated that they had been visited by an ALNF mentor, and one control site (10%) indicated that they had been visited by a mentor.

## Educator characteristics

Response rate for educators in 2018 was excellent with 100% of educators returning a completed questionnaire. Table 26 disaggregates the characteristics of the study rooms by EL&L and control sites. The educators in the control sites were slightly more educated (with 83% holding a bachelor's degree compared to 75% in EL&L rooms). Conversely, EL&L rooms included more children with a diagnosed disability or language delay (62% compared to 58%).

As expected, EL&L sites were more exposed, but not perfectly so to the EL&L intervention. Whilst 85% (compared to 8% in control sites) were working toward or completed their Certificate IV qualification in EL&L, and nearly 90% (compared to 30% in control sites) had been visited by an EL&L mentor in the last two years, there were some educators in EL&L sites that had experienced no training in the program.

*Table 26. Characteristics of educators and study rooms.*

Room or group characteristics	EL&L	Control
Total study rooms	14	13
Returned educator's questionnaires (N)	14	13
Returned educator's questionnaires (%)	100	100
Educator has a bachelor's degree or higher (%)	75	83
Educator has more than 5 years of experience (%)	92	92
Room includes more than 3 children with a disability or language delay (%)	62	58
Educator feels they don't have the skills/training to support all children (%)	38	45
Educator completed or working towards Cert IV EL&L (%)	85	8
Educator completed no EL&L training (%)	15	83
Educator visited by EL&L mentor in 2017 or 2018 (%)	89	30

More detailed analysis shows that there was heterogeneity with the intervention groups relative to the control group. For example, in response to a question about the extent to which they had undertaken training to deliver the EL&L program, two EL&L educators indicated they were working towards their qualification and two had received no training at all (see Table 27).

*Table 27. Completion of training in ALNF's EL&L program.*

Response option	EL&L (n)	Control (n)
Completed the Cert IV	9	0
Working towards Cert IV	2	1
Training other than Cert IV	0	1
No training	2	10

Regardless, there was good separation of the two groups, with only two control site educators indicating they had done any training in the program, and none reporting that they had completed the training. The majority of educators received their training in EL&L prior to 2016. Table 28 provides a summary of the results for both intervention and control groups.

*Table 28. Year completed ALNF's EL&L program.*

Response option	EL&L (n)	Control (n)
I have never been trained to deliver the EL&L program	2	11
I am currently being trained to deliver the EL&L program	2	1
2018	0	0
2017	0	0
2016	1	0
2015	1	0
2014	1	0
2013	0	0
2012	3	0
2011	0	0
2010 or earlier	2	0
Not answered	1	0

Results indicated that many more participants in the intervention group had undertaken EL&L training than in the control group, and that some of the participants in the intervention group had undertaken training up to and over a decade prior. As with the directors' questionnaire, many intervention educators indicated they had been delivering the program for more than 24 months, and all but one control site indicated that they did not deliver the program (see Table 29).

*Table 29. Total months of delivery of EL&L program in Centre.*

Response option	EL&L (n)	Control (n)
I have not delivered the program	2	11
1-6 months	3	1
7-12 months	0	0
13-18 months	1	0
19-24 months	1	0
More than 24 months	6	0

In terms of frequency of visits from ALNF mentors to educators, five intervention educators indicated they had not received mentoring in the past 3 years (i.e., 2015-2017). Whereas 3 control educators indicated having received mentoring in 2018. It is likely that this is due to the educators moving on from an ALNF site to their current site (Table 30).

*Table 30. Frequency of visits from ALNF EL&L mentors.*

Response option	EL&L (n)	Control (n)
2018	7	3
2017	6	0
2016	6	0
2015 or earlier	5	0

In 2018, educators were asked about frequency with which they taught generic language activities with children. Response options for the six specific activities were (5) Daily, (4) Weekly, (3) Monthly, (2) Less frequently than monthly, and (1) Never. Table 31 provides a summary of the differences in frequency between the two groups.

Table 31. Frequency with which participants are explicitly taught to...

	EL&L (Mean)	EL&L (Percent Daily)	Control (Mean)	Control (Percent Daily)
... make sounds and practice articulation?	4.75	75.00	4.08	58.33
... write letters and words?	5.00	100.00	4.33	53.85
... sound out words or syllables?	4.83	83.33	3.42	25.00
... speak in sentences (e.g., sequencing ideas)?	4.62	84.62	4.83	76.92
... identify letters by name or sound?	4.92	92.31	4.17	30.77
... identify the sounds that phonograms make (e.g., the sounds of printed or written letters)?	4.42	58.33	3.50	33.33
GRAND MEAN	4.76	82.27	4.05	46.37

Results suggested that educators in the intervention group explicitly taught children language and literacy more often than educators in the control group. The difference however was small. Of note was the largest differences related to the use of writing, sounding out syllables, and phonics. It is clear that educators were indicating that they undertook oral language and literacy instruction on most days.

When asked about specific EL&L resources the responses were expectedly different. A small number (between zero and one educators in the control group) reported using any ALNF resources (Table 32). The majority of EL&L practitioners relied mostly on Book Kits, Tommy Turtle, Sound/Spelling Sticks, and Sound Cards. Other materials were used less than half the week (e.g., on between two or three days per week). Very few practitioners used the SEAPART (assessment) in their practice. While this is not surprising, it would be expected that the results from the SEAPART would be used to inform the educators' daily practice.

*Table 32. Frequency with which participants use...*

	EL&L (Mean)	EL&L (Percent Daily)	Control (Mean)	Control (Percent Daily)
... SEAPART?	1.67	0.00	1.00	0.00
... Game Kits?	2.83	8.33	1.27	0.00
... Book Kits?	3.62	15.38	1.27	0.00
... Tommy Turtle?	4.08	46.15	1.36	9.09
... Sound/Spelling Sticks?	3.08	16.67	1.00	0.00
... Arthur Puppet?	2.58	8.33	1.36	9.09
... Sound Cards?	3.50	33.33	1.36	9.09
... Sentence Cycle?	2.17	0.00	1.00	0.00
... Sounds to Sentences?	2.58	8.33	1.00	0.00
GRAND MEAN	2.90	15.17	1.18	3.03

Both intervention and control sites undertook similar rates of professional learning and development. This was not unexpected as PD is required for ECEC educators. The largest difference between the two groups is the higher rate of professional mentoring in the intervention group – this is a specified part of the EL&L program (Table 33).

*Table 33. Participation in professional development activities in past 12 months - type*

Professional Development Activities	EL&L (percentage yes)	Control (percentage yes)
Courses or seminars	83.33	91.67
Formal qualification programme (e.g. a degree programme)	18.18	16.67
Observation visits to other early childhood centres	54.55	50.00
Formal coaching by a peer or an external person	45.45	25.00
Read professional literature (e.g., books on early childhood development)	91.67	100.00
Participation in a network of staff formed specifically for the purpose of professional development	75.00	83.33
Induction or mentoring activities	33.33	41.67
GRAND MEAN	57.36	58.33



In terms of the amount of time spent on professional learning, educators were asked to estimate the total number of hours that they had spent on PD activities in the past twelve months (i.e., during 2017) (see Table 34). Participants were provided with the following four response options: (4) More than 35 hours, (3) 16 to 35 hours, (2) 6 to 15 hours, and (1) Less than six hours. The mean for the intervention group was 2.92 while the mean for the control group was 3.17 (both close to response option 3, 16-35 hours). This indicates that educators were spending relatively little time on professional learning activities – fewer than two to five days per year.

The topics of focus in professional learning activities were mostly the same between groups, with the EL&L educators more likely to focus on oral language and literacy and less likely to focus on transitions to school. In addition, the intervention group were more likely to focus on health and hygiene issues as well as how to work with children from non-English speaking backgrounds. This might point to some relative disadvantage on the EL&L group relative to the control group.

*Table 34. Participation in professional development activities in past 12 months - content*

Professional Development Activities	EL&L (percentage yes)	Control (percentage yes)
Content related to child development and wellbeing	91.67	91.67
How to encourage play-based learning, creativity and problem solving	76.92	83.33
How to facilitate children's transitions from Preschool to Kindergarten	23.08	66.67
How to work with parents or guardians/families	41.67	50.00
How to develop children's literacy, oral language, and mathematics skills	76.92	58.33
Content related to health awareness (e.g., ear infections or diet advice)	76.92	41.67
Content related to child protection (e.g., signs of neglect, trauma, abuse)	53.85	75.00
How to work with children who do not speak English as their first language	23.08	0.00
How to support children with special needs or disability	75.00	75.00
How to support children from diverse backgrounds (e.g., multicultural, economically disadvantaged, religious)	38.46	41.67
Classroom management	23.08	54.55
Monitoring, assessing and reporting child development and learning	38.46	75.00
GRAND MEAN	53.26	59.41

## Classroom characteristics and quality

ECEC quality was measured by the CLASS, scaled according to the instrument manual (Pianta et al., 2008). Simple Ordinary Least Squares (OLS) regression models looking at the mean difference between EL&L sites and control sites were not significant: on average the quality of the sites were similar, without controlling for other factors. There were differences in the variation of the distribution of quality of intervention and control group, with EL&L sites generally showing greater variation. For example, the highest quality instructional support was seen in an EL&L site.

Figure 2 shows the distribution of observed quality. These scores were consistent with other studies in Australia, including large-scale studies that are representative of Australian classrooms. In particular, it should be noted that it is consistently found that instructional support – the use of open discussions and open activities, along with feedback and language to facilitate and stimulate deeper thinking and learning - is low in most classrooms in Australia (Tayler et al., 2013).

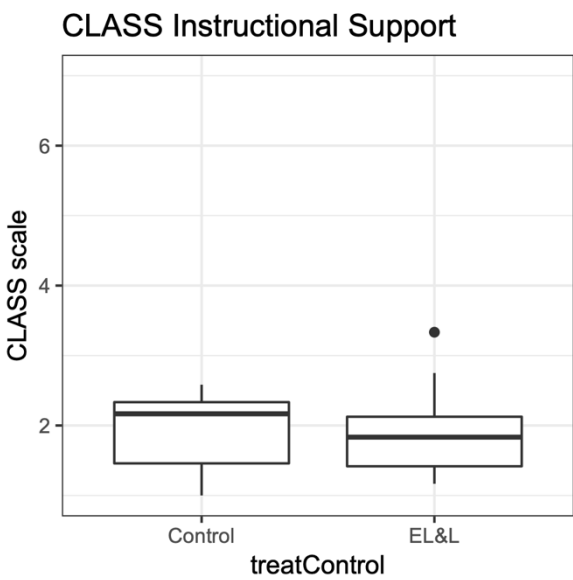
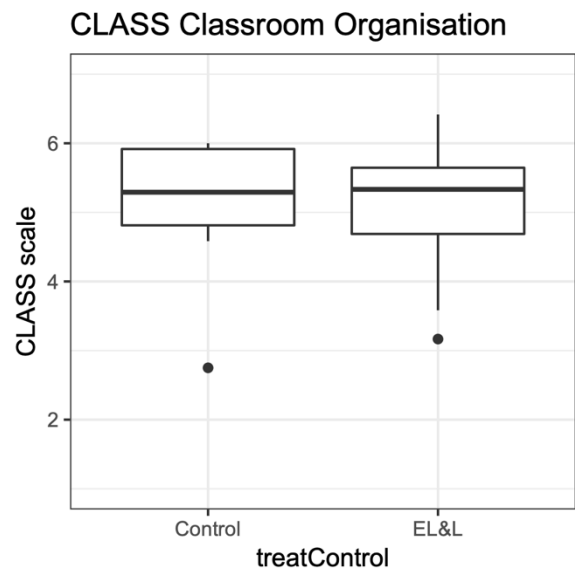
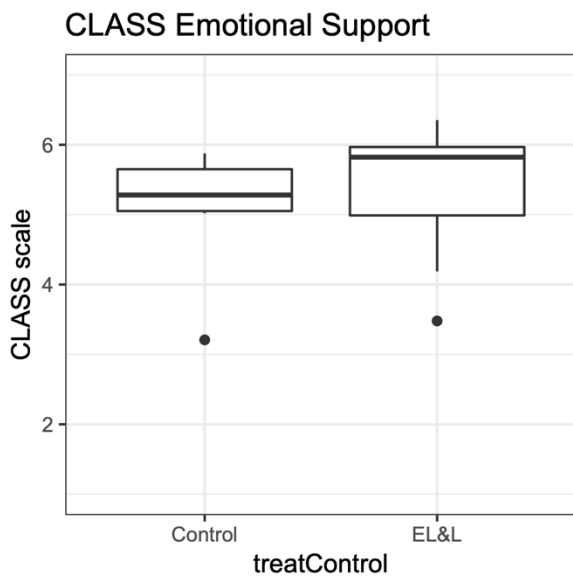


Figure 2. Boxplots of the distribution of ECEC quality

## Key contextual covariates

Frequencies for key contextual categorical covariates from the 2018 cohort can be found in Table 35. Notable and significant ( $p < 0.05$ ) differences (using a chi-square test with Yates' continuity correction) between the intervention and control groups are:

- Aboriginal and Torres Strait Islander – Intervention group has 27.9% Aboriginal and Torres Strait Islander children, control group has 16.3% Aboriginal and Torres Strait Islander children ( $x^2 = 6.8773$ ,  $df = 1$ ,  $p = 0.00873$ )
- Parents highest level of education – Parents of children in the control group had completed higher levels of education (on average) than parents of children in the intervention group ( $x^2 = 6.0468$ ,  $df = 2$ ,  $p = 0.04864$ )
- Parents relationship status – 16.8% of parents of control group children are not living with partner/no partner, compared with 32.9% of parents of intervention group children ( $x^2 = 10.308$ ,  $df = 1$ ,  $p = 0.001324$ )
- Parents household income – 24.1% of parents of control group children earn less than 50k, compared with 40.8% of parents of intervention group children ( $x^2 = 11.622$ ,  $df = 2$ ,  $p = 0.002995$ )
- Educators highest level of education - Educators of children in the control group had completed lower levels of education (on average) than educators of children in the intervention group ( $x^2 = 30.906$ ,  $df = 2$ ,  $p = 1.945e-07$ )
- Number of children in class – Children from the control groups were enrolled in larger classrooms (on average) than intervention group children ( $x^2 = 169.01$ ,  $df = 2$ ,  $p = 2.2e-16$ )
- Educators and ALNF – Educators of children in the control group were less likely (on average) to have both heard of ALNF EL&L and working towards or completed EL&L training, compared with educators of children in the intervention group ( $x^2 = 99.551$ ,  $df = 1$ ,  $p = 2.2e-16$ )
- Directors' highest level of education - Directors of children in the control group had completed higher levels of education (on average) than directors of children in the intervention group ( $x^2 = 54.112$ ,  $df = 2$ ,  $p = 1.777e-12$ )
- Directors' level of experience - Directors of children in the control group had more experience (on average) than directors of children in the intervention group ( $x^2 = 41.758$ ,  $df = 1$ ,  $p = 1.033e-10$ )
- Directors and ALNF – Directors of children in the control group were less likely (on average) to have heard of ALNF EL&L, compared with directors of children in the intervention group ( $x^2 = 127.62$ ,  $df = 1$ ,  $p = 2.2e-16$ )

Table 35. Frequencies of key contextual categorical variables (N (%))

Variable	Description	Group	0	1	2	Notes
Gender	Gender	All		265 (47.8)	289 (52.2)	1 – Female, 2 - Male
		Control		134 (48.6)	142 (51.4)	
		Intervention		131 (47.1)	147 (52.9)	

Variable	Description	Group	0	1	2	Notes
P01	Child's Aboriginal or Torres Strait Islander status	All	301 (77.8)	86 (22.2)		0 - No, 1 - Yes
		Control	159 (83.7)	31 (16.3)		
		Intervention	142 (72.1)	55 (27.9)		
P02	Child's ESL status	All		385 (98.7)	5 (1.3)	1 - No, 2 - Yes
		Control		193 (100)		
		Intervention		192 (97.5)	5 (2.5)	
P09a	Parent reads to their child from a book	All	122 (37.3)	103 (31.5)	102 (31.2)	0 - 0 to 3 days per week, 1 - 4 to 6 days per week, 2 - everyday
		Control	50 (30.9)	55 (34)	57 (35.2)	
		Intervention	72 (43.6)	48 (29.1)	45 (27.3)	
P12	Books in the child's home	All	119 (36.8)	106 (32.8)	98 (30.3)	0 - less than 50, 1 - 50 to 100, 2 - more than 100
		Control	52 (33.1)	48 (30.6)	57 (36.3)	
		Intervention	67 (40.4)	58 (34.9)	41 (24.7)	
P13a	Child had low birth weight	All		24 (7.5)	296 (92.5)	1 - Yes, 2 - No
		Control		11 (7)	146 (93)	
		Intervention		13 (8)	150 (92)	
P13b	Child had hearing difficulties in the past	All		7 (2.2)	315 (97.8)	1 - Yes, 2 - No
		Control		1 (0.6)	156 (99.4)	
		Intervention		6 (3.6)	159 (96.4)	
P15	Child had ear infection / glue ear in the past 12 months	All		47 (14.5)	278 (85.5)	1 - Yes, 2 - No
		Control		25 (15.9)	132 (84.1)	
		Intervention		22 (13.1)	146 (86.9)	
P20	Parents highest level of education	All	118 (38.6)	100 (32.7)	88 (28.8)	0 - Year 12 at most, 1 - VET, 2 - Uni
		Control	48 (31.8)	53 (35.1)	50 (33.1)	
		Intervention	70 (45.2)	47 (30.3)	38 (24.5)	
P21	Parents currently employed	All		174 (54.2)	147 (45.8)	1 - Yes, 2 - No
		Control		89 (57.1)	67 (42.9)	
		Intervention		85 (51.5)	80 (48.5)	
P24	Parental relationship status	All		241 (74.8)	81 (25.2)	1 - Living with partner (either married or not married), 2 - Not living with partner or no partner
		Control		129 (83.2)	26 (16.8)	
		Intervention		112 (67.1)	55 (32.9)	
P29	Parental household income	All	93 (32.4)	105 (36.6)	89 (31)	0 - less than 50k, 1 - 50k to 100k, 2 - More than 100k
		Control	35 (24.1)	65 (44.8)	45 (31)	
		Intervention	58 (40.8)	40 (28.2)	44 (31)	

Variable	Description	Group	0	1	2	Notes
E04	Educators highest level of education	All	30 (6)	74 (14.9)	393 (79.1)	0 - Year 12 at most, 1 - VET, 2 - Uni
		Control	30 (11.9)	34 (13.4)	189 (74.7)	
		Intervention		40 (16.4)	204 (83.6)	
E08	Number of children in the class	All	110 (22.4)	249 (50.8)	131 (26.7)	0 - less than 20, 1 - 20 to 30, 2 - More than 30
		Control	32 (12.6)	90 (35.6)	131 (51.8)	
		Intervention	78 (32.9)	159 (67.1)		
E13	Educator has heard of ALNF EL&L	All		398 (76.8)	120 (23.2)	1 - Yes, 2 - No
		Control		146 (57.7)	107 (42.3)	
		Intervention		252 (95.1)	13 (4.9)	
E15	Educator's level of EL&L training	All	268 (53.9)	54 (10.9)	175 (35.2)	0 - never, 1 - currently doing, 2 - completed
		Control	189 (74.7)		64 (25.3)	
		Intervention	79 (32.4)	54 (22.1)	111 (45.5)	
D04	Directors highest level of education	All	39 (7.3)	342 (63.8)	155 (28.9)	0 - VET, 1 - Bachelor, 2 - Postgrad
		Control		176 (63.3)	102 (36.7)	
		Intervention	39 (15.1)	166 (64.3)	53 (20.5)	
D06b	Director's level of experience	All	82 (16.2)	423 (83.8)		0 - Less than 5 years' experience as a director, 1 - 5 or more years' experience as a director
		Control	18 (6.5)	260 (93.5)		
		Intervention	64 (28.2)	163 (71.8)		
D15	Director has heard of ALNF EL&L	All		400 (77.4)	117 (22.6)	1 - Yes, 2 - No
		Control		161 (57.9)	117 (42.1)	
		Intervention		239 (100)		

Note: The labels in the variable column are used in later tables.

Descriptive statistics for key contextual continuous covariates from the 2018 cohort can be found in Table 36. Whilst there are some small differences in the means of these variables for intervention and control groups, none of them are statistically significant (0.05 level).

Table 36. Descriptive statistics of key contextual continuous variables

Variable	Description	Group	N	Mean	SD	SE	Median	Min	Max
ageAtAssessment	Age at the time of first assessment	All	367	0.00	1.00	0.05	0.09	-4.13	1.68
		Control	162	-0.10	1.11	0.09	0.09	-4.13	1.58
		Intervention	205	0.08	0.89	0.06	0.06	-3.22	1.68
P10_SDQ_Ext	Strengths and difficulties questionnaire (SDQ) - Externalising	All	327	6.37	3.59	0.20	6.00	0.00	16.00
		Control	161	6.38	3.72	0.29	6.00	0.00	16.00
		Intervention	166	6.37	3.47	0.27	6.00	0.00	16.00
P10_SDQ_Int	SDQ - Internalising	All	327	3.57	2.82	0.16	3.00	0.00	13.00
		Control	161	3.61	2.94	0.23	3.00	0.00	13.00
		Intervention	166	3.54	2.71	0.21	3.00	0.00	13.00
P11_STSC_Flex	Short temperament scale for children (STSC) - Flexibility	All	322	2.68	1.16	0.06	2.75	0.00	5.00
		Control	157	2.71	1.22	0.10	2.75	0.00	5.00
		Intervention	165	2.66	1.10	0.09	2.75	0.00	5.00
P11_STSC_Persist	STSC - Persistence	All	322	3.13	1.08	0.06	3.25	0.00	5.00
		Control	157	3.05	1.14	0.09	3.25	0.00	5.00
		Intervention	165	3.20	1.02	0.08	3.25	0.00	5.00
P11_STSC_Social	STSC - Sociability	All	322	2.70	0.99	0.06	2.75	0.25	5.00
		Control	157	2.74	1.02	0.08	2.75	0.25	5.00
		Intervention	165	2.68	0.96	0.07	2.75	0.25	5.00
P19_Kessler6	Kessler K6 non-specific distress scale	All	246	20.73	3.48	0.22	22.00	2.00	24.00
		Control	119	20.55	3.65	0.33	22.00	6.00	24.00
		Intervention	127	20.90	3.33	0.30	22.00	2.00	24.00

The differences described in this section, particularly for categorical variables, show that the intervention group of children are less advantaged in many of the key areas that are anticipated to have an impact on their learning trajectory. These variables were included in subsequent analyses and explored in more depth in the results section of the report.



# Item response theory (IRT) analysis

## Calibration

At each stage of data collection, IRT models were fit to the data and item reviews were undertaken to ensure that items with inadequate psychometric properties were removed from the analyses and assessment forms for subsequent stages of data collection. To reflect the longitudinal nature of the study, additional items of increasing difficulty (including items from new sub-domains (sub-strands)) needed to be included in the assessments. For example, items relating to book orientation, compound word awareness, reading comprehension, receptive oral language, syllables and phoneme segmentation, phoneme manipulation, and phoneme deletion were added to later rounds of the study as children's skills and abilities grow in sophistication. Following the last round of data collection, further item review activities were undertaken and a final set of adequately functioning items (N=234) from across all four rounds of the study were retained and included in the calibration. A mapping of items to the round(s) and position within booklet(s) in which they appeared can be found at Appendix 11: Item and booklet mapping. A number of key results for the final calibration model, including item Interventions, model convergence and item-level statistics, are presented below.

Several different item treatments were applied to the items included in the final calibration model, including splitting items to be cycle specific (where, for example, suspected curriculum or program effects are observed), recoding items (e.g., collapsing response categories), and item bundling (i.e., combining multiple items to form a single item). This is done to address misfit of the model to the observed data. In extreme cases items are removed from the calibration (that is removed from the assessment).

Four items required splitting due to the changes made to either the stimulus (and scoring in one case) or the items' position in the assessment in subsequent assessment forms. Each of these split items for later rounds used the same item labelling convention with “\_1” pasted on to the end (e.g., C04d became C04d\_1). The following items (and details of the reasons for splitting) are detailed here:

- C04d
  - 2018 - Tell me name of this body part (Point to **hand**) – Accepted answer: **Hand or palm**
  - 2021 - Tell me name of this body part (Point to **palm**) – Accepted answer: **Palm**
- C04h
  - 2019 – Tell me the name of this body part (**Point** to eyebrow) - Accepted answer: Eyebrow
  - 2021 – Tell me the name of this body part (**Trace finger over** eyebrow) - Accepted answer: Eyebrow
- C15c

- 2019 – Tell me what happened in the story (Wombat trouble) – **Item was positioned after a series of initial items about the book are administered.** These initial items (C11a-d and C15a) allow the respondents to turn the pages and look through the book themselves, answer a question about the title page, and read some words/sentences from the first two pages.
- 2021 – Tell me what happened in the story (Wombat trouble) – **Item was positioned at the start of the Wombat trouble section** (i.e., it was the first time looking at the book when they responded to this item).
- C35a
  - 2019 – Tell me a word that rhymes with rope – Max score: **1**
  - 2021 – Tell me a word that rhymes with rope – Max score: **2**
    - Follow up stimulus - If child identifies one word correctly say: Tell me another word that rhymes with 'rope'.

Items that were recoded can be seen in Table 37. In some cases, it might be a single response category getting collapsed into the closest category below it (e.g., reading anything on the page scoring for C15a changes from 0-3 to 0-2). In other cases, multiple response categories are collapsed (e.g., oral language scoring for *storyNarrative* changes from 0-5 to 0-3). These recoding decisions were based on the results of the item reviews. It was likely that either the fit of the item was inadequate, there were few (<10) or no cases in a particular response category, or both. The recoded items were only retained in the model if their psychometric properties improved to an acceptable level.

*Table 37. Item recodes*

item	raw code	recode
C10b	1	0
C10b	2	1
C15	1	0
C15	2	1
C15	3	1
C15a	3	2
C16	2	1
C17c	1	0
C17c	2	1
C22o	1	0
C22o	2	1
C36a	1	0
C36a	2	1
C36c	1	0

item	raw code	recode
C36c	2	0
C36c	3	1
sceneClarity_Play	2	1
sceneClarity_Play	3	2
sceneClarity_Play	4	3
sceneExpressiveness_Play	2	1
sceneExpressiveness_Play	3	2
sceneExpressiveness_Play	4	2
sceneRelevance_Water	1	0
sceneRelevance_Water	2	0
sceneRelevance_Water	3	1
sceneRelevance_Park	1	0
sceneRelevance_Park	2	0
sceneRelevance_Park	3	1
sceneRelevance_Play	1	0
sceneRelevance_Play	2	0
sceneRelevance_Play	3	1
sceneStructure_Water	3	2
sceneStructure_Water	4	3
sceneStructure_Water	5	4
sceneStructure_Park	3	2
sceneStructure_Park	4	3
sceneStructure_Park	5	4
sceneStructure_Play	3	2
sceneStructure_Play	4	3
sceneStructure_Play	5	4
storyNarrative	1	0
storyNarrative	2	1
storyNarrative	3	2
storyNarrative	4	3
storyNarrative	5	3

*Note:* For all items, 77 and 88 are recoded to 'missing' and 99 is recoded to '0'

A method of item bundling was applied with 55 raw items being bundled in different combinations, resulting in 13 bundled items. Most of these bundling decisions were based on either the design of the items, the difficulty of the items, or both. For example, C05 asks children to “Name as many things that you can eat, as you can” with each response being treated as a separate raw item C05a-C05j, resulting in a sum score that was then recoded to create a bundled item scored 0-2 (C05sum). Another example is C33j-l, which asked children “Do these words have the same last sound?”. As these items were testing the same theoretical concept, were using the same word for comparison, and were very easy, these raw items were bundled (C33jkl). One of the bundled items is a bit different to the others, where the second raw item (C47b) is dependent upon the first raw item (C47a). C47a asks children to “Write your name at the top of this piece of paper”, with C47b stating “If incorrect, describe what the child wrote. Check one circle.”. This bundled item (C47) was scored 0 – incorrect (non-name letters/symbols/scribbles), 1 – incorrect (letters in name but incorrect order or other mistakes), 2 – correct. Full details of all the bundled items can be found in Table 38.

Table 38. Bundled items

Modelled item	Original items	Score category		
		0	1	2
C47	C47a, C47b_Incorrect	Incorrect - non name letters/symbols/scribbles	Incorrect - letters in name but incorrect order or other mistakes	Correct
C05sum	C05a_Code, C05b_Code, C05c_Code, C05d_Code, C05e_Code, C05f_Code, C05g_Code, C05h_Code, C05i_Code, C05j_Code	None correct	1 to 5 correct	6 or more correct
C06sum	C06a_Code, C06b_Code, C06c_Code, C06d_Code, C06e_Code, C06f_Code, C06g_Code, C06h_Code, C06i_Code, C06j_Code	None correct	1 to 5 correct	6 or more correct
C08_Bun	C08c, C08e, C08f, C08h, C08i, C08m	0-5 correct	6 correct	
C33abc	C33a, C33b, C33c	0-2 correct	3 correct	
C33def	C33d, C33e, C33f	0-2 correct	3 correct	
C33ghi	C33g, C33h, C33i	0-2 correct	3 correct	
C33jkl	C33j, C33k, C33l	0-2 correct	3 correct	
C33mno	C33m, C33n, C33o	0-2 correct	3 correct	
C33pqr	C33p, C33q, C33r	0-2 correct	3 correct	
C34abc	C34a, C34b, C34c	0-2 correct	3 correct	
C34def	C34d, C34e, C34f	0-2 correct	3 correct	
C34mno	C34m, C34n, C34o	0-1 correct	2-3 correct	

## Model convergence

Diagnostic plots can be used to determine whether the estimated model converged adequately. IRT models estimated by marginal maximum likelihood use an iterative process where the parameters of the model are moved incrementally towards the most likely solution given the data. A stable solution is reached when the incremental change over successive iterations become increasingly small. To determine if a stable solution is reached, convergence criteria need to be specified (e.g., largest change in any parameter estimate between successive iterations is less than  $f$ ). Conservative criteria were used; estimation terminates if change in deviance or largest change in item parameter estimates is less than 0.000001. Figure 3 shows the iteration history for the item parameters ( $X_{si}$ ) (also known as item location or difficulty parameters) for the calibration model. This shows that by the accepted iteration (286 on x-axis), the model had converged on a stable solution for item parameter estimates. Model convergence plots for likelihood and variance (see Appendix 10: Model convergence plots) also show that stable solutions were reached by the accepted iteration. Appendix 10: Model convergence plots also provides the model convergence plots for the population model. These also show that from approximately iteration 70 onwards, the model had converged on a stable solution. The parameter estimates from the 1000<sup>th</sup> (and last) iteration were chosen as the accepted solution.

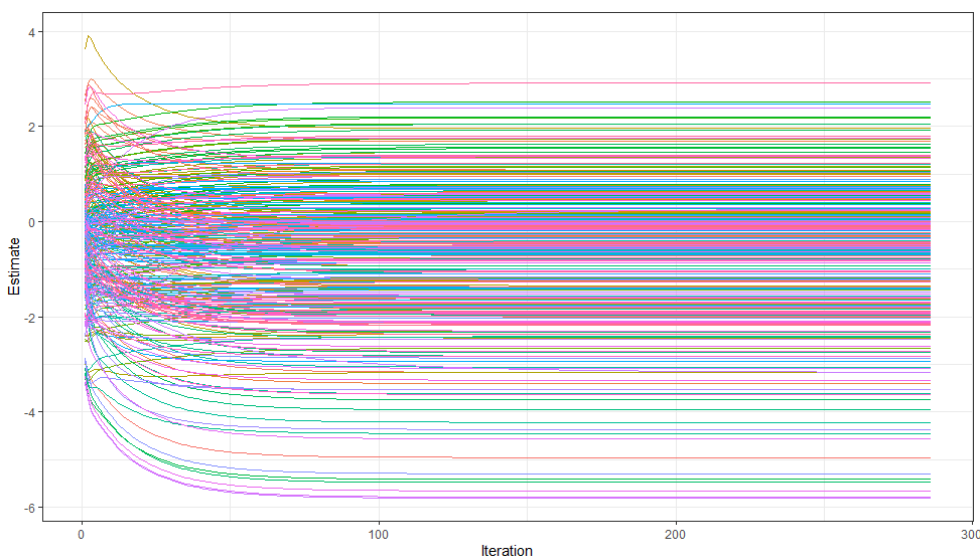


Figure 3. Iteration history for  $X_{si}$  (Item difficulty)

## Reliability

For the calibration model, person separation reliability estimates are provided. Values range from 0 to 1, with values approaching 1 indicating good reliability. The maximum likelihood estimate (MLE) for test reliability was 0.904. The Warm's mean weighted likelihood estimate (WLE: Warm, 1989) for test reliability was 0.902. A value of 0.991 for item separation was also estimated.

## Response category frequencies

The proportion of responses in the different response categories for each item are provided in Appendix 12: Raw frequencies. The frequencies are broken down by cycle, resulting in frequencies (and percentages) of cases who responded to each item by cycle combination. In all, there are 320 item by cycle combinations. The proportion of cases who respond correctly to an item is also known as item facility. Item facilities of less than 5% or greater than 95% may be indicative of an item which is too hard or too easy, respectively. Thirteen items by cycle combinations were too easy (>95% correct) and two combinations were too hard (<5% correct). In a similar fashion, for polytomous items we hope to see at least 5% of cases belonging to each of the response categories. Eight item by cycle combinations had less than 5% respond to the “0” response category and three had less than 5% respond to the “1” category. Twenty-seven other polytomous item by cycle combinations had less than 5% respond to multiple response categories. However, these items were retained as they showed adequate psychometric properties using other indicators and/or are important for assessment framework or construct coverage purposes.

## Item-rest correlation

Item-rest correlation estimates are used to indicate how well the vector of item responses is correlated with the vector of raw scores summed across the rest of the items. Values range from -1 to 1, with values greater than 0.2 generally considered adequate. A list of the items included in the final calibration model along with various item-level statistics, including item-rest correlation estimates, can be seen at Appendix 13: Item-level statistics. Values of the average item-rest correlation estimates for the final set of items in the calibration range from 0.09 to 0.72. The mean (SD) of the item-rest correlation estimates is 0.39 (0.12).

## Item fit

Item fit statistics are produced to indicate how well the model fits the data. Weighted (infit) mean square (MNSQ: Wu, 1997) estimates are used here with values approaching 1 indicative of well-fitting items, with values in the range 0.8 to 1.2 generally considered adequate. MNSQ statistics for the final set of items included in the calibration model can be seen in Appendix 13: Item-level statistics. The MNSQ values for this model range from 0.72 to 1.33. The mean (SD) of these statistics is 1.00 (0.14).

## Item characteristic curves

Item characteristic curves (ICCs) are typically used in item response models to examine the relationship between the latent ability of respondents and their probability of success on an item. In this way, it allows for degree to which the observed response align with model expectations. The empirically derived observed item characteristic curve (dashed line) is placed together with the theoretically derived modelled item characteristic curve (solid line). If these two lines are aligned, this indicates adequate the fit of the model to the data. If the observed ICC is steeper than the model ICC, this represents an over-discriminating item. Conversely, if the observed ICC is flatter than the model ICC, this represents an under-discriminating item. The ICCs for the final set of items can be seen at Appendix 14: Item characteristic curves (ICCs).

An example of an adequately discriminating dichotomous and polytomous item can be seen in Figure 4 and Figure 5, respectively. These are examples of ICCs where the probability curve based on observed responses matches model expectations.

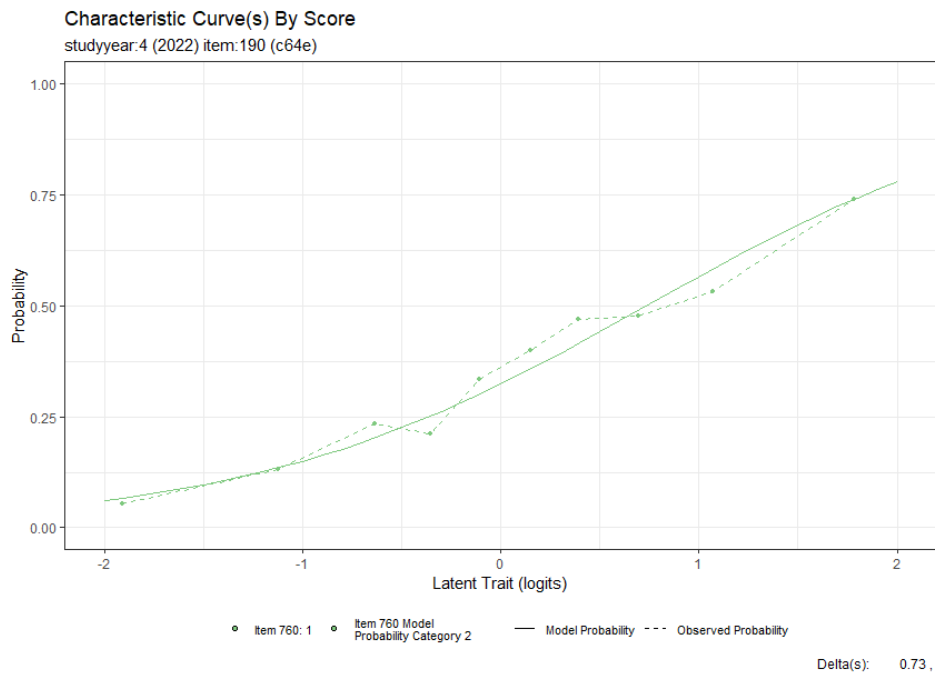


Figure 4. Adequately discriminating dichotomous item

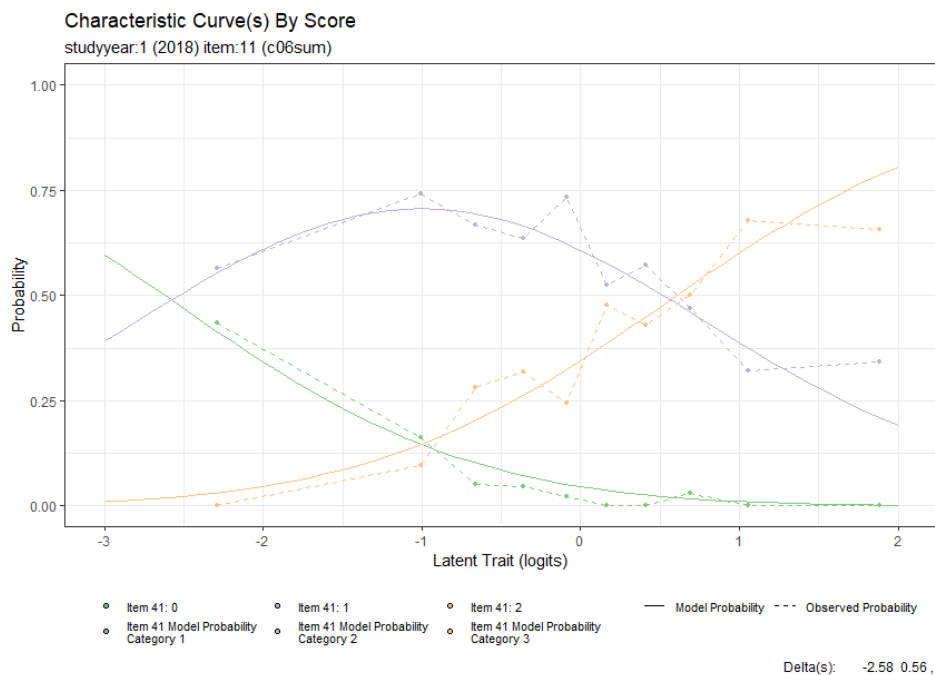


Figure 5. Adequately discriminating polytomous item

An example of an over-discriminating item is Figure 6. This shows that the item discriminates between higher and lower ability children to a greater degree than the theoretical curve based on model expectations.



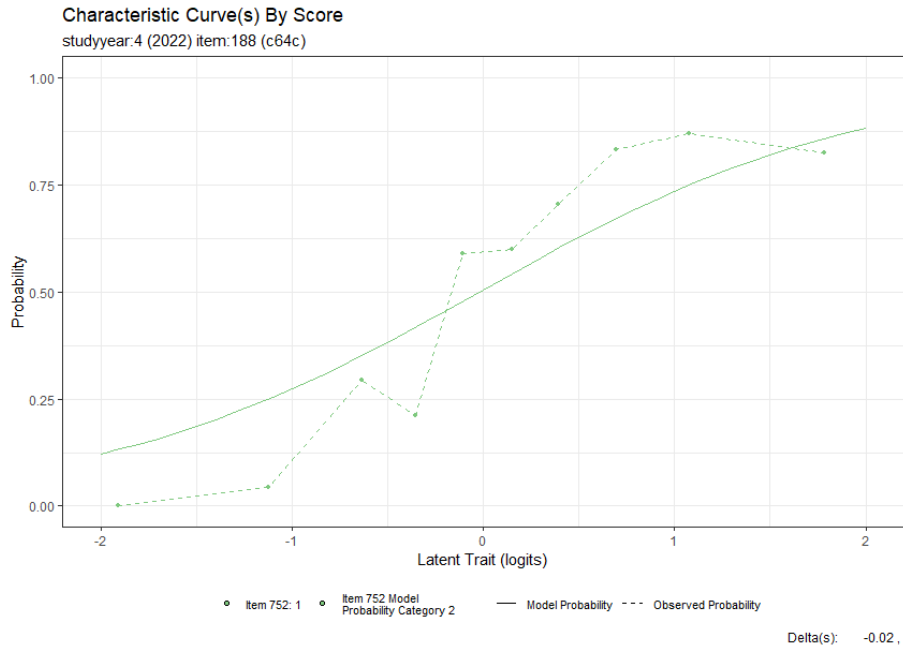


Figure 6. Over-discriminating item

An example of an under-discriminating item is Figure 7. This shows that the item does not discriminate well between higher and lower ability children (i.e., the probability of responding correctly to this item does not increase as quickly relative to ability as would be expected based on model expectations).

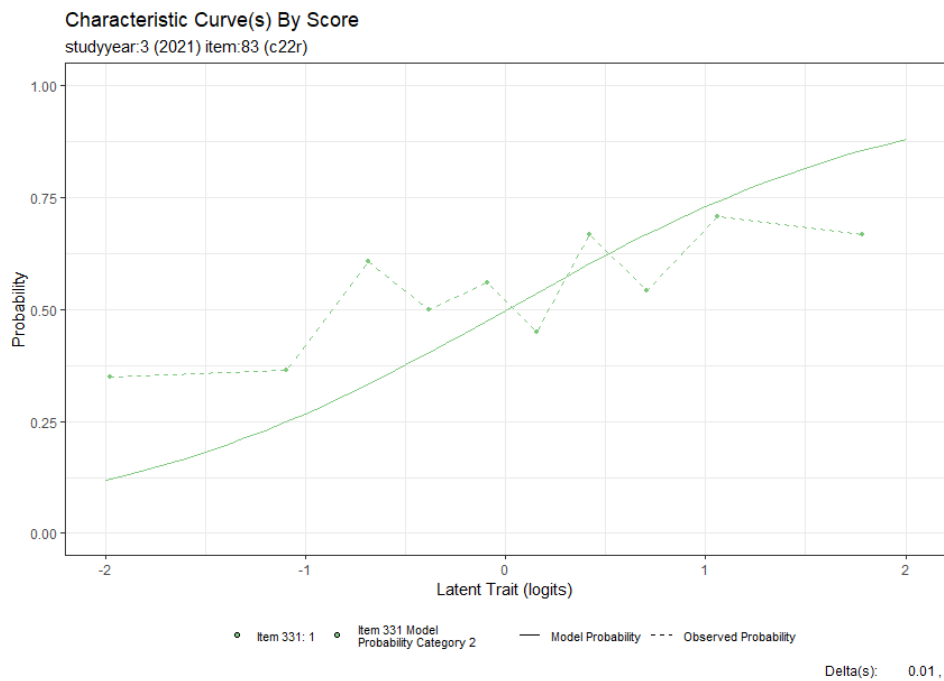


Figure 7. Under-discriminating item

## Population model

A population model is estimated by adding a latent regression model to the item response model with a number of key contextual covariates entered as regressors. The item difficulty estimates from the final calibration model are used as anchors (fixed values across all cycles). This approach results in a full latent variance/covariance matrix with means and variances estimated amongst all cycles. Therefore, plausible values produced from a fully conditional model with all covariates included in the model can be exported and used to measure the growth of children over time. A number of analytical results for the population model are presented below.

### Reliability

For the population model, expected a-posteriori/plausible value (EAP/PV: Adams, 2005) reliability estimates of 0.89, 0.87, 0.88 and 0.85 were produced for cycle 1 (2018), cycle 2 (2019), cycle 3 (2021) and cycle 4 (2022), respectively. EAP/PV reliability is a measure of the degree to which the item responses increase the certainty in the estimate of the case abilities. Estimates can range from zero to one, with values approaching one indicating good reliability.

### Test targeting

Item-person maps provide a graphical representation of the targeting of an assessment (i.e., how well the range of difficulties of items aligns with the range of abilities of children). Figure 8, Figure 9, Figure 10 and Figure 11 provide indications of the targeting of the assessment items in 2018, 2019, 2021 and 2022, respectively. Note that the plotted red points represent the average item difficult (“delta dot”), and items that receive partial credit (polytomies) span out wider than implied by these plots. These results should be taken together with the high reliability results and the later pictures of items thresholds (Figure 14) that show there is good targeting of the test to the abilities of the sample of children in the study.

The range of items included in the 2018 assessment adequately cover the range of student abilities from that cycle.

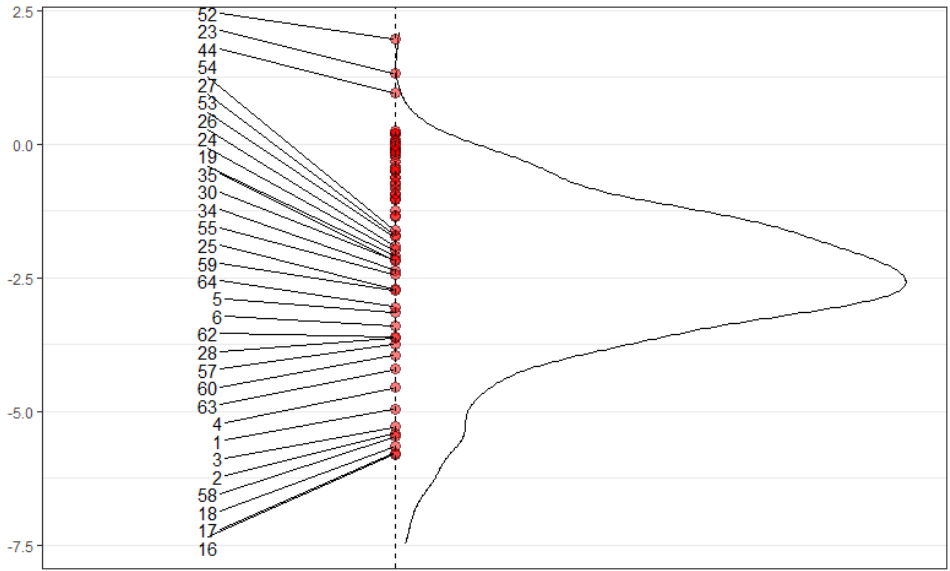


Figure 8. Item-person map for 2018

The range of items included in the 2019 assessment adequately cover the range of student abilities from that cycle.

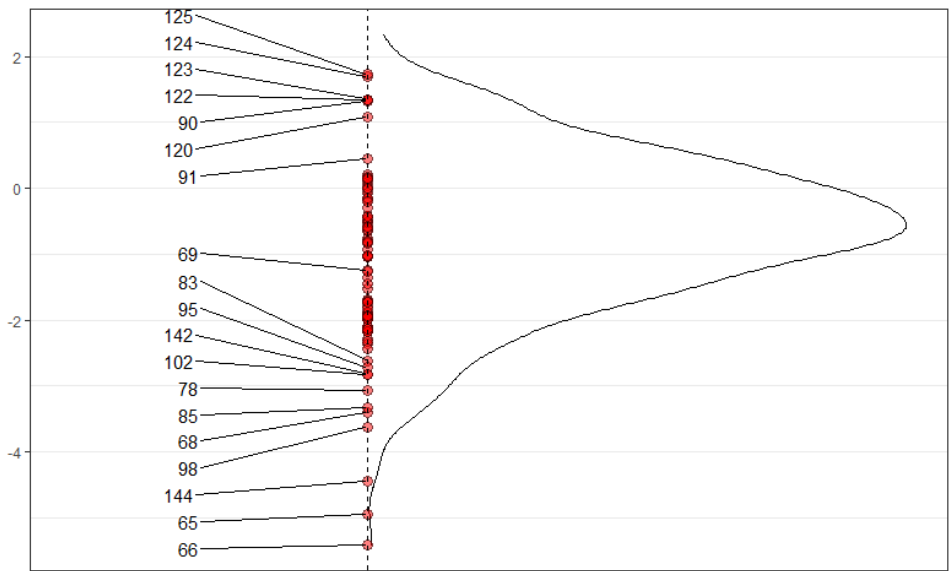


Figure 9. Item-person map for 2019

The range of items included in the 2021 assessment mostly cover the range of student abilities from that cycle. However, there are fewer item delt dots targeted to the highest ability children in the study.

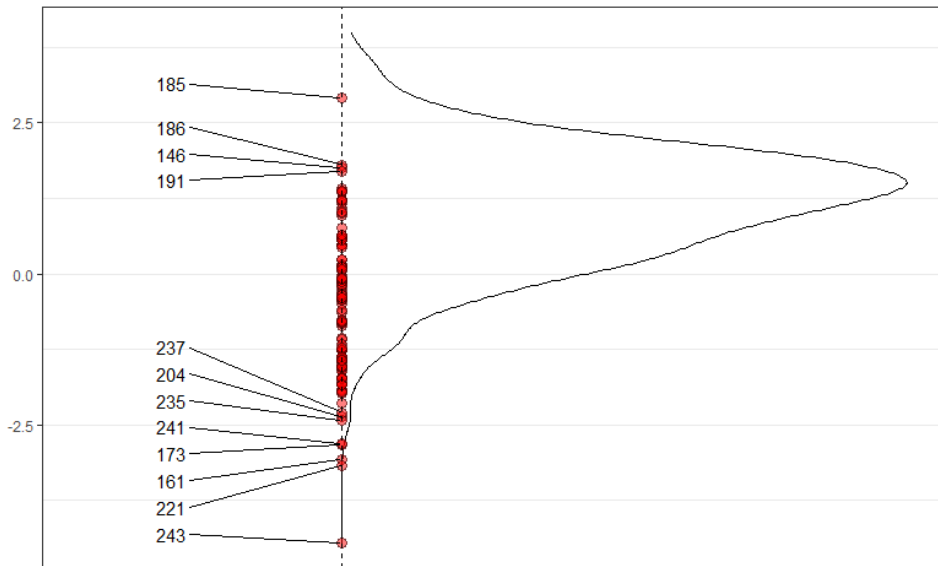


Figure 10. Item-person map for 2021

The range of items included in the 2022 assessment mostly cover the range of student abilities from that cycle. However, there are fewer items targeted to the highest ability children in the study.

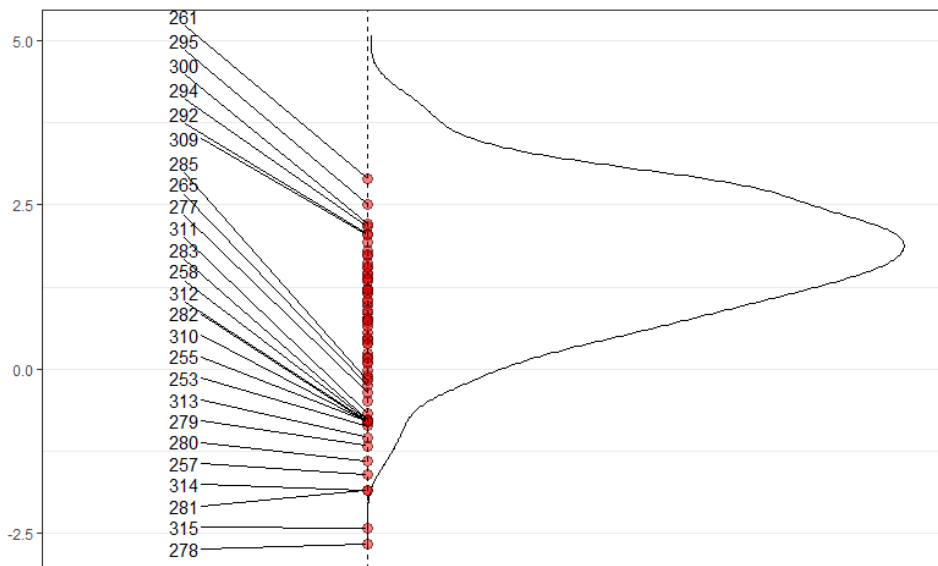


Figure 11. Item-person map for 2022

### Correlations, covariances, and variances

Unconditional correlations between cycles (lower left of the diagonal), covariances between cycles (upper right of diagonal) and variances of cycles (last row) can be seen in Table 39. Correlations between cycles are reasonable, with the estimates decreasing as the time between cycles increases (e.g., largest correlation between 2021 and 2022 (only one year apart) and lowest correlation between 2018 and 2022 (four years apart)).

Table 39. Unconditional correlation/covariance/variance matrix

Cycle	(1)	(2)	(3)	(4)
(1) 2018		1.24	0.95	0.89
(2) 2019	0.76		0.99	0.84
(3) 2021	0.68	0.86		0.80
(4) 2022	0.59	0.68	0.75	
Variance	1.98	1.34	0.99	1.15

### Multivariate analysis

Of particular interest is the average achievement of children who had an ear infection/glue ear in the year prior to the start of the study compared with the rest of the cohort. A comparison between these two sub-groups of the sample across each timepoint can be seen in Table 40. This shows that children with an ear infection/glue ear prior to the start of the study had slightly higher ELLDI achievement. However, this gap closed by 2019 and then for the remainder of the study those same children achieved lower ELLDI scores on average compared with their peers.

Table 40. Average oral language and literacy (ELLDI) achievement of children with ear infection/glue ear compared with the rest, by year

Group	Year			
	2018 Mean (SD)	2019 Mean (SD)	2021 Mean (SD)	2022 Mean (SD)
Ear infection/glue ear (N=47)	-2.32 (1.21)	-0.51 (1.17)	1.19 (0.92)	1.69 (1.06)
Did not have ear infection/glue ear (N=278)	-2.39 (1.32)	-0.48 (1.10)	1.33 (0.98)	1.89 (1.05)

The latent regression estimates produced from the population model can be seen in Table 41. A number of regression estimates show statistically significant (at the 0.05 level) differences between reference groups and contrast groups. A summary of key findings is presented here:

- Intervention group membership
  - There were significant differences for all of the contrast groups in 2018 relative to the reference group (control group)
    - Negative: ‘Intervention – Low Intensity’
    - Positive: ‘Intervention – No Intensity’, ‘Intervention – Medium Intensity’, ‘Intervention – High Intensity’
  - These differences reduced in subsequent cycles with only two estimates statistically significant (negative)
    - ‘Intervention – Low Intensity’ in 2019 and ‘Intervention – High Intensity’ in 2022

- Gender
  - Boys demonstrated lower average ability than girls in the first two cycles (2018 and 2019), with the gap closing by the later cycles (2021 and 2022)
- Age
  - Older children demonstrated higher average ability than younger children in the 2018, 2019 and 2021 cycles, but this effect decreased over time to the point where there was no difference in 2022
- Aboriginal and Torres Strait Islander
  - There were statistically significant differences between Aboriginal and Torres Strait Islander and non-Aboriginal and Torres Strait Islander children (lower) in 2019, 2021 and 2022
- Language other than English
  - LOTE children demonstrated lower average ability compared with non-ESL children in 2021
- Parent reads to their child from a book
  - In 2018, children whose parents read to them everyday had higher average ability than children whose parents read to them 0-3 days per week
- Books in the home
  - In 2018, children with 50 to 100 books in the home demonstrated lower average ability than children with 0-50 books in the home
  - In 2019, children with more than 100 books in the home had higher average ability than children with 0-50 books in the home
- Low birth weight
  - Children with low birth weight had higher average ability than children who did not have low birth weight in 2019
- Hearing difficulties
  - Children with hearing difficulties in the past had higher average ability than children who did not have hearing difficulties in the past in 2018
- Parents level of education
  - Children whose parents had university level education had higher average ability than children who had parents with Year 12 at most in 2019 and 2022
- Parents employment status
  - Children whose parents were employed had higher average ability than children whose parents were unemployed in 2018 and 2022
- SDQ

- Children who had higher scores (of more concern) on the externalising scale demonstrated lower average ability than children with lower scores in 2021 and 2022
- Children who had higher scores (of more concern) on the internalising scale had higher average ability than children with lower scores in 2018, 2019 and 2022
- STSC
  - Children who had higher scores (more desirable temperament or nature) on the sociability scale had higher average ability than children with lower scores in 2019, 2021 and 2022
  - Children who had higher scores (more desirable temperament or nature) on the flexibility scale had higher average ability than children with lower scores in 2018 and 2019
- Kessler 6
  - Children who had higher scores (more desirable feelings) on the Kessler 6 scale had higher average ability than children with lower scores in 2021
- Educators highest level of education
  - Children who had an educator with either VET or University level qualifications had higher average ability than children with an educator with Year 12 at most in 2018
- Children in the class
  - Children who were enrolled in a class with either 20-30 children or more than 30 children had higher average ability than children who were enrolled in a class with less than 20 children in 2018
- Educators' awareness of ALNF
  - Children who had an educator who had heard of ALNF demonstrated lower average ability than children with an educator who had not heard of ALNF in 2018
- Educator completed EL&L training
  - Children who had an educator who was either currently doing or completed EL&L training demonstrated lower average ability than children with an educator who had never completed EL&L training in 2018
- Directors highest level of education
  - Children who had a director with either Bachelor or Postgraduate level qualifications had higher average ability than children with a director with a VET qualification in 2018

Table 41. Regression estimates from population model

	2018			2019			2021			2022		
	Est	SE	t	Est	SE	t	Est	SE	t	Est	SE	t
(Intercept)	-6.72	0.65	-10.38	-1.37	0.54	-2.52	0.05	0.53	0.10	0.75	0.58	1.28
ellintensity_2018_1	0.66	0.19	3.44									
ellintensity_2018_2	-0.69	0.26	-2.63									
ellintensity_2018_3	1.52	0.19	7.89									
ellintensity_2018_4	0.78	0.19	4.03									
ellintensity_2019_1				0.03	0.16	0.19						
ellintensity_2019_2				-0.45	0.13	-3.60						
ellintensity_2019_3				-0.48	0.26	-1.84						
ellintensity_2019_4				0.08	0.12	0.68						
ellintensity_2021_1							0.10	0.34	0.29			
ellintensity_2021_2							0.13	0.13	1.00			
ellintensity_2021_3							-0.06	0.20	-0.32			
ellintensity_2021_4							-0.19	0.11	-1.69			
ellintensity_2022_1										0.03	0.44	0.06
ellintensity_2022_2										-0.10	0.12	-0.79
ellintensity_2022_3										-0.17	0.44	-0.39
ellintensity_2022_4										-0.26	0.13	-1.99
gender_imp_1	-0.32	0.08	-4.19	-0.21	0.08	-2.58	-0.11	0.08	-1.36	-0.16	0.08	-1.87
ageatassessment_2018_imp	0.66	0.04	16.65									
ageatassessment_2019_imp				0.49	0.04	12.22						
ageatassessment_2021_imp							0.23	0.04	5.89			
ageatassessment_2022_imp										0.00	0.04	-0.10
p01_2018_imp_1	0.14	0.10	1.43	-0.37	0.10	-3.74	-0.33	0.10	-3.46	-0.34	0.10	-3.33
p02_2018_imp_1	0.19	0.29	0.65	0.24	0.31	0.80	0.75	0.30	2.54	-0.28	0.33	-0.85
p09a_2018_imp_1	0.18	0.10	1.90	-0.04	0.10	-0.35	0.04	0.10	0.37	0.19	0.11	1.83
p09a_2018_imp_2	0.27	0.11	2.46	0.13	0.11	1.16	0.07	0.11	0.61	0.09	0.12	0.71
p12_2018_imp_1	-0.28	0.10	-2.83	0.02	0.10	0.19	0.00	0.10	0.04	-0.02	0.10	-0.20
p12_2018_imp_2	0.19	0.12	1.65	0.27	0.12	2.29	0.10	0.12	0.89	0.20	0.12	1.61
p13a_2018_imp_1	0.05	0.12	0.40	-0.29	0.13	-2.28	-0.06	0.12	-0.48	-0.04	0.14	-0.26



	2018			2019			2021			2022		
	Est	SE	t	Est	SE	t	Est	SE	t	Est	SE	t
p13b_2018_imp_1	-0.58	0.20	-2.89	-0.29	0.21	-1.39	-0.22	0.20	-1.08	0.30	0.22	1.38
p15_2018_imp_1	0.19	0.10	1.84	0.02	0.11	0.21	-0.01	0.10	-0.05	-0.06	0.11	-0.52
p20_2018_imp_1	0.18	0.09	1.89	0.05	0.09	0.56	0.07	0.09	0.77	0.00	0.10	-0.01
p20_2018_imp_2	0.18	0.12	1.55	0.29	0.12	2.39	0.21	0.12	1.80	0.25	0.13	1.99
p21_2018_imp_1	0.29	0.08	3.45	0.12	0.09	1.39	0.10	0.08	1.23	0.27	0.09	2.84
p24_2018_imp_1	0.01	0.11	0.07	0.13	0.11	1.17	0.13	0.10	1.22	0.06	0.11	0.49
p29_2018_imp_1	0.06	0.12	0.52	-0.01	0.12	-0.05	-0.08	0.11	-0.73	-0.15	0.12	-1.19
p29_2018_imp_2	0.06	0.14	0.41	0.06	0.14	0.47	0.01	0.13	0.10	0.14	0.14	0.95
p10_sdq_ext_2018_imp	0.00	0.01	0.07	-0.02	0.01	-1.69	-0.06	0.01	-4.08	-0.04	0.01	-2.81
p10_sdq_int_2018_imp	0.04	0.02	2.47	0.04	0.02	1.99	0.03	0.02	1.72	0.07	0.02	3.59
p11_stsc_social_2018_imp	0.04	0.04	1.20	0.11	0.04	3.00	0.11	0.04	3.06	0.13	0.04	3.25
p11_stsc_flex_2018_imp	0.09	0.04	2.01	0.12	0.04	2.70	0.00	0.04	0.05	0.07	0.05	1.53
p11_stsc_persist_2018_im	0.04	0.04	0.90	0.02	0.05	0.37	0.01	0.04	0.11	-0.02	0.05	-0.33
p19_kessler6_2018_imp	0.01	0.01	1.23	0.01	0.01	1.01	0.03	0.01	2.18	0.02	0.01	1.68
e04_2018_imp_1	1.59	0.26	6.19									
e04_2018_imp_2	1.76	0.21	8.45									
e08_2018_imp_1	0.75	0.17	4.38									
e08_2018_imp_2	1.18	0.16	7.53									
e13_2018_imp_1	-0.46	0.16	-2.94									
e15_2018_imp_1	-0.65	0.20	-3.19									
e15_2018_imp_2	-0.43	0.16	-2.67									
d04_2018_imp_1	1.26	0.24	5.18									
d04_2018_imp_2	1.06	0.23	4.63									
d06b_2018_imp_1	-0.12	0.16	-0.73									
d15_2018_imp_1	-0.01	0.15	-0.05									

Note: t-statistic values >1.96 and <-1.96 indicate positive and negative statistically significant differences of the contrast groups relative to the reference groups, respectively.

## Longitudinal analysis

The focus of secondary analysis is to build towards a growth model that adequately reflects the trajectories of children in the study and explores which key contextual covariates of interest impact the growth of sub-groups of the sample. Initially, an unconditional latent growth model is estimated which establishes the most appropriate model to build on for conditional latent growth modelling.

As a first step, it is useful to visualize the distribution of abilities of the sample at each cycle (see Figure 12). This shows that while the average abilities of the children included in the study increases over time, there is significant overlap in the distributions across the cycles. For example, the highest ability children in 2018 are at or above the lowest ability children in 2022 on the ELLDI Scale.

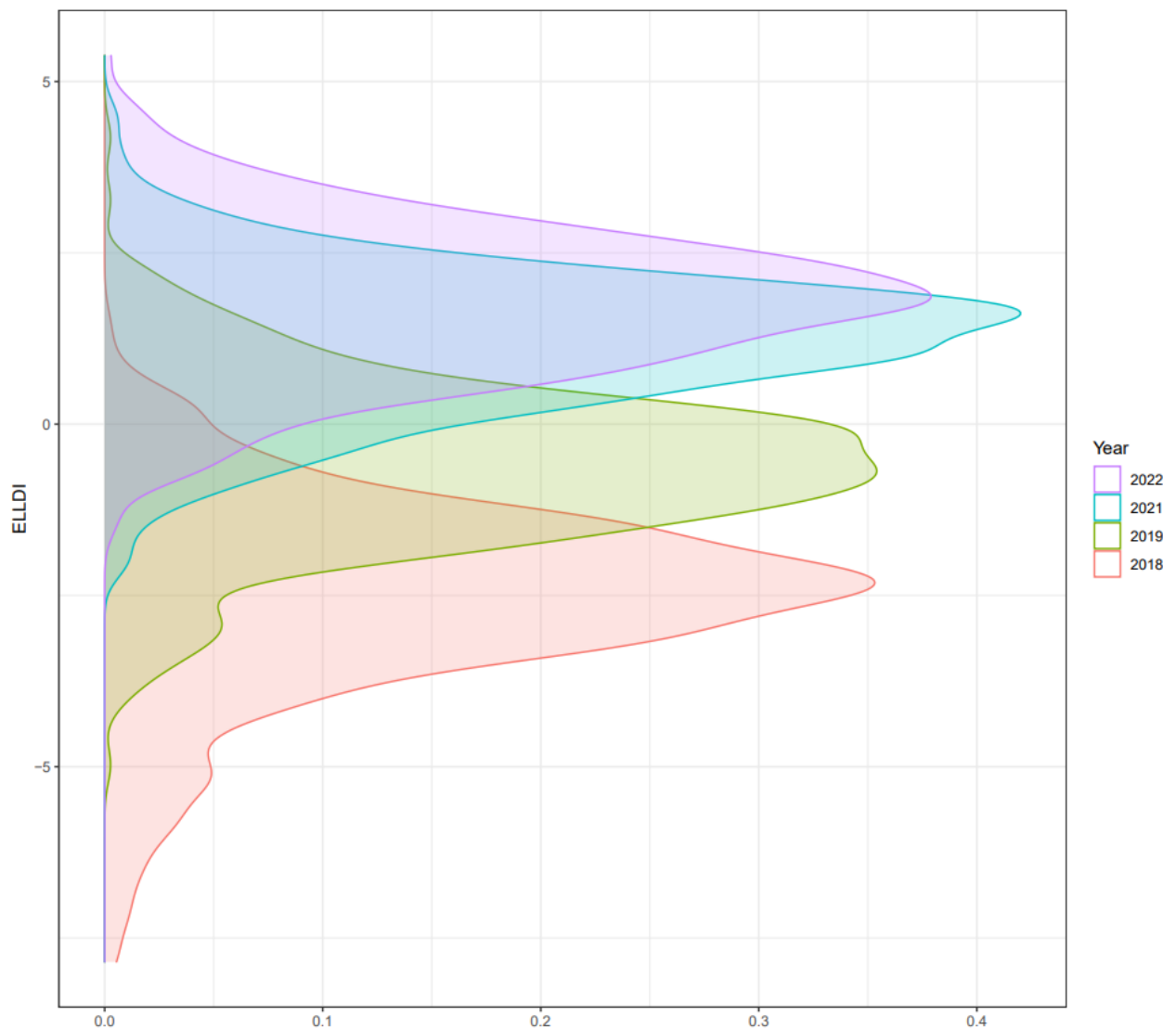


Figure 12. Density plot of abilities by cycle

The average ability (aggregate of fifteen plausible values) of the cohort of children (N=558) from each cycle (in logits) is presented in Table 42. This shows that children (on average) grew by 1.96 logits (1.53 SDs based on the pooled SD between 2018 and 2019) between 2018 to 2019, 1.79 logits (1.67 SDs based on the pooled SD between 2019 and 2021) from 2019 to 2021 and 0.58 logits (0.56 SDs based on the pooled SD between 2021 and 2022) from 2021 to 2022. The average growth from the beginning to the end of the study was 4.33 (3.49 SD based on the pooled SD between 2018 and 2022).

*Table 42. Mean abilities of children across cycles (logits)*

Cycle	Mean	SD	Variance
2018	-2.58	1.41	1.98
2019	-0.62	1.16	1.34
2021	1.17	0.99	0.99
2022	1.75	1.07	1.15

In order to compare the average ability of children from the intervention and control groups across each cycle (in logits), ordinary least squares regression was used. Taking the set of plausible values drawn from the population model and using appropriate pooling techniques a model was fit to each timepoint with a categorical predictor of intervention or control group membership along with gender and age at assessment included as covariates in the model. The conditional means and variances can be used to test the difference between the groups at each time point (see Table 43). As can be seen from this table, the intervention group started behind the control group and caught up in 2019 after a year of exposure to the program. The difference between the groups reemerges 2 years later with the control group demonstrating higher oral language and literacy skills. The gap between the groups remains lower at the end of the study than it was at the beginning in the study.

Taking the unconditional means and variances from Table 42, it is possible to contextualise this finding. Taking the mean difference between cycles and pooling the variance, gives an approximate metric of expected growth, such that the catch up observed between 2018 and 2019 represents approximately one month of growth (a value-add).

*Table 43. Mean abilities of children by intervention and control group and cycle.*

	2018	2019	2021	2022
Control	-2.323	-0.405	1.373	1.968
Intervention	-2.540	-0.517	1.211	1.78
Delta	-0.217	-0.112	-0.162	-0.188
SE	0.096	0.093	0.077	0.083
p	0.031	0.193	0.043	0.031

### Unconditional latent growth model

In building a growth model the simplest model is first estimated (empty model) and subsequent models are estimated with increasing complexity until parsimony is achieved (simplest model with adequate explanatory power). The process of comparing models and selecting the most parsimonious involves estimating ANOVAs for competing nested models, examining the variance of residuals for individual models (lower variances are equivalent to

more explanatory power), comparing the model-implied means with the actual means at each cycle, and then visually inspecting whether the selected type of growth aligns well with the change in abilities over time for each child in the study.

The first step is to compare the empty model (ELLDI ~ 1 + (1 | ChildID)) with a random intercept fixed slope model (ELLDI ~ Cycle + (1 | ChildID)). The ANOVA statistic comparing these models is 8673.43 ( $p < 0.00$ ), showing that the inclusion of the random intercept term in the model improved the fit of the model to the data. Additionally, the variance of the residuals for the empty model and the random intercept fixed slope model were 4.218 and 0.589, respectively. This shows a large increase in explanatory power from the simpler to more complex model.

The next step is to determine whether growth should be modelled as linear or otherwise (e.g., quadratic, cubic). The variance of residuals for the linear growth model, quadratic growth model and cubic growth model can be seen in Table 44. This shows that the fit of the cubic growth model is the best with adequate improvement in explanatory power based on the variance of the residuals.

*Table 44. Comparing growth models*

Model	Variance of residuals	ANOVA statistic (p-value) - comparing current model with previously accepted model
Linear	0.589	
Quadratic: (ELLDI ~ (Cycle + I(Cycle <sup>2</sup> )) + (1   ChildID))	0.431	330.851 ( $p < 0.00$ )
Cubic: ELLDI ~ (Cycle + I(Cycle <sup>2</sup> ) + I(Cycle <sup>3</sup> )) + (1   ChildID)	0.413	63.090 ( $p < 0.00$ )

It also important to compare the model-implied means with the actual means to ensure that the results are comparable. As can be seen in Table 45, the model-implied means from the estimated cubic model are equal to the actual means (to two significant figures).

*Table 45. Comparing actual means with model-implied means by cycle*

Cycle	Actual	Model-implied
2018	-2.58	-2.58
2019	-0.62	-0.62
2021	1.17	1.17
2022	1.75	1.75

A visual inspection of a cubic curve against the individual abilities estimates of children over time (see Figure 13) also shows good alignment.

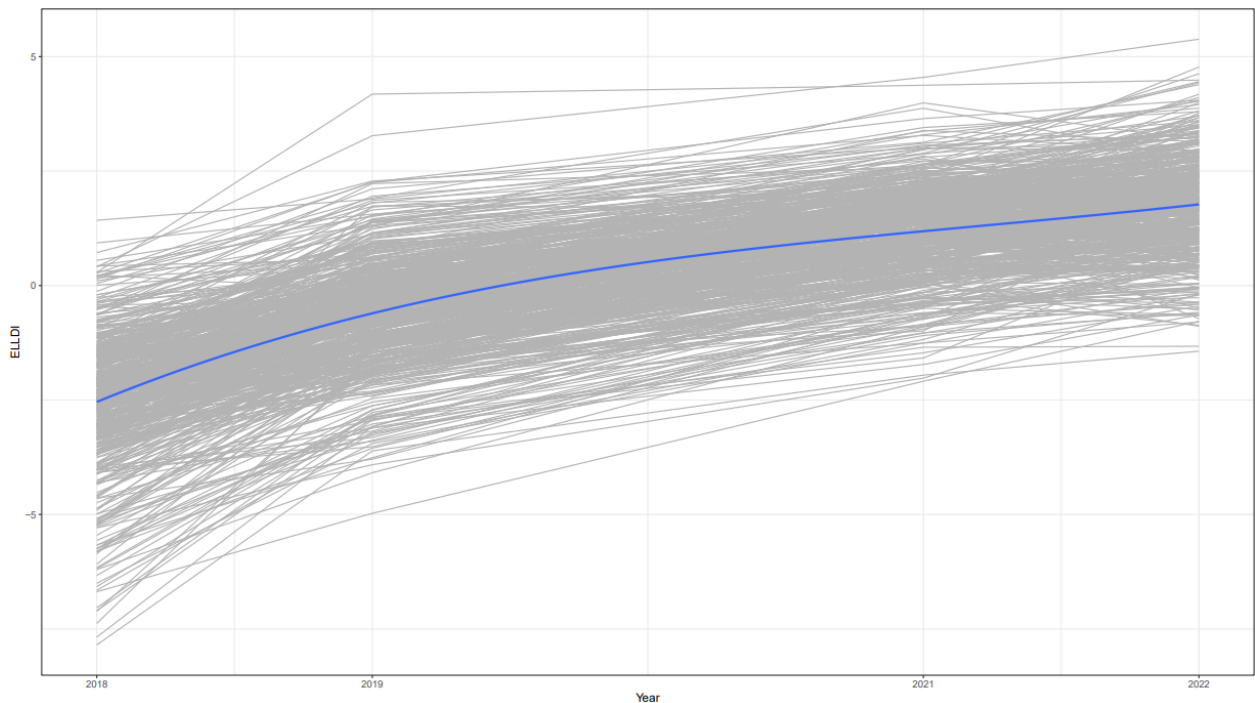


Figure 13. Cubic growth curve (blue) plotted over individual child abilities (piecemeal growth, in grey) over time

Finally, a model which includes a random slope term ( $ELLDI \sim (\text{Cycle} + I(\text{Cycle}^2) + I(\text{Cycle}^3)) + (\text{Cycle} | \text{ChildID})$ ) is estimated and compared to the cubic growth model with a fixed slope. The model with a random slope term had better fit to the data (ANOVA estimate of 68.920 ( $p < 0.00$ )) and better explanatory power (variance of residuals = 0.286). Therefore, a random intercept with cubic growth and random slope model was selected as the final unconditional model, which was then used to build further into a more complex conditional model (see next section).

### Conditional latent growth model

Once the most parsimonious unconditional latent growth model was established, it was important to explore the impact that key contextual covariates may have on the trajectory of children included in the study. This was done by adding effects to the model in a step-wise approach until the list of suitable covariates (that were included in the population model) were exhausted, significant effects were retained and explanatory power was improved with their addition. A “Cycle” interaction term was included with all covariates of interest to examine the change in the effect over time.

#### *Intervention effect*

The first conditional model explored was one which included an effect for the intervention (i.e., Intervention vs Control). As can be seen by Table 46, there is a statistically significant (0.05 level) difference between the control (reference) group and intervention (contrast) group from 2018, with the average ability of intervention children lower than the control group.

*Table 46. Regression estimates for conditional latent growth model – Intervention effect*

Term	Estimate	SE	t	p
Intercept	1.863	0.069	26.862	0.000
Cycle	0.710	0.092	7.735	0.000
Cycle <sup>2</sup>	0.143	0.055	2.575	0.010
Cycle <sup>3</sup>	0.060	0.009	6.498	0.000
Intervention	-0.217	0.097	-2.246	0.025
Cycle * Intervention	-0.096	0.138	-0.699	0.485
Cycle <sup>2</sup> * Intervention	0.002	0.082	0.028	0.978
Cycle <sup>3</sup> * Intervention	0.004	0.014	0.326	0.744

### *Intervention plus contextual factors*

The second conditional model (see Table 47) built upon the first by adding additional key covariates of interest. The statistically significant (at the 0.05 level) fixed effects are:

- Gender – Males demonstrated lower average ability than females
- Age at assessment – Older children demonstrated higher average ability
- Aboriginal and Torres Strait Islander – Aboriginal and Torres Strait Islander children demonstrated lower average ability than non-Aboriginal and Torres Strait Islander children
- SDQ (Externalising behaviour) – Children who demonstrated higher (concerning) levels of externalising behaviour demonstrated lower average ability than those who do not
- Parent/caregiver education – Children of parent/caregiver with university level education demonstrated higher average ability than children of parent/caregiver who has Year 12 education at most
- Parent/caregiver employment - Children of an unemployed parent/caregiver demonstrated lower average ability than children of an employed parent/caregiver

Of note is that in this model, once the strong disadvantage factors are added into the model, there is no significant difference between children in the intervention and control groups. That is, holding constant age, gender, Aboriginal and Torres Strait Islander status, child behaviour and temperament as well as parental education and employment, children in the EL&L group have caught up to their peers at the end of 2022. This is in stark comparison to the unconditional findings that show the EL&L group behind their peers. This is due to strong effects of disadvantage on learning: children from the most vulnerable groups miss out on important home and program inputs that lead to learning gaps.

Table 47. Regression estimates for conditional latent growth model – Intervention effect + Additional child and family effects

Term	Estimate	SE	t	p
Intercept	2.008	0.136	14.797	0.000
Cycle	0.922	0.200	4.621	0.000
Cycle <sup>2</sup>	0.238	0.113	2.110	0.036
Cycle <sup>3</sup>	0.071	0.018	3.900	0.000
Intervention	-0.110	0.094	-1.171	0.243
Male	-0.193	0.095	-2.038	0.043
Age at assessment	0.116	0.050	2.333	0.022
ATSI	-0.351	0.120	-2.935	0.004
SDQ - Externalising behaviour	-0.034	0.014	-2.409	0.018
Parent/Caregiver: VET Education	0.056	0.122	0.456	0.649
Parent/Caregiver: University Education	0.409	0.136	3.009	0.003
Parent/Caregiver: Unemployed	-0.296	0.118	-2.518	0.014
Cycle * Intervention	-0.052	0.136	-0.383	0.702
Cycle <sup>2</sup> * Intervention	0.017	0.084	0.198	0.843
Cycle <sup>3</sup> * Intervention	0.007	0.014	0.477	0.634
Cycle * Male	-0.146	0.131	-1.112	0.267
Cycle <sup>2</sup> * Male	-0.084	0.077	-1.093	0.275
Cycle <sup>3</sup> * Male	-0.011	0.013	-0.878	0.380
Cycle * Age at assessment	-0.113	0.069	-1.646	0.102
Cycle <sup>2</sup> * Age at assessment	-0.022	0.038	-0.576	0.565
Cycle <sup>3</sup> * Age at assessment	-0.007	0.006	-1.191	0.234
Cycle * ATSI	-0.146	0.170	-0.863	0.390
Cycle <sup>2</sup> * ATSI	-0.103	0.100	-1.032	0.303
Cycle <sup>3</sup> * ATSI	-0.015	0.016	-0.943	0.347
Cycle * SDQ - Externalising behaviour	0.042	0.021	1.987	0.050
Cycle <sup>2</sup> * SDQ - Externalising behaviour	0.025	0.012	1.984	0.050
Cycle <sup>3</sup> * SDQ - Externalising behaviour	0.004	0.002	1.733	0.086
Cycle * Parent/Caregiver: VET	-0.212	0.181	-1.172	0.244
Cycle * Parent/Caregiver: University	0.128	0.182	0.703	0.483
Cycle <sup>2</sup> * Parent/Caregiver: VET	-0.098	0.111	-0.883	0.379
Cycle <sup>2</sup> * Parent/Caregiver: University	0.074	0.111	0.670	0.504
Cycle <sup>3</sup> * Parent/Caregiver: VET	-0.014	0.019	-0.732	0.466
Cycle <sup>3</sup> * Parent/Caregiver: University	0.008	0.018	0.446	0.657
Cycle * Parent/Caregiver: Unemployed	-0.186	0.180	-1.034	0.306
Cycle <sup>2</sup> * Parent/Caregiver: Unemployed	-0.046	0.095	-0.483	0.630
Cycle <sup>3</sup> * Parent/Caregiver: Unemployed	0.000	0.015	-0.009	0.993

## Intervention intensity

An alternate way of exploring the magnitude of the intervention is to consider the actual degree to which children were intervened upon. That is, in the traditional intervention and control model, we assume that all children received the intention. This is, of course, not true. It is rather the *intention* to treat the children in the intervention group, but this is a matter of the degree to which the intervention is faithfully implemented in all sites. Instead, it is possible to judge the fidelity of the implementation and consider the actual treatment effect: the effect on the group of children who get high quality provision of the intervention for an extended period of time.

ALNF were tasked with evaluating the level of intensity of the EL&L activities in the intervention sites for each of the years in which data was collected, resulting in the population model variables ELLIntensity\_2018, ELLIntensity\_2019, ELLIntensity\_2021 and ELLIntensity\_2022. Each of these variables include the levels:

- 0 – Control
- 1 – None
- 2 – Low
- 3 – Medium
- 4 – High

The criteria for assigning a site to one of these levels is based on teachers, support staff and dosage (see Table 48).

*Table 48. Criteria for rating intervention intensity*

Intensity	Teacher	Support staff	Dose
High	≥1 Trained	Trained (sometimes)	Daily
Medium	New or partially trained	≥1 Trained	≤ Weekly
Low	Not trained	New or partially trained	Irregularly
None	None	None	None

Further to this, the concept of continuity of intervention intensity was explored. Several variables were created which reflected the extent to which children were exposed to different levels of intervention intensity from the beginning of the study and whether that intensity was sustained over time. The alternative variables were coded as follows:

- ELLContInt1
  - 4 – Low, medium or high intensity in 2018, 2019, 2021 and 2022
  - 3 – Low, medium or high intensity in 2018, 2019 and 2021 (but not 2022)
  - 2 – Low, medium or high intensity in 2018 and 2019 (but not 2021)
  - 1 – Low, medium or high intensity in 2018 (but not 2019)
  - 0 – No intensity in 2018



- ELLContInt2
  - 4 – Medium or high intensity in 2018, 2019, 2021 and 2022
  - 3 – Medium or high intensity in 2018, 2019 and 2021 (but not 2022)
  - 2 – Medium or high intensity in 2018 and 2019 (but not 2021)
  - 1 – Medium or high intensity in 2018 (but not 2019)
  - 0 – No or low intensity in 2018
- ELLContInt3
  - 4 – High intensity in 2018, 2019, 2021 and 2022
  - 3 – High intensity in 2018, 2019 and 2021 (but not 2022)
  - 2 – High intensity in 2018 and 2019 (but not 2021)
  - 1 – High intensity in 2018 (but not 2019)
  - 0 – No, low or medium intensity in 2018

Descriptive statistics of ELLDI scores for each of these variables was explored (see Table 49, Table 50 and Table 51). It is interesting to note that:

- 47 (out of 558) children were exposed to some level of intervention intensity for all four cycles (see Table 49)
- 15 (out of 558) children were exposed to continuous medium or high levels of intervention intensity beyond the first two years of the study (see Table 50)
- No children were exposed to continuous high levels of intervention intensity beyond the first two years of the study (see Table 51)

*Table 49. Descriptive statistics for ELLDI by continuity of intervention intensity – ELLContInt1: Low, Medium and High*

Group	N	2018		2019		2021		2022	
		Mean	SE	Mean	SES	Mean	SE	Mean	SE
No intensity	326	-2.48	0.08	-0.57	0.06	1.22	0.05	1.88	0.06
2018 EL&L Only	109	-2.49	0.13	-0.41	0.11	1.12	0.10	1.69	0.10
2018-2019 EL&L	48	-3.04	0.21	-1.07	0.17	0.95	0.14	1.44	0.14
2018-2021 EL&L	28	-1.87	0.23	-0.08	0.21	1.59	0.19	1.79	0.20
2018-2022 EL&L	47	-3.35	0.24	-1.20	0.20	0.58	0.17	1.08	0.15

Table 50. Descriptive statistics for ELLDI by continuity of intervention intensity variables – ELLContInt2: Medium and High

Group	N	2018		2019		2021		2022	
		Mean	SE	Mean	SES	Mean	SE	Mean	SE
No or low intensity	398	-2.67	0.07	-0.69	0.06	1.10	0.05	1.75	0.05
2018 EL&L Only	111	-2.23	0.14	-0.25	0.11	1.33	0.10	1.77	0.10
2018-2019 EL&L	34	-2.82	0.23	-0.97	0.18	1.00	0.16	1.49	0.17
2018-2021 EL&L	7	-2.06	0.46	-0.26	0.27	1.38	0.23	1.69	0.19
2018-2022 EL&L	8	-2.09	0.35	-0.49	0.25	0.86	0.06	1.49	0.12

Table 51. Descriptive statistics for ELLDI by continuity of intervention intensity variables – ELLContInt3: High

Group	N	2018		2019		2021		2022	
		Mean	SE	Mean	SES	Mean	SE	Mean	SE
No, low or medium intensity	495	-2.58	0.06	-0.63	0.05	1.14	0.04	1.75	0.05
2018 EL&L Only	49	-2.47	0.26	-0.46	0.20	1.26	0.18	1.66	0.17
2018-2019 EL&L	14	-2.54	0.28	-0.66	0.19	0.98	0.19	1.32	0.28
2018-2021 EL&L	0	-	-	-	-	-	-	-	-
2018-2022 EL&L	0	-	-	-	-	-	-	-	-

## Contextualising findings

To provide meaning to the ELLDI measure it is possible to use the statistical properties of the measure to describe the knowledge, skills and abilities that are observed along the scale and to describe how these knowledge, skills and abilities are becoming more sophisticated as children grow. The underlying item response model has the property of constant discrimination – this means that the relative challenge of the items on the measure remains in the same order no matter where children are located on the measure.<sup>14</sup> This allows for the description of a sequence of learning that is, on average, similar for all children. Such described scales are therefore useful, because educators (and researchers and policy makers) can look at the items that lie above a child's current location to get a sense for what skills come next, and how the application of children's current knowledge will change as they grow. This allows for targeted practice – teaching at the right level – to support children to consolidate what they are currently demonstrating and to be scaffolded to learn what comes next (to be given an appropriate amount of challenge and to be set realistic expectations). Most importantly, such a described scale gives educators a common language to understand what oral language and literacy looks like at a range of different levels. Such a common language allows for more productive transitions – as children move through the education system, educators can engage in professional discussions where they can explicitly describe what this child can do and what they will be tackling next.

### An ELLDI Scale

Table 52 presents the ODEC children's language and literacy development as a described scale, based on the results of the ELLDI from the four cycles of assessment. The levels of the measure are defined by the item locations (difficulty) transferred to Thurstonian Thresholds – in this case the response probability is set to 65% - so called, RP65s. RP65s are useful to interpret measures like the ELLDI as they represent cumulative probabilities. That is, higher adjacent response categories on an item are also located higher on the measure: the RP65 of an item that is scored zero, one, two will have two thresholds: between zero and one, and between one and two and threshold two will always be above threshold one. In this example, the first threshold would be the location on the scale where the probability of being in the first response category is 35% and the probability of being in the second or third category is 65%. The second threshold is the location on the scale where the probability of being in the first response or second response categories is 35% and the probability of being in the third category is 65%.

The RP65s therefore, represent consolidation of a skill: children located at the same level as an RP65 can most likely demonstrate the underlying skill. For items scores zero, one, two (and so on) we take the highest scored category only. The location therefore represents the location that a child at the same location is 65% likely to be able to demonstrate the skill. In this sense, the descriptions of the levels below, reflect something akin to mastery: children who are towards the top of these levels can most likely demonstrate the skills (consolidation) whilst the next level above reflects where they should be challenged to progress with scaffolded and integrated activities.

---

<sup>14</sup> And analysis has been undertaken to remove items that violate this assumption. See section on DIF.

For ease of interpretation the original scale (in logits) is transformed to a more user-friendly ELLDI Scale that has mean 500 and standard deviation 50. This ensures those interpreting the measure do not need to deal with negative numbers and can work with integers:

$$ELLDI = (Ability \times 50) + 500$$

The original item parameter estimates are presented along with the ELLDI transformed values in Appendix 15: Item parameter estimates. It can be seen in the Table 52 that the scale is broken into 9 levels ranging in difficulty from the easiest item found below 293 (providing limited additional information to describe an image) to the most difficult item found above 689 (coherently describes an image using a wide vocabulary and complex sentences to link ideas and explain relationships). The described scale covers all strands of the Big Six except fluency (Konza, 2014). It also includes a print convention strand and a writing strand to capture pre-literacy knowledge and skills, as well as children's skills and understanding of shape and letter formation, spelling, and punctuation.

Table 52. ELLDI Scale.

Level	Location	Major literacy strands	Sub-strand	Skills
9	>689	Oral language	Expressive	Describes an image using complex sentences and a wide vocabulary. Provides coherence by linking ideas and describing relationships based on the image
8	619-689	Comprehension	Listening comprehension	Listens to a story and makes an inference drawing on subtle evidence from the text to justify their reasoning
			Reading comprehension	Reads a short information text and identifies the purpose of a common convention in a diagram (used to illustrate size) Reads a short story and makes a generalisation about several key events Reads a short information text and uses contextual clues to infer the meaning of a word used in an unfamiliar way
		Oral language	Expressive	Describes favourite picture using simple and compound sentences. Provides coherence using simple vocabulary and describing elements of the picture
		Phonics	Reading	Reads a sentence (12 words), including CVC and high frequency words accurately
		Vocabulary	Expressive	Generates names for up to 10 different things that can take us places
		Writing	Punctuation	Accurately punctuates a dictated sentence
		Spelling	Writes a dictated sentence (8 words), spelling all words correct	
7	564-619	Comprehension	Listening comprehension	Listens to a story and provides evidence to support a judgement about a character's behaviour when this is supported by clues in the text and makes familiar inferences Listens to a story and draws on personal knowledge and clues from the context to infer the meaning of uncommon vocabulary
			Reading comprehension	Reads a short, simple information text (88 words) and makes a simple inference combining prominent clues in an image and the text Reads a short, simple text (68 words) and infers a character's feeling when this is familiar, and clues are obvious
		Phonemic awareness	Medial phoneme match	Consistently identifies whether multiple pairs of simple words have the same or different middle sounds

Level	Location	Major literacy strands	Sub-strand	Skills	
6	515-564	Phonological awareness	Final blend: Phoneme deletion	Deletes the first sound in a blend at the end of a spoken word and says the new word	
			Final phoneme match	Consistently identified whether multiple pairs of simple words have the same or different final sounds	
			Final phoneme identification	Provides the last sound in a given simple word	
			Rhyme	Produces two rhymes for a given spoken word	
			Syllables	Counts the number of syllables in a one syllable word	
			Writing	Formation	Forms letters using correct shape, size, and orientation
				Spelling	Writes a dictated simple sentence (3 words), spelling all words correct
				Comprehension	Reading comprehension
			Oral language	Clarity, volume, fluency	Speaks in a way that is clear and easy to understand Uses volume and pace to emphasise meaning
		Phonemic awareness	Phoneme segmentation	Segments a short, simple 2-3 letter word articulating each sound separately	
			Final phoneme deletion	Deletes the final sound of a given spoken word and says the new word	
			Initial phoneme match	Consistently identified whether multiple pairs of simple words have the same or different first sounds	
			Initial phoneme: word production	Produces a spoken word with the same first sound as the one provided	
Initial phoneme identification	Provides the first sound in a given simple word				
Phonics	Decoding	Matches a simple word with a suffix (a verb ending in 'ing') from a group of 4 to a picture, with multiple distractors			
	Letter name	Names 20 alphabet letters (mix of upper and lower case)			
Phonological awareness	Rhyme	Identifies words that rhyme from a set of given words or from a text			
	Syllables	Counts the number of syllables in an open 3 syllable word (ba-na-na)			

Level	Location	Major literacy strands	Sub-strand	Skills
		Print conventions	Text awareness	Knows where to begin reading a story and where to go next when they get to the end of a line of text (return sweep)
		Vocabulary	Expressive	Names less common body parts (hip, spine) Names some less common objects (market, stall)
			Receptive	Follows temporal directions when the word 'after' is used in the middle of the direction
5	470-515	Comprehension	Listening comprehension	Listens to a story and draws an inference from simple clues in the text and illustrations (e.g., ideas, feelings) Listens to a story and identifies and connects several details stated in the text, about the same event
		Oral language	Clarity, volume, fluency	Speaks with a level of clarity but a few words maybe difficult to understand Uses some variation in volume and pace to emphasise meaning
		Phonemic awareness	Medial phoneme match	Identifies if a single pair of simple words have the same or different medial sounds when this is obvious
		Phonics	Decoding	Matches a simple word with a suffix (a verb ending in 'ing') from a group of 4 to a picture, with limited initial letter distractors
			Letter name	Names 13 alphabet letters (mix of upper and lower case)
			Letter sound	Provides phonemes for 7 letters
		Phonological awareness	Rhyme	Produces a single rhyme for a given spoken word
			Syllables	Counts the number of syllables in a closed 3 syllable word (after-noon)
		Print conventions	Text awareness	Shows that reading goes from left to right
		Vocabulary	Receptive	Follows temporal directions when the word 'before' is used at the beginning of the direction
4	419-470	Comprehension	Listening comprehension	Listens to a story and makes simple inferences based on prominent clues in the illustrations Provides evidence from the story about a character's actions (what they did, what happened) Listens to a story and draws on personal knowledge and clues from the text to infer the meaning of simple vocabulary
		Oral language	Clarity, volume, fluency	Makes self generally understood when speaking but there may be some hesitation

Level	Location	Major literacy strands	Sub-strand	Skills
				Uses limited variation in volume and pace to emphasise meaning
			Expressive	Describes an image, using some simple, relevant words but may go off topic
		Phonemic awareness	Final phoneme match	Identifies if a single pair of simple words have the same or different last sounds when this is obvious
			Initial phoneme match	Identifies if a single pair of simple words have the same or different first sounds when one word includes a blend
		Phonics	Decoding	Matches a simple CVC word from a group of 4 to a picture, with limited initial letter distractors
			Letter name	Names 3 out of 3 upper case alphabet letters
			Letter sound	Provides phonemes for 4 out of 4 different letters (mix upper and lower case)
		Phonological awareness	Compound words	Identifies the first word in a spoken compound word
			Syllables	Counts the number of syllables in two syllable words, whether they are open or closed syllables, but not one-syllable words
		Print conventions	Text awareness	Finds pairs of matching letters when they are the same shape and font
		Vocabulary	Expressive	Generates names for up to 10 different animals or types of foods
			Receptive	Follows a 3-step sequence of verbal instructions (do this then this then this)
		Writing	Formation	Copies simple shapes and writes their name
3	363-419	Comprehension	Listening comprehension	Listens to a story and provides simple evidence using the text and the image
			Oral language	Clarity, volume, fluency
		Expressive		Speaks mainly in phrases of 2-3 words
		Phonemic awareness	Initial phoneme match	Identifies if a single pair of simple words have the same or different first sound when this is obvious
		Phonological awareness	Compound word	Identifies the final word in a spoken compound word
		Print conventions	Text awareness	Indicates where a word is written on the cover of a book
			Book orientation	Turns pages right to left in order, looking through a book
		Vocabulary	Expressive	Names a variety of common things in an image



Level	Location	Major literacy strands	Sub-strand	Skills
			Receptive	Follows a 2-step sequence of verbal instructions (do this then this)
2	293-363	Comprehension	Listening comprehension	Listens to a story and provides information using the image
		Oral language	Expressive	Describes an image mainly using single words often with pauses or fillers (um, ah)
1	<293	Oral language	Expressive	Provides one or two additional words to describe an image when prompted
		Vocabulary	Expressive	Names three common body parts

The proportion of children who are working at each of the nine ELLDI reporting levels at each stage of the study can be seen in Table 53. This table shows the spread of children at each timepoint across the different ELDDI reporting levels, and hence spread across the scale. It also allows for an approximate examination of the change (growth) in students from year to year in reference to the reporting levels.

*Table 53. Proportions of children in each level on the ELLDI, by year*

Level	Year			
	2018	2019	2021	2022
1	6.7%	0.3%		
2	31.7%	4.8%		
3	42.0%	13.4%	0.7%	0.1%
4	16.0%	31.3%	3.8%	1.7%
5	3.3%	30.9%	13.3%	7.8%
6	0.4%	15.1%	31.3%	21.4%
7		4.1%	41.0%	40.1%
8		0.2%	9.7%	26.3%
9			0.2%	2.5%

The distribution of the 2018 sample relative to the ELLDI Scale can be seen in Figure 14. This includes all the RP65 thresholds used to produce the described scale across all four rounds of assessment administration. Additional images of the distribution of the 2019, 2021 and 2022 samples relative to the ELLDI Scale can be seen in Appendix 16: Item-Person maps of the sample distribution relative to ELLDI Scale levels.

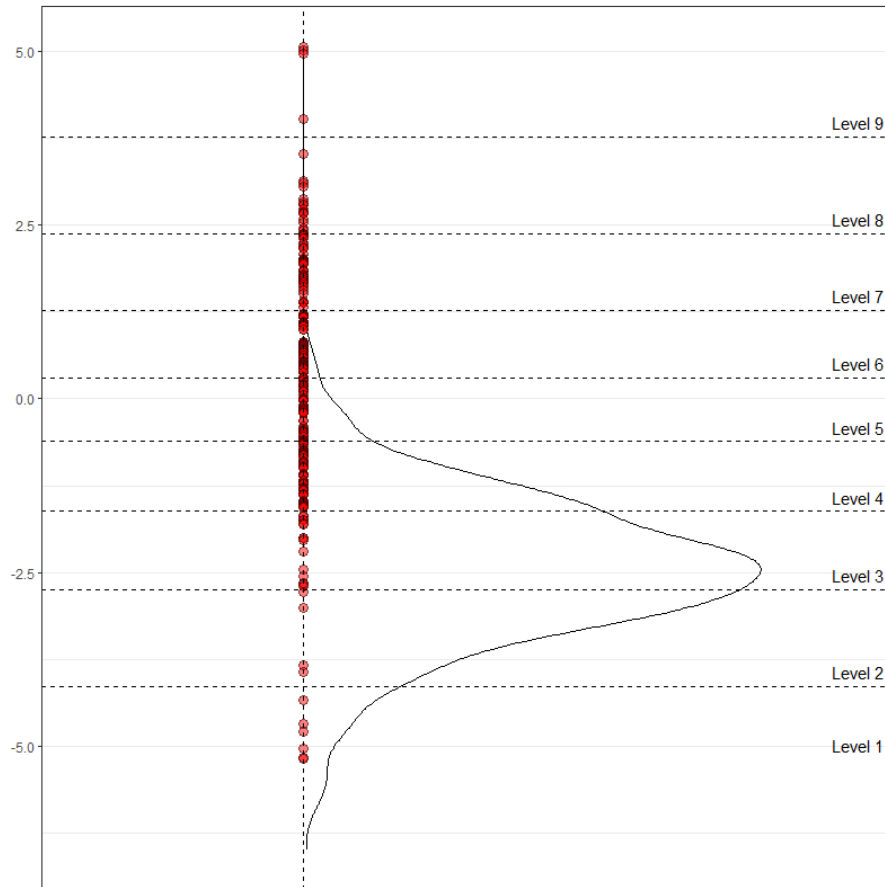


Figure 14. Item-person map for the ELLDI Scale, including RP65 thresholds from all cycles (2018 sample only)

### Interpretation of the ELLDI described language and literacy scale

When interpreting the ELLDI Scale it can be seen that skills such as those found in the oral language, vocabulary and comprehension strands are distributed throughout the scale but they are more prominent at the higher and/or lower points. Unlike the constrained skills found in the other strands, such as phonics and phonemic awareness, which for most children can be mastered over time with practice, skills in the oral language, vocabulary and comprehension strands are considered unconstrained emphasising that they continue to develop over a lifetime (Paris, 2005). Added to the notion that there is no ceiling in the development of these unconstrained skills is the interrelationship that exists across these three strands. When considered in terms of children's language and literacy learning, rich oral language environments provide the foundation and vehicle for larger and more diverse vocabulary development, which in turn equip children with a strong understanding of language structures that support effective comprehension (Carlson et al., 2013; Nation and Snowling, 2004; Tunmer & Chapman, 2012).

Understanding these three strands and the unconstrained skills that sit within them is key for two reasons.

Firstly, it is plausible that the learning and development of these skills will be different to those of the constrained skills found in the other strands. This is evident in research focusing on reading comprehension where the emphasis is on teaching children a range of strategies to support comprehension. This includes but is not limited to strategies such as comprehension monitoring and graphic organisers, which children can choose to use to assist comprehension as they learn to self-regulate their reading (NICHD, 2000). Similarly, oral language and vocabulary development require the use of different teaching strategies. This includes but is not limited to the use of direct and indirect instruction across contexts, as well as instructional techniques that integrate, “sophisticated vocabulary, contextual support for learning new vocabulary, repeated interactive book reading, engaging children in extended conversations and use of complex syntax” (Grifenhagen et al., 2016, p. 2).

Secondly, it is conceivable that there will be children who present with either limited or exceptional oral language, vocabulary, and comprehension skills and therefore their ability may sit outside of the ELLDI Scale. In such instances, it may be necessary to conduct further assessments with these children or consult with allied health professionals to understand and cater for their specific learning needs. What is being emphasised here, is that if the teaching of oral language, vocabulary and comprehension is restricted to only those skills described on the ELLDI Scale for these children then it is unlikely to result in significant improvements in their oral language, vocabulary, and comprehension skills.

As with the unconstrained skills it is important to understand the development of constrained skills on the ELLDI Scale. On the scale, constrained skills can be found in print conventions, phonological awareness, phonemic awareness, phonics and writing strands. Unlike the strands containing unconstrained skills, these strands are situated between Level 3 to Level 8 on the scale. As previously stated, constrained skills are those that have a ceiling because once they are learnt they can be applied with a certain level of precision, and they do not need to be taught again. An example of this on the ELLDI Scale can be seen in the print convention strand. At Level 3, children’s mastery of print conventions is reflected in their ability to peruse each page in a book in order, as well as distinguishing words from other content on the cover of a book. The most difficult print convention skill is mastered at Level 6 when children know that once you finish reading a line of text you use a return sweep to continue reading the next line. Once a child has learned the return sweep, according to the scale they will have mastered all the skills in the print convention strand, so teaching can be focused on mastery of other strands and skills at that level. Here, it is important to remember that the ELLDI Scale is not exhaustive and is limited by the type of knowledge and skills assessed and represented in each of the unconstrained and constrained strands. This also means the scale has the potential to be extended, with items from future assessments being able to be calibrated onto the scale. This could result in additional skills being added to the different strands at various levels or an increase in levels. For example, additional items could mean the inclusion of extra print convention skills or oral language skills that extend what has been achieved at Level 9.

When reviewing the skills in the phonics strand it can be seen they have been classified under four sub-strands: letter names, letter sounds, decoding and reading. While the letter name and letter sound sub-strands pertain to children mastering the alphabetic principle, the decoding and reading sub-strands address the application and mastery of the alphabetic principle to attach sounds to strings of letters to articulate them as words. In particular, the focus of the decoding sub-strand is at the word level, whereas the focus of

the reading sub-strand is at the sentence level. The classification of items into decoding and reading sub-strands in the phonics strand differentiates them from the reading comprehension sub-strand in the comprehension strand. This is to aid in distinguishing between decoding simple words or reading a sentence, with no or limited focus on making meaning from what was read (sometimes referred to as 'barking at print'), compared to reading words to make meaning and respond to comprehension questions.

Of interest in the phonics strand is the mastering of skills involving the alphabetic principle and decoding. The ELLDI Scale indicates these skills share similar locations and they were mastered simultaneously (see the Phonics strand at Levels 4, 5 and 6 in Table 52). The simultaneous mastering of the alphabetic principle and decoding provides support for the use of a synthetic phonics approach to teaching phonics, whereby letter sounds are taught in groups in a systematic sequence (e.g., s, a, t, p, i, n...) to maximise opportunities for children to blend sounds to read and write in the form of vc (at), and cvc (pin) words. Research examining the use of synthetic phonics to teach decoding has shown that children taught using this approach outperform children who were taught using other phonics approaches (Johnston & Watson, 2005; Wheldall et al., 2019).

### Further applications of a described scale

#### *Sustainable Development Goals (SDGs)*

With descriptions of the scale and exemplar items available to illustrate these skills at many locations, it is possible to extend the use of the ELLDI to contextualise other, external criteria, benchmarks, measures and assessments. For example, to identify the location on the ELLDI at which children are likely to have met Sustainable Development Goal SDG 4.2.1 or 4.1.1a. This allows observations of the ELLDI to reference external criteria and serve a dual purpose: to be used as a formative assessment and to provide this additional information without the need for the child and educator to complete many assessments.

**SDG indicator 4.2.1** refers to the "Proportion of children aged 24–59 months who are developmentally on track in health, learning and psychosocial well-being, by sex" (UNICEF, 2022a). The Early Childhood Development Index 2030 (ECDI2030) tool is designed to measure SDG indicator 4.2.1 (UNICEF, 2021). Children aged 5 years are said to be developmentally on track for learning if they answer at least 15 of 20 items correctly on the ECDI2030. Some examples of the types of learning related items from the ECDI2030 are:

- ECD5. Can (name) say 10 or more words, like 'mama' or 'ball'?
- ECD6. Can (name) speak using sentences of 3 or more words that go together, for example, "I want water" or "The house is big"?
- ECD7. Can (name) speak using sentences of 5 or more words that go together, for example, "The house is very big"?
- ECD8. Can (name) correctly use any of the words 'I,' 'you,' 'she,' or 'he,' for example, "I want water" or "He eats rice"?
- ECD9. If you show (name) an object (he/she) knows well, such as a cup or animal, can (he/she) consistently name it?
- ECD10. Can (name) recognize at least 5 letters of the alphabet?
- ECD11. Can (name) write (his/her) name?

Note that not all ECDI2030 items relate to oral language and literacy – and therefore caution should be taken in interpreting this alignment, however, it is assumed that if a child can demonstrate all the skills listed above, they have met SDG 4.2.1, and this location can be reliably identified on the measure. A child who has consolidated the skills described in SDG 4.2.1 is estimated to be located at Level 4 on the ELLDI Scale. Using this as a benchmark, 81.6% of the ODEC children were in Level 4 or above (meeting minimum standard) and 50.3% of children were in Level 5 or above (exceeding minimum standard) in 2019 (approx. 5 years of age). It is to be expected that most children should meet this minimum standard – it is, after all, a global minimum indicator developed to be also used in developing contexts. The prospect of up to 20% of children in this sample not meeting this standard is an indicator of the overall low level of learning observed and likely related to the high level of disadvantage in the region.

**SDG indicator 4.1.1 (a)** refers to the “Proportion of children and young people in grades 2/3 achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex” (UNICEF, 2022b). Minimum proficiency levels (MPLs) have been developed for benchmarking against SDG 4.1.1 (a) at key stages of schooling (UNICEF, 2022b). Broad MPL definitions for Grades 2 and 3 are as follows (UNICEF, 2022b):

- Reading
  - Grade 2 - They read and comprehend most of written words, particularly familiar ones, and extract explicit information from sentences.
  - Grade 3 - Students read aloud written words accurately and fluently. They understand the overall meaning of sentences and short texts. Students identify the texts’ topic.

More detailed descriptions of the MPLs for sub-domains of reading are also available in the Global Proficiency Framework (GPF: United States Agency for International Development *et al.*, 2020) for assessing approximate locations of the MPLs on the ELLDI Scale. Descriptors are available for aural language comprehension, decoding and reading comprehension (retrieving information, interpreting information). Some examples are as follows (United States Agency for International Development *et al.*, 2020):

- a. Aural language comprehension
  - Grade 2 - Identify key events, ideas or major characters
  - Grade 3 - Make simple inferences that require connecting explicit clues from different parts of the text
- b. Decoding
  - Grade 2 - Decode most words in a connected text, including some unfamiliar ones
  - Grade 3 - Read at a pace and a level of accuracy that meets minimum country standards for fluency
- c. Reading comprehension
  - Grade 2 - Locate most pieces of explicit information in a sentence when the information is prominent and there is no competing information

- Grade 3 - Understand most simple, implicit information in a text by filling in obvious missing information in one sentence or in two consecutive sentences (e.g., Sally blew out the candles on her cake. What was she celebrating?)

If a child has consolidated the skills described above (along with the other skills described in the GPF), then they are deemed to be located at approximately Level 5 to Level 6. Given SDG 4.1.1 (a) is concerned with the proportion of children achieving the MPL at Grade 2/3, the benchmark of achieving Level 6 (on the ELLDI Scale) in 2022 (start of Grade 3) is used. In this case, 90.3% of the ODEC children were in Level 6 or above (meeting minimum standard) and 68.9% of children were in Level 7 or above (exceeding minimum standard) in 2022 (start of Grade 3). Similar to the finding (but slightly less stark) relating to SDG 4.2.1, most children should be expected to be meeting the MPL by Grade 3, given it is designed as a global indicator for use in developing contexts. However, approximately 10% of children in the sample did not meet this minimum standard by Grade 3 further highlights the level of disadvantage in the region and the need for more support.

### *Australian Curriculum*

The idea that the ELLDI Scale can be used to contextualise other benchmarks can be extended to important Australian transition points to help reduce the uncertainty about the skills and abilities of children as they enter preschool and exit preschool (transition to school). For example, the descriptions in the **Australian Curriculum for the Foundation year** (first year of school) can be identified – we assume these to be the skills needed to successfully engage in school. Such locations can therefore decouple the child's learning from the system – we can know ahead of time if a child is on track to be ready to engage with preschool or school and intervene before on-entry assessments are conducted. The results on the measure again serve a dual purpose – the formative part of the assessment allows educators to intervene at the right level which the results of the assessment allow communication against external criteria that often require additional assessment or refer to third parties (such as allied health services).

The English strand of the Australian Curriculum aims to help children develop effective language and literacy skills as they progress from Foundation to Year 10. The curriculum itself comprises of three interrelated strands of language, literature and literacy, covering the development of speaking, listening, viewing, reading, writing and creating. Students' attainment of these skills are described as a sequence of achievement from Foundation to Year 10 and are classified as receptive learning (listening, reading and viewing) and productive learning (speaking, writing and creating). Examples of achievement standards and the types of learning that is expected by students at the end of the Foundation year (first year of school) are:

- a. Receptive learning
  - Recall one or two events from texts with familiar topics.
  - Recognise the letters of the English alphabet, in upper and lower case and know and use the most common sounds represented by most letters.
  - Blend sounds orally to read consonant-vowel-consonant words

b. Productive learning

- Listen for rhyme, letter patterns and sounds in words.
- Retell events and experiences with peers and known adults.
- Identify and use rhyme, and orally blend and segment sounds in words

Based on the English achievement standards for the Foundation year (first year of school) of the Australian Curriculum (ACARA, 2021), the skills required to be working at this level are approximately aligned with Level 6 on the ELLDI Scale. In 2019 (start of Prep), 19.4% of the sample were working at Level 6 or above (meeting minimum standard) and 4.3% were already working at Level 7 or above (exceeding minimum standard). However, the achievement standard for Foundation refers to the skills that need to be required to meet the grade level expectations by the end of the Foundation year. Given that no data exists for the end of Prep (or start of Grade 1) it is more difficult to assess the achievement of the ODEC sample against Foundation. The average growth across the study was 53.1 ELLDI Scale points which equates to the sample moving up by approximately one ELLDI level for each year of the study. This would likely mean that approximately half of the sample would be in Level 6 or above (meeting minimum standard) and approximately 15-20% would be in Level 7 or above (exceeding minimum standard) by the start of Grade 1. In contrast, and concerningly so, that would mean that half of the sample would not meet the minimum achievement standard of the Foundation level of the Australian Curriculum by the time they enter Grade 1.



# Discussion and recommendations

The aim of the Overcoming Disadvantage in Early Childhood study is to understand the impact the EL&L program has on the language and literacy development of preschool-aged children. The following discussion synthesises the results from the study and provides recommendations for the implementation of the EL&L program to support the continued advancement of children's language and literacy learning.

## The Early Language & Literacy (EL&L) program

### EL&L impact

An expectation of preschool is for it to offer comprehensive programs that positively impact children's language and literacy learning and that these programs help close the gap between those from advantaged and disadvantaged backgrounds. The EL&L program did this. Results reveal that children in the intervention group, receiving the EL&L program in preschool, went from being significantly behind the control group in their language and literacy learning at timepoint one (2018), to catching up at timepoint two (2019). This is despite the intervention children being from more disadvantaged backgrounds than the control group. Added to this, there were also more children with diagnosed disability or language delay in the intervention centres. Understanding this is not only important in terms of the children's growth but also in terms of how well the EL&L program balances and targets the early language and literacy knowledge, skills and concepts children need in order to become literate. This finding supports that of Chambers et al. (2016) and the argument that preschool programs should include explicit and targeted teaching of early language and literacy skills, such as phonological awareness and oral language if they are to positively impact children's learning.

At subsequent timepoints, when the intervention children transitioned from preschool into school, the accelerated trajectories were not sustained and by the final timepoint (2022) a gap between the intervention and control children had re-emerged. Along with the program being implemented in a small number of schools, it was implemented with less rigor and consistency. This could be due to a number of reasons such as the level of commitment to the program, an overcrowded curriculum and EL&L trained staff moving schools.

The results described in the paragraph above come from models controlling only for child gender and age. In subsequent models where the full set of covariates were controlled for – in particular disadvantage - it was observed that the gap between the EL&L intervention group and the control group was explained-away. That is, differences between the two groups can be explained by the relative disadvantage of the EL&L group. This reinforces the finding that the EL&L program targets the most disadvantaged children in the region. These children suffer multiple disadvantages: they receive the least preschool and other ECEC inputs and come from the most vulnerable families. In the context of the intervention, it is clear that the EL&L program has the potential to close achievement gaps, and can do so when the intensity of the program is higher. To successfully close, and keep closed, achievement gaps, the program must be sustained over multiple years.

While it is likely that the move from preschool to school and the change from a flexible program to a specified curriculum contributed to the reduced language and literacy growth of the intervention children, consideration needs to be given to the impact of the number of natural disasters that took place across the Mid North Coast from the end of 2019 onwards. In the first instance, in 2019 the weather in NSW was the warmest and driest on record and the Mid North Coast experienced extreme heat and fires, which periodically closed preschools and schools. This was followed by a worldwide pandemic in 2020, which included lockdowns that interrupted school attendance, through closures, and extended absences with families hesitant to send their children to school and children being infected and absent with COVID-19. The effect of the pandemic delayed the third timepoint of data collection and impacted the study until the fourth and final timepoint. Recent COVID-19 related research conducted by Cloney and Jeffries (unpublished) indicates that the detrimental effects felt by the COVID-19 pandemic were more substantial for low SES children and as a result the learning gaps widened. Finally, the region experienced unprecedented rainfall in 2021/2022, which again led to school closures and further interruptions to school attendance due to flooding. Hence, added stresses and reduced time at preschool and school meant less exposure to the EL&L program and reduced outcomes for children, particularly those from low SES households.

### Program commitment

We know from professional development research that PD programs involving the 'enactment' - professional experimentation and the opportunity to put new ideas into action, along with support from leadership, are more effective than those that do not involve these elements (Clarke & Hollingsworth, 2002; Kennedy, 2016; Timperley et al., 2007). After careful analysis it was shown there was strong support for the implementation of the EL&L program in the majority of intervention centres in the early part of the study. This was evident in the many directors and educators that had completed their EL&L training up to 8 years prior to the study, along with some directors and educators currently undertaking EL&L training. There was also continued use of the program for longer than 24 months at intervention centres. Added to this, there were other elements of the EL&L training and program identified in research as contributing to the effective and lasting use of what is learned during PD. This includes situated practice - connecting what was taught during the PD to ECEC programming and classroom instruction (Borko et al., 2010), and mentoring (Kraft et al., 2018). These identified, long-term and sustained practices are what Timperley et al. (2007), refer to as 'maintaining momentum,' which have the potential to translate into improved outcomes in children's learning.

As the children transitioned out of the 22 initially recruited ECEC services into well over 100 additional schools, the same pattern was not observed. Indeed, many schools are outside the programming of ALNF and therefore do not provide the program. Many schools that had teachers that had completed EL&L training either did not implement the program due to staff turnover or other reasons (e.g., change of leadership).

### Overcoming disadvantage

Analysis of the intervention children's language and literacy learning across the duration of the study, highlighted and reinforced the impact of multiple disadvantages. This has been identified consistently in the literature and shown to impact children's language and

literacy growth. As might be expected, parent and caregivers' level of education and employment impacted children's language and literacy ability. Intervention children received less educational stimulation, such as being read to at home. Boys demonstrated lower ability than girls in preschool. Aboriginal and Torres Strait Islander children demonstrated lower average ability than other children in the intervention group (and were proportionally more likely to be in the intervention group). This difference in ability was statistically significant in the study from timepoint two onward, when the children left the EL&L program and graduated from preschool to school. In addition, the intervention sites operated in low socio-economic status (SES) neighbourhoods and generally rated as lower quality on the National Quality Standard (ACECQA, 2020), and were more likely to be *working towards* the Standard, and less likely to be *exceeding* the Standard. At the end of the study in 2022, the lowest ability children from the study had not reached the language and literacy level of the highest achieving children from the first timepoint in 2018. Here there is considerable evidence to show that if these low ability children do not have the reading basics by the end of grade 3, then it is likely they will continue to struggle to learn to read proficiently to support their learning, when the focus shifts from learning to read, to reading to learn (Juel, 2006; Lonigan et al., 2009). Each of these findings on their own, has the potential to impact a child's language and literacy learning but in many cases the intervention children are defined by multiple disadvantages, putting them further behind their more advantaged peers and making them less likely to be ready for school. These findings provide a clear picture of the level of disadvantage in the study area and shed light on the extensive intervention these children will need in preschool and beyond to help them reach their potential.

It is well documented that disadvantaged children or those with learning difficulties require intense use of programs promoting long term explicit and systematic instruction to consolidate and support learning. To understand the fidelity of the EL&L program the intervention sites were classified into groups according to the recency of engagement with ALNF mentors and trainers and qualitative judgements of the degree to which the program is being implemented as designed. With intervention children going from significantly behind in their language and literacy learning in timepoint one to catching up to the Control group in timepoint two, there is no surprise that 90% of these children experienced reasonable levels of program intensity during the first year of the study. As the children transitioned from preschool to school, however, the number of children exposed to the EL&L program decreased. During the final two years of the study less than 20% of the intervention children were exposed to a combination of medium and high or low, medium, and high intervention intensity, while attending school. Therefore, it is not surprising that no intervention children were exposed to continuous high levels of EL&L program across the duration of the study.

Understanding the intensity of the EL&L program and the impact it had on children's language and literacy learning provides two valuable pieces of information. Firstly, the importance of continuous, explicit and systematic exposure to the program cannot be underestimated given the positive impact it had on the children's language and literacy development in preschool. Secondly, even when the program was delivered with varied intensity, it achieved the desired effect of not only improving children's language and literacy learning, but also accelerating it, to closing the learning gap between the intervention and control group. Consequently, it could be hypothesised that if the intervention children were exposed to the EL&L program with continuous high intensity then the learning gains would be greater than what was observed between timepoint one and timepoint two. There is also the possibility that this gap could be reversed with the

intervention children surpassing the control children, which is consistent with research where socioeconomic disadvantage is disentangled from access to high quality, and intensity ECEC programs (Duncan & Sojourner 2013).

### Integration of EL&L

To understand the quality of classroom interactions and the possible reach of the EL&L strategies and pedagogies, educators in the intervention and control groups were observed using the CLASS. CLASS is an observational tool used to understand differences in the quality of interactions between educators and the children in their class, along with interactions between children. Results from these observations indicated there were no significant differences in the quality of the educator-child interactions across the CLASS domains, even though there were greater variations in the quality of the classroom interactions across the intervention educators. That said, on average, intervention educators displayed higher levels of emotional support than control educators. The emotional support domain of the CLASS is founded on the building of strong and respectful relationship, and the educator's ability to be responsive to the children's needs, as well as support responsibility and autonomy. Studies focused on the benefits of high levels of emotional support and strong educator-child relationships have been shown to positively influence early learning and counter children's early disadvantage and learning difficulties (Moen et al. 2019; Hamre & Pianta, 2001). Furthermore, with the intervention educators' EL&L training it could be said they have specialized language and literacy knowledge to support their interactions with the children. Here, Picker (2022) established that educators with higher levels of content knowledge had stronger relationships with the children in their class and this was evident in instances such as, the acknowledgement of the children's emotional needs during pedagogical interactions and gratitude for children's learning efforts. It is possible, that this may also be the case for the intervention educators who used the knowledge gained from the EL&L training to be sensitive to the children's emotional needs, while being responsive to their perspectives, which supported improved language and literacy learning. Nevertheless, as defined by the CLASS, the instructional support of the intervention educators did not reflect that of the educators in Picker's (2022) research where higher levels of content knowledge resulted in higher levels of instructional support.

The inclusion of mentoring in the EL&L program is designed to provide educators with the opportunity to practice applying the knowledge, skills and strategies from the training, while receiving explicit and targeted feedback to improve the implementation of the program and reflect on its use. These mentoring practices align with the PD research indicating that PD followed by mentoring increases the effectiveness of educators enacting what they have learnt during the training (Kraft et al., 2018). However, results from the CLASS Instructional Support analysis, suggest that the strategies and pedagogies taught as a part of the EL&L training, had limited impact on other instructional interactions during day-to-day preschool experiences. This is a common CLASS finding, with numerous studies indicating that instructional support is of a low-quality in ECEC settings (Hamre & Pianta, 2005; Soliday et al., 2021), despite the researchers emphasizing that quality instructional support leads to improved learning, particularly for children from disadvantaged backgrounds (Chambers et al., 2016). Therefore, if we consider that the mentor's role is to ensure quality implementation of the EL&L program and the program's overarching aim is to improve children's language and literacy skills to increase their opportunities in life, then it would be expected that elements of the EL&L

program, such as the use of various pedagogies ('Integrated Practice', 'Serve and Return' and 'Learning through Conversation' strategies in the EL&L materials) would be present in general class experiences. As there is evidence of variable implementation of these pedagogies, it is recommended that the EL&L training and mentoring be extended to include awareness of domain general application of these pedagogies and strategies to help educators improve their instructional support, with the aim of optimizing and integrating language and literacy learning into the less formal and child-centred parts of the day such as during transitions times, and meal time, in addition to specific EL&L experiences.

Finally, to truly understand the impact of the EL&L program on children's language and literacy development, Chambers and colleagues (2016) stress that comprehensive programs are enhanced by the inclusion of measures that provide educators with information about children's achievement. Added to this assessment results are used to plan for targeted educator-directed and child-centred experiences that continue to progress children's language and literacy learning (Chambers et al., 2016, p. 100). It is noted, that ALNF have been using the ELLDI to formatively assess children to support improved professional discussion, explicit planning and learning outcomes (Groom et al., 2022) based on the following interim recommendations:

1. Implement the use of the ELLDI assessment with the EL&L program as a formative assessment.
2. Construct a described scale to simplify the interpretation of the results of the ELLDI.
3. Construct supporting materials to support the interrogation of ELLDI results for individual children – particularly to quantify growth, and to highlight strengths and limitations.

## Final EL&L recommendations

1. **Increase children's exposure to EL&L** in the preschool year (and earlier) through intensifying the program expectations (hours per day and days per weeks). Such intensity is required time for consolidation and mastery of language and literacy skills when it is likely that most children will not get long-term exposure and support in the primary school years under the current scope of the program.
2. **Build further quality control into the EL&L program** to ensure fidelity of implementation and maximise engagement with the program. For example, ALNF mentors could use measures of pedagogical quality as a lens to focus and improve practice (Cloney & Hollingsworth, 2018). If one or more ALNF staff were trained to use the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2009), they could continue to support the implementation of the EL&L program while also supporting educators to improve the quality of their interactions with children within and beyond EL&L experiences. In this context, *interactions* refer specifically to the pedagogical strategies used by educators to model language, to give feedback and to introduce and connect concepts.
  - In addition to this, ALNF may elect to use more general measures of ECEC program quality, including the MELQO MELE to monitor or track quality in its settings. This could be done referring to external benchmarks to ensure the quality in EL&L sites is considered high. An observation schedule was developed as part of the ODEC study, and this could be deployed or adapted as required.
3. **Expand the scope of EL&L:** Seek opportunities to expand the scope of the program to cover more school sites or to implement intervention models that follow the child, no matter their location. This may include partnerships within local communities to reduce barriers to participation. For example, overcoming transport costs by providing free transport, or partnering with home-based services to deliver programs in the homes of the most vulnerable children.

## The Early Language and Literacy Developmental Index (ELLDI)

The need for an assessment to capture children's growth in the early years, which is built on frameworks of best practice in early learning (quality interactions) and principles that define quality assessments (cutting edge psychometrics), has been acknowledged in this report. The Early Language and Literacy Developmental Index (ELLDI) was spawned from this need and six separate assessment booklets (two booklets from timepoint 2 onwards), including the development of supporting resources and oral language coding rubrics, to accurately measure children's early language and literacy learning across the duration of the study were constructed (see Appendix 1: ELLDI assessment booklets).

Analysis of the ELLDI results, indicate it meets Cloney and colleagues' (2019) principles that define quality early years assessments. Firstly, the ELLDI was found to validly measure language and literacy constructs, as well as provide valid and specific information about children's learning. Secondly, it was considered to be reliable, as it produced consistent and valid results across the many education centres in the study. Finally, the ELLDI accurately measures and allows for the mapping of language and literacy components such as those identified in the science of reading, such as phonemic awareness, onto a scale of difficulty. An example of this can be seen with phonemic awareness - the ability to break spoken words into their separate sounds. The skill of phoneme isolation (identifying the first sound in a word) was typically located at the lower, less difficult end of the scale. Whereas, phoneme manipulation was located higher on the scale, as its more difficult to manipulate sounds in words, than it is to isolate them.

The ELLDI was deemed to be well-targeted in its ability to assess the language and literacy skills across the range of children's abilities in the study sample, particularly in the first two timepoints. It was noted at the third and fourth timepoints that while the range of items accurately captured the development of low achieving children, these items were less adequately targeted to higher achieving children. Considering the ELLDI is aimed at preventing children from falling behind in their language and literacy learning, this was not perceived to be a problem in the context of its current use. However, if the intention is to continue developing ELLDI assessments for administration with older and more able children, then more difficult items would need to be developed and calibrated onto the ELLDI Scale to ensure that the assessment is adequately targeting the full range of children's abilities.

Not only was the ELLDI carefully crafted to align the principles of quality assessment in the early years with best practice in early learning (Cloney et al., 2019). It was also constructed using authentic resources and developmentally appropriate items to measure and map the language and literacy growth of the children in the study successfully and accurately. When administered in the field, the elements of the ELLDI aligned with best practice in early learning, such as building relationships, enhancing engagement and the use of authentic experiences, contributed to the success and completion of 1700+ child interviews across the life of the study. Initially this success was based on the ACER fieldworker establishing a warm and responsive one-to-one relationship with the child in the moments after being introduced and prior to the beginning of the interview. During the interview the relationship with the child was maintained using ELLDI administration connector scripts acknowledging the child's contribution and providing details about what was happening next. There were also opportunities during authentic experiences such as



the open-ended oral language items, for the ACER fieldworker to facilitate extended conversations. Anecdotally, for the most part children were relaxed and seemed to feel comfortable being interviewed by an ACER fieldworker. On many occasions children not in the study felt like they were missing out and asked when they would get a turn to go with a fieldworker. Children involved in the study were engaged in the assessment process and seemed to enjoy the focused attention.

Other contributing factors to the successful administration of the ELLDI were the scripts, prompts and the sequence of items. During the administration the use of scripts and prompts meant fieldworkers could give their full attention to the child and not be distracted by thinking about what to say next. This also provided every child with the same assessment experience, which means we can have confidence in using the results to make comparison within and across children, centres and the intervention and control groups. Furthermore, the sequencing and grouping of items helped the children stay focused and answer questions that moved from listening comprehension activities, to naming alphabet letters and then onto the next set of items aimed at measuring other language and literacy skills.

There is no disputing that language and literacy skills are predictors of a range of academic outcomes, including proficiency in reading and writing. With the development of the ELLDI, an innovative measure, educators implementing the EL&L program will soon have the ability to accurately measure children's individual growth. The educator's ability to interpret these assessments and understand what children know and can do will be extended with the use of the ELLDI Scale (a type of learning progression, see Table 52). Children's abilities will be able to be plotted on the scale, with the locations providing educators with information about where children currently are in their learning and by examining higher locations on the scale, educators will be provided with information about where the children need to go next. This is of particular importance for children who are located at either end of the ELLDI Scale, as those who are positioned on the lower end of the scale will require extra support, whereas those positioned higher on the scale will require extension. Examples of such varying needs were observed in timepoint one (2018) when one study child read the book 'Tricky Monkey' from cover to cover, compared to another child who could not identify the letters 'M' and 'T' in the Tricky Monkey title or the produce the sounds these letters make.

### Results of the ELLDI in context

A described scale was able to be produced for the ELLDI, which provides users with a detailed description of the sequence of learning for oral language and literacy across the span of skills covered in the ELLDI. It allows for educators (and researchers and policy makers) to gain an understanding of the skills that each student who is placed on the scale (and in turn located at a level with associated descriptions of learning) currently possesses and what skills are likely to be developed next as they continue to grow. The described scale includes nine different defined levels ranging from foundational skills in expressive vocabulary and oral language (e.g., can name body parts or use words to describe an image) through to more advanced skills in listening and reading comprehension (e.g., makes inferences and generalisations about a story) and writing.

When looking at the strands and skills within the ELLDI Scale, it can be seen that the scale itself explains what children can do within each level. This information can be used by educators to inform decisions about where to next for the children's learning. As the



scale includes a combination of unconstrained and constrained skills consideration needs to be given to the types of strategies used to facilitate learning. Here, it is suggested that the teaching of unconstrained skills be facilitated using a range of learning experiences that are at times explicit but coupled with open and integrative activities. Whereas, constrained skills are taught explicitly so children consolidate and master these skills, with instructional support and integration into everyday experiences. It is acknowledged that the ELLDI Scale is not exhaustive and there is potential to add to the existing skills using further assessments to increase the depth and/or breath of skills within or across levels, or to increase the spread of levels.

The children in the study grew (on average) by approximately one ELLDI level each year, from Level 3 in 2018 to Level 7 in 2022. Some useful benchmarks were established using external criterion to better understand the developmental stages of the sample included in this study. Alignment of the ELLDI Scale with SDG 4.2.1 showed that 81.6% of the ODEC children were meeting minimum standard and 50.3% of children were exceeding minimum standard in 2019 (approx. 5 years of age). Given this is a global indicator for use in developing contexts, we should expect that almost all children should be meeting this minimum standard, which highlights the level of disadvantage in the region ALNF is operating within. This finding persisted through to later stages of the study where 90.3% of the ODEC children were meeting minimum standard and 68.9% of children were exceeding minimum standard against SDG 4.1.1 (a) in 2022 (start of Grade 3). Perhaps the starkest finding was that approximately half of the ODEC children would likely have been meeting minimum standard and approximately 15-20% would have been exceeding minimum standard of the Australian Curriculum Foundation level (first year of school) by the start of Grade 1. Conversely, this means that half of the children would not have met the minimum achievement standard of the Foundation level of the Australian Curriculum by the time they enter Grade 1, an indication of not being ready for school.

### Future development of the ELLDI

Future development of the ELLDI should focus on the potential of the measure to yield rich descriptions of children's learning trajectories that can allow educators to intervene early. For example, it is well-established that children from vulnerable backgrounds are more likely to be behind on their learning by the time they sit NAPLAN at Grade 3. The ELLDI, however, with strong alignment to the NAPLAN reading measure, can describe trajectories that predict success: for example, providing evidence of the necessary trajectory to achieve NAPLAN Band 2 (minimum proficiency on Reading) at Grade 3. Specifically, what skills and abilities are typically acquired, by when, to predict success. In this way, intervention can happen at age 2 and 3 rather than waiting until children reach school.

Similarly, many schools assess children early or "on entry" to find out about the skills and abilities they can demonstrate. Such measures (including Best Start in New South Wales, the Year 1 Phonics Checklist, the PAT assessments), can be linked to the learning trajectories observed using the ELLDI. Measures such as Best Start and PAT that have strong common elements in their frameworks could also be formally equated to allow teachers to report against those scales using the ELLDI.

Further still, where schools require clinical diagnosis or interpretation of psychological assessments to seek targeted funding, the ELLDI could be aligned with this measure to pre-screen children and identify those likely to be in sub-populations that are eligible for

additional support. Whilst the ELLDI is not a clinical measure, it does have the breadth to cover many constructs covered in psychological assessments including the WJIV (Verbal Ability) and SEAPART (Phonics and Phonological Awareness). This may increase opportunities to broaden the use of the ELLDI to support referral into high quality preschool programs, including those offering the EL&L program. If, for example, maternal and child health nurses could identify children with developmental delays they could administer the ELLDI and could refer children into programs where they may otherwise miss out. This would reinforce the nexus between health and development and learning: there is a well-established link between some developmental challenges, like persistent ear infections, and learning.

Further, there is significant opportunity to translate and adapt the ELLDI into other languages. The psychometric techniques used to construct and validate the ELLDI are the same as those used in large scale assessments, including the OECD Programme for International Student Assessment (PISA) where assessments are conducted in close to 100 languages, and over 115 combinations of country and language. The assessments, including for reading are reported on common, international metrics that can facilitate comparisons between languages without bias introduced by language or other cultural differences. The ELLDI, in a different way, could be translated into Indigenous languages, to support assessment of children in communities where English is not their first language. This may apply in settings where ALNF already works and builds capacity in the local community by training local Aboriginal and Torres Strait Islander people to be EL&L practitioners. Such an approach would allow children to demonstrate their abilities before they translate to working in English in formal educational settings. This would provide educators with information about whether a child's oral language development is what would be expected for their age in their home language or whether they need extra support. This is consistent with a strengths-based approach, often called Mother Tongue Education in international contexts. Such approaches are shown to be effective ways to transition children to the school language of instruction.

Lastly, there is significant opportunity to expand the domains assessed in a suite of tools similar to the ELLDI. Children are expected to engage with numeracy concepts in early years settings in school. Like literacy, numeracy is a general skill that is applied in many curriculum contexts and is essential for lifelong success. The development of a numeracy measure would allow ALNF to develop additional interventions in mathematical literacy or the language of mathematics – to analyse, reason, and communicate mathematical ideas and to solve problems in a wide range of contexts. Similar to the big six, existing frameworks can be used to describe, in detail, how children develop abilities in quantity, space and shape, change and relationship, and uncertainty – a framework used in OECD studies including PISA and the International Early Learning and Child Well-being Study (IELS).

## Final ELLDI recommendations

1. **Use the ELLDI Scale to describe a learning progression**, developing tools to orient educators to what learning is being observed, what comes next, and what EL&L materials and strategies are well-targeted to these needs. Such a learning progression has the potential to empower educators with a deep understanding of the reciprocal relationship between language and literacy, and how they develop. It could help educators be less reliant on repeatedly assessing children or teaching to the test and build their confidence in making judgements about how children's language and literacy skills are progressing.
2. **Use the ELLDI to link to the learning outcomes of children on NAPLAN.** For example, children who are in the lowest performance group (those in NAPLAN bands 1 (below the national standard) and 2 (at the national standard)) – the ELLDI can be used to describe the specific skills and abilities these children can demonstrate. Importantly, the significant overlap in the ELLDI and NAPLAN Reading measure means that children's developmental trajectories on the ELLDI can be used to identify children requiring additional support to meet national minimum standards on NAPLAN well before they arrive at school. This may involve empirical work, including formal equating where possible.
3. **Embed the use of the ELLDI into the EL&L program.** The ELLDI becomes a part of the all-important plan, teach, assess and reflect cycle used for effective teaching. This should be embedded in the EL&L program.
4. **Align the ELLDI with major assessments used in school to allow earlier intervention.** For example, what trajectories starting from age 2 predict success on PAT, the Phonics Check, and NAPLAN Reading. Further, align the ELLDI with measures used diagnostically to allow pre-screening of children likely to be eligible for formal additional support.
5. **Broaden the use of the ELLDI to support referral into high quality preschool programs**, including those offering the EL&L program. If, for example, maternal and child health nurses could identify children with, for example, persistent ear infections, they could administer the ELLDI and could refer children into programs where they may otherwise miss out.
6. **Translate and adapt the ELLDI into other languages**, including Indigenous languages, to support children who do not speak English at home. Such approaches can allow children to demonstrate that they are acquiring the language skills that will later help them successfully transition to the language of instruction in school.
7. **Develop a numeracy measure that assesses mathematical literacy** – to analyse, reason, and communicate mathematical ideas and to solve problems in a wide range of contexts. Like literacy, numeracy is a general skill that is applied in many curriculum contexts and is essential for lifelong success.

## Concluding remarks

The EL&L program is demonstrated to have significant promise. Importantly, across the span of the study, ALNF has implemented interim findings to strengthen the program and this resulted in a program that is significantly more focused on teaching at the right level: educators can use good quality assessment (ELLDI) to generate robust evidence of what children can do – their strengths and limitations – and use this to select EL&L materials and to plan at the right level to effectively scaffold children's development. This also increases the capacity of educators by building a shared understanding of children's – supporting professional discussions and improved planning – hallmarks of quality embed in the National Quality Standard. These findings are *not* reflected in the final results of this study – in order to preserve the fidelity of the research design, ALNF did not modify their programming on the Mid North Coast of NSW with the cohort of children participating in the study. It is likely the EL&L program of 2023 is substantially stronger than the one initially observed during the study presented here. It is expected that this will result in stronger impacts on children's learning.

There is a crisis of learning in the region that ALNF are operating the EL&L program. Children's learning in the region, an area typified by significant disadvantage, is lower than we would expect in order to have strong outcomes in school. This is true for all children in the study – Intervention and Control. This is coupled with an underinvestment in intervention in general. Despite efforts to provide programs to children, the EL&L program operates in a small fraction of schools and services in the region. When children move schools – a common occurrence for children from less advantaged backgrounds – there is a high chance they slip out of the program. While it may be assumed that children could or should get the EL&L program for five (or more years) - from the year before preschool to Grade 3 - most children get no more than 12 months of support (this is because they transition out of preschool into a school that is not providing EL&L or some EL&L sites have disengaged from the program, for example due to staff turnover). This should be a stark insight to those wishing to successfully intervene on children's early learning through the school system. The quantum of investment needed to reach children in a sustained way must ensure the support can either: (1) follow the child (no matter which school they attend or move to) or, (2) there is sufficient coverage of schools to ensure children do not fall through the cracks. This would require an expansion of coverage not likely possible given the size of the current program.

It should be noted that a lack of sustained impact as the intensity of the intervention is reduced over time (by attrition, or the movements of families into non-intervention sites) is not to say that ongoing positive effects may not be observed.

*Preschool effects, however, reduce over time and thus there is a need for systematic, sustained, and constant attention to enhancing learning if these early gains are to be maximized (Hattie 2009, p. 58).*

The OECD summarises longitudinal findings in relation to the often-quoted ECEC *fadeout effect* – where benefits are lost, and disadvantage-related gaps remerge in primary school – that shows that despite this, the long-term benefit of high quality ECEC interventions reemerges in adult life. For example, adult earnings are better predicted by achievement at the end of interventions in the early childhood programs than in primary school, and that in studies in the UK, despite observed fadeout effects, benefits remerge in the later years of school.

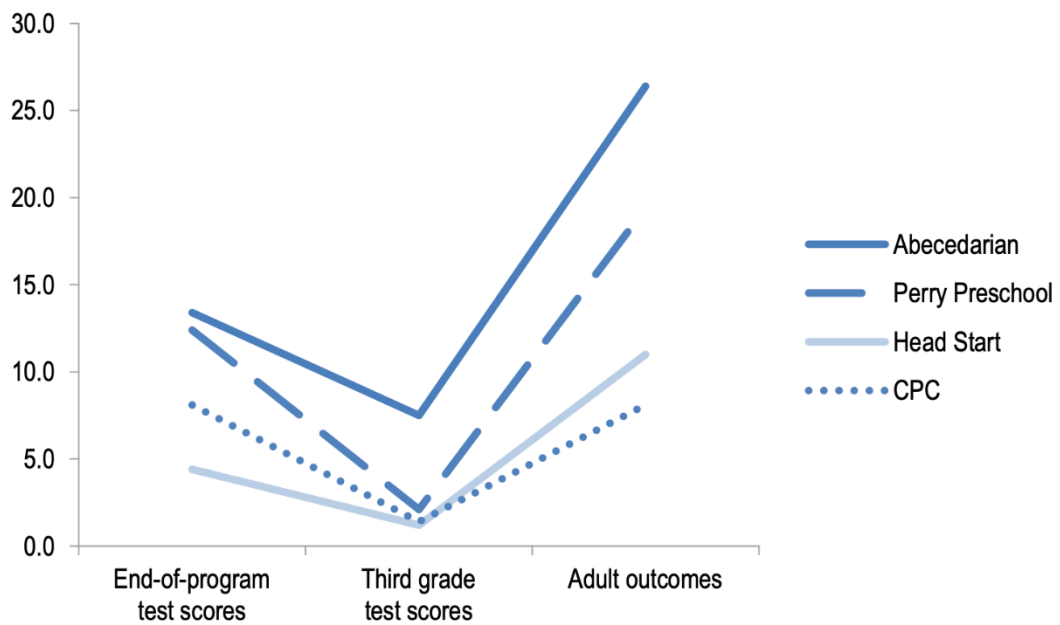


Figure 15. "Predicted percentage effects on adult earnings of early childhood programmes, based on test scores versus adult outcomes" Reproduced from OECD (Shuey & Kankaraš, 2018)

The EL&L program demonstrates a capacity to close achievement gaps between children who access the program. This is notable given the highly disadvantaged population who access the program – disadvantaged even relative to local community that is almost entirely in the lowest 25% of SES in Australia. The sample of children who enter the EL&L program in 2018 begin the study behind their peers. After 12 months of exposure, the learning gap is closed for children in the EL&L program. Over time, these gains are lost as children leave intervention EL&L programs and enter business-as-usual schooling. This is consistent with findings that show that persistent intervention is needed across the early years and lower primary to build a solid literacy and reading foundation.

# References

- Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational Evaluation*, 31(2–3), 162–172. <https://doi.org/10.1016/j.stueduc.2005.05.008>
- Adams, R. J., Cloney, D., Wu, M. L., Osses, A., & Vista, A. (2020). *ACER ConQuest: Generalised Item Response Modelling Software (Version 5)* [Computer software]. Australian Council for Educational Research.
- Adams, R. J., Wilson, M., & Wang, W. (1997). The Multidimensional Random Coefficients Multinomial Logit Model. *Applied Psychological Measurement*, 21(1), 1–23. <https://doi.org/10.1177/0146621697211001>
- Australian Children’s Education and Care Quality Authority (ACECQA). (2020). National Quality Standard. <https://www.acecqa.gov.au/nqf/national-quality-standard>
- Australian Curriculum, Assessment and Reporting Authority. (2021). *Primary Curriculum Review: All learning areas – Foundation year*. Australian Curriculum Review. [https://www.australiancurriculum.edu.au/media/7162/primary-scope-and-sequence\\_-\\_foundation.pdf](https://www.australiancurriculum.edu.au/media/7162/primary-scope-and-sequence_-_foundation.pdf)
- Borko, H., Jacobs, J., Koellner, K. (2010). Contemporary approaches to teacher professional development. *International Encyclopedia of Education*, 7, 548–556.
- Cadima, J., Peixoto, C., & Leal, T. (2014). Observed classroom quality in first grade: Associations with teacher, classroom, and school characteristics. *European Journal of Psychology of Education*, 29(1), 139–158. <https://doi.org/10.1007/s10212-013-0191-4>
- Carlson, E., Jenkins, F., Li, T., & Brownell, M. (2013). The Interactions of vocabulary, phonemic awareness, decoding, and reading comprehension. *The Journal of Educational Research*, 106(2), 120–131. <https://doi.org/10.1080/00220671.2012.687791>
- Chambers, B., Cheung, A. C., & Slavin, R. E. (2016). Literacy and language outcomes of comprehensive and developmental-constructivist approaches to early childhood education: A systematic review. *Educational Research Review*, 18, 88–111. <https://doi.org/10.1016/j.edurev.2016.03.003>
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18(8), 947–967. [https://doi.org/10.1016/S0742-051X\(02\)00053-7](https://doi.org/10.1016/S0742-051X(02)00053-7)
- Cloney, D., & Adams, R. J. (2021). *Conquest version 0.9.95*. <https://cran.r-project.org/package=conquest>
- Cloney, D., & Hollingsworth, H. (2018). Using measures of pedagogical quality to provide feedback and improve practice. In J. Lodge, J. Cooney Horvath, & L. Corrin (Eds.), *Learning Analytics in the Classroom Translating Learning Analytics Research for Teachers* (pp. 221–244). Routledge. <https://doi.org/10.4324/9781351113038-14>
- Cloney, D., & Picker, K. (2021). Developing an assessment of oral language and literacy: Measuring growth in the early years. *Excellent Progress for Every Student: Proceedings and Program*. ACER Research Conference 2021. [https://doi.org/10.37517/978-1-74286-638-3\\_2](https://doi.org/10.37517/978-1-74286-638-3_2)

- Cloney, D., & Picker, K. (2021). Developing an assessment of oral language and literacy: Measuring growth in the early years. *Excellent Progress for Every Student: Proceedings and Program*. ACER Research Conference 2021. [https://doi.org/https://doi.org/10.37517/978-1-74286-638-3\\_2](https://doi.org/https://doi.org/10.37517/978-1-74286-638-3_2)
- Cloney, D., Cleveland, G., Hattie, J., & Tayler, C. (2016). Variations in the availability and quality of early childhood education and care by socioeconomic status of neighbourhoods. *Early Education and Development*, 27(3), 384–401. <https://doi.org/10.1080/10409289.2015.1076674>
- Cloney, D., Jackson, J., & Mitchell, P. (2019). *Assessment of Children as Confident and Involved Learners in Early Childhood Education and Care*. Victorian Curriculum and Assessment Authority. <https://www.vcaa.vic.edu.au/Documents/earlyyears/EYLitReviewLearning.pdf>
- Cloney, D., Picker, K., Jefferies, D., Groom, S., Mendel, M-R., Brace, E., Geddes, M., & Palethorpe, L. (2022, August 22-25). *Assessment in the early years: A symposium about measurement, applications and going to scale*. ACER Research Conference 2022: Reimagining assessment. <https://research.acer.edu.au/rc21-30/rc2022/rc2022/13>
- Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher-child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, 101(4), 912–925. <https://doi.org/10.037/a0016647>
- Egert, F., Fukkink, R. G., & Eckhardt, A. G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis. *Review of Educational Research*, 88(3), 401-433. <https://doi.org/10.3102/0034654317751918>
- Foorman, B. R., Herrera, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). The structure of oral language and reading and their relation to comprehension in kindergarten through Grade 2. *Reading and Writing: An Interdisciplinary Journal*, 28(5), 655–681. <https://doi.org/10.1007/s11145-015-9544-5>
- Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British parents, teachers and children. *Journal of Abnormal Child Psychology*, 38(8), 1179–1191. <https://doi.org/10.1007/s10802-010-9434-x>
- Grifenhagen, J. F., Barnes, E. M., Collins, M. F., & Dickinson, D. K. (2016). Talking the talk: Translating research to practice. *Early Child Development and Care*, 187(3-4), 509-526. <https://doi.org/10.1080/03004430.2016.1246444>
- Groom, S., Mendel, M. R., Brace, E., & Pryor, P. (2022). Evidence-informed early language and literacy practice: Implementation of a new formative assessment for early language and literacy development. ACER Research Conference 2022: Reimagining assessment. <https://doi.org/10.37517/978-1-74286-685-7-11>
- Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child development*, 72(2), 625-638. <https://doi.org/10.1111/1467-8624.00301>



- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 76*(5), 949-967. <https://doi.org/10.1111/j.1467-8624.2005.00889.x>
- Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., Brown, J. L., Cappella, E., Atkins, M., Rivers, S. E., Brackett, M. A., & Hamagami, A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *Elementary School Journal, 113*(4), 461-487. <https://doi.org/10.1086/669616>
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.
- Hattie, J.A.C. (2009). *Visible learning : a synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Hoff, E., & Ribot, K. M. (2015). Language development: Influence of socio-economic status. In J. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2<sup>nd</sup> ed.) (pp. 324-328). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.23132-2>
- Johnston, R. S., & Watson, J. E. (2005). A seven year study of the effects of synthetic phonics teaching on reading and spelling attainment. *Insight, 17*, 1-9. <https://webarchive.nrscotland.gov.uk/3000/https://www.gov.scot/Resource/Doc/933/0044071.pdf>
- Juel, C. (2006). The impact of early school experiences on initial reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2) (pp. 410-426). The Guilford Press.
- Kennedy, M. M. (2016). How Does Professional Development Improve Teaching? *Review of Educational Research, 86*(4), 945-980. <https://doi.org/10.3102/0034654315626800>
- Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*(2), 184-189. <https://doi.org/10.1001/archpsyc.60.2.184>
- Khoo, S. T., & Meiers, M. (2006). Literacy in the first three years of school: A longitudinal investigation. *The Australian Journal of Language and Literacy, 29*(3), 252-267.
- Konza, D. (2014). Teaching reading: Why the “Fab Five” should be the “Big Six.” *Australian Journal of Teacher Education, 39*(12), 153-169. <https://doi.org/10.14221/ajte.2014v39n12.10>
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research, 88*(4), 547-588. <https://doi.org/10.3102/0034654318759268>
- La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the pre-kindergarten year. *The Elementary School Journal, 104*(5), 409-426. <https://doi.org/10.1086/499760>
- Levickis, P., Cloney, D., Roy-Vallières, M., & Eadie, P. (2023). Associations of specific indicators of adult-child interaction quality and child language outcomes: What teaching



practices influence language?. *Early Education and Development*, 1-20.  
<https://10.1080/10409289.2023.2193857>

Linacre, J. (1994). *Many-faceted Rasch measurement*. MESA Press.

Lonigan, C. J., & Shanahan, T. (2009). Developing Early Literacy: Report of the National Early Literacy Panel. Executive Summary. A Scientific Synthesis of Early Literacy Development and Implications for Intervention. *National Institute for Literacy*.

Marsman, M., Maris, G., Bechger, T., & Glas, C. (2016). What can we learn from Plausible Values? *Psychometrika*, 81(2), 274–289. <https://doi.org/10.1007/s11336-016-9497-x>

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149–174. <https://doi.org/10.1007/BF02296272>

Meiers, M., Khoo, S. T., Rowe, K., Stephanou, A., Anderson, P., & Nolan, K. (2006). *Growth in Literacy and Numeracy in the First Three Years of School* (ACER research monograph 61). Australian Council for Educational Research.  
[https://research.acer.edu.au/acer\\_monographs/1](https://research.acer.edu.au/acer_monographs/1)

Moen, A. L., Sheridan, S. M., Schumacher, R. E., & Cheng, K. C. (2019). Early childhood student–teacher relationships: What is the role of classroom climate for children who are disadvantaged?. *Early Childhood Education Journal*, 47, 331-341.  
<https://doi.org/10.1007/s10643-019-00931-x>

Monseur, C., & Adams, R. J. (2009). Plausible values: How to deal with their limitations. *Journal of Applied Measurement*, 10(3), 320–334.

Morrow, L. M., & Rand, M. K. (1991). Promoting literacy during play by designing early childhood classroom environments. *The Reading Teacher*, 44(6), 396-402.  
<https://www.jstor.org/stable/20200675>

Morrow, L. M., Strickland, D. S., & Wood, D. G. (1998). *Literacy instruction in half-and whole-day kindergarten*. International Reading Association.  
<https://archive.org/details/literacyinstruct0000morr/page/n3/mode/2up>

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, 27(4), 342–356. <https://doi.org/10.1111/j.1467-9817.2004.00238.x>

National Institute of Child Health and Human Development (NICHD). (2000). *Teaching children to read: An evidence-based assessment of the scientific literature on reading and its implications for reading instruction*. US Department of Health and Human Services.  
<https://www.nichd.nih.gov/publications/pubs/nrp/smallbook>

New South Wales Department of Education. (2023). Best Start Kindergarten Assessment.  
<https://education.nsw.gov.au/teaching-and-learning/curriculum/literacy-and-numeracy/assessment-resources/best-start-kindergarten>

Niklas, F., Nguyen, C., Cloney, D., Tayler, C., & Adams, R. (2016). Self-report measures of the home learning environment in large scale research: Measurement properties and associations with key developmental outcomes. *Learning Environments Research*, 19(2), 181–202. <https://doi.org/10.1007/s10984-016-9206-9>

- OECD. (2011). How do some students overcome their socio-economic background? *PISA in Focus*, 5. <https://doi.org/10.1787/5k9h362p77tf-en>
- Paatsch, L., & Nolan, A. (2020). Supporting oral language development in young children. In A. Kilderry & B. Raban (Eds). *Strong foundations: Evidence informing practice in early childhood and care* (pp. 156-169). ACER Press: Camberwell, Victoria. <https://doi.org/10.37517/978-1-74286-555-3>
- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, 40(2), 184-202. <https://doi.org/10.1598/RRQ.40.2.3>
- Pianta, R. C., La Paro, K. M., & Hamre, B. (2008). Classroom assessment scoring system manual: Preschool (pre-k) version. Brookes.
- Picker, K. (2022). *Examining the relationship between teacher reading content knowledge, pedagogy, and children's learning experiences* [Doctoral dissertation, University of Melbourne]. Minerva. <https://minerva-access.unimelb.edu.au/bitstreams/ac544b0c-50ce-40b4-92b8-a2898966d3e6/download>
- Pink, B. (2008). *Socio-Economic Indexes for Areas (SEIFA): Technical Paper. ABS Catalogue no. 2039.0.55.001* (Socio-Economic Indexes for Areas (SEIFA): Technical Paper B2 - Socio-Economic Indexes for Areas (SEIFA): Technical Paper). Australian Bureau of Statistics.
- Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, 45(3), 605. <https://doi.org/10.1037/a0015365>
- Prior, M., Smart, D., Sanson, A. N. N., & Oberklaid, F. (2000). Does Shy-Inhibited Temperament in Childhood Lead to Anxiety Problems in Adolescence? *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 461–468. <https://doi.org/10.1097/00004583-200004000-00015>
- Queensland Department of Education. (2023). Early Start. <https://education.qld.gov.au/schools-educators/school-types/early-years/early-start>
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <http://www.R-project.org>
- Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, 45(4), 958. <https://doi.org/10.1037/a0015861>
- Rose, J. (2006). *Independent review of the teaching of early reading: Final report*. UK Department for Education and Skills. <https://dera.ioe.ac.uk/id/eprint/5551/2/report.pdf>
- Rowe, K. (2005). *Teaching reading: National inquiry into the teaching of literacy*. Australian Government Department of Education, Science and Training. [https://research.acer.edu.au/tll\\_misc/5](https://research.acer.edu.au/tll_misc/5)
- Sanson, A. V., Smart, D. F., 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*, 40(2). <https://www.jstor.org/stable/23087863>

- Shuey, E. A., & Kankaraš, M. (2018). The power and promise of early learning (OECD Education Working Papers No. 186). OECD. <https://doi.org/10.1787/f9b2e53f-en>
- Siraj, I., Melhuish, E., Howard, S. J., Neilsen-Hewett, C., Kingston, D., De Rosnay, M., Huang, R., Gardiner, J., & Luu, B. (2023). Improving quality of teaching and child development: A randomised controlled trial of the leadership for learning intervention in preschools. *Frontiers in Psychology, 13*, 1092284. <https://doi.org/10.3389/fpsyg.2022.1092284>
- Soliday Hong, S., Zadrozny, S., Walker, J., Love, E. N. G., Osborne, J. D., Owen, J. L., Jenkins, G., & Peisner-Feinberg, E. (2021). *Longitudinal Study of Georgia's Pre-K Program: Third Grade Report*. The University of North Carolina, FPG Child Development Institute.
- Sylva, K., Siraj-Blatchford, I., & Taggart, B. (2010). *ECERS-E: The Early Childhood Environment Rating Scale Curricular Extension to ECERS-R*. Trentham Books. <https://eric.ed.gov/?id=ED511422>
- Taylor, 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), 13–21. <https://doi.org/10.1177/183693911303800203>
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration (BES)*. <https://www.educationcounts.govt.nz/publications/series/2515/15341>
- Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities, 45*(5), 453–466. <https://doi.org/10.1177/0022219411432685>
- UNICEF. (2021). *Early Childhood Development Index 2030*. <https://data.unicef.org/resources/early-childhood-development-index-2030-ecdi2030>
- UNICEF. (2022a). *SDG indicator metadata (4.2.1)*. <https://unstats.un.org/sdgs/metadata/files/Metadata-04-02-01.pdf>
- UNICEF. (2022b). *SDG indicator metadata (4.1.1)*. <https://unstats.un.org/sdgs/metadata/files/Metadata-04-01-01.pdf>
- United States Agency for International Development, UNESCO Institute for Statistics, The World Bank Group, Foreign, Commonwealth & Development Office, Australian Council for Educational Research, & Bill and Melinda Gates Foundation. (2020). *Global proficiency framework for mathematics: Grades 1 to 9*. USAID. <https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2019/05/GAML6-REF-16-GLOBAL-PROFICIENCY-FRAMEWORK.pdf>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate Imputation by Chained Equations in R. *2011, 45*(3), 67. <https://doi.org/10.18637/jss.v045.i03>
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*, 427–450. <https://doi.org/10.1007/BF02294627>
- Werner, C. D., Linting, M., Vermeer, H. J., & Van IJzendoorn, M. H. (2016). Do intervention programs in child care promote the quality of caregiver-child interactions? A

meta-analysis of randomized controlled trials. *Prevention Science*, 17, 259-273. <https://doi.org/10.1007/s11121-015-0602-7>

Wheldall, K., Bell, N., Wheldall, R., Madelaine, A., & Reynolds, M. (2019). Performance of Australian children on the English Phonics Screening Check following systematic synthetic phonics instruction in the first two years of schooling. *Australian Journal of Learning Difficulties*, 24(2), 131-145. <https://doi.org/10.1080/19404158.2019.1635500>

Willms, J. D., & Shields, M. (1996). *A measure of socioeconomic status for the National Longitudinal Study of Children*. Atlantic Center for Policy Research in Education.

Wu, M. L. (1997). The development and application of a fit test for use with marginal maximum likelihood estimation and generalised item response models [M. Ed, The University of Melbourne]. <http://hdl.handle.net/11343/57769>

Wu, M. L. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2-3), 114-128. <https://doi.org/10.1016/j.stueduc.2005.05.005>

# Appendices

Note: Appendices 2-16 are located in a separate document.

Appendix 1: ELLDI assessment booklets

Appendix 2: Oral language coding rubrics

Appendix 3: Room Record form

Appendix 4: Director questionnaire

Appendix 5: Educator questionnaire

Appendix 6: Parent/Caregiver questionnaire

Appendix 7: Centre information sheet

Appendix 8: ODEC Field Operation Manual

Appendix 9: Data entry manual

Appendix 10: Model convergence plots

Appendix 11: Item and booklet mapping

Appendix 12: Raw frequencies

Appendix 13: Item-level statistics

Appendix 14: Item characteristic curves (ICCs)

Appendix 15: Item parameter estimates

Appendix 16: Item-Person maps of the sample distribution relative to ELLDI Scale levels

# Appendix 1: ELLDI Assessment Booklets



ODEC: Child Literacy and Language Response Book: **ELLDI example items**

EXPRESSIVE LANGUAGE
<p><i>Materials: None</i></p> <p><i>STOP RULES: None</i></p>

**I will point to some parts of my body. Tell me the name of this body part, for instance (pointing to nose), this is my nose.**  
*Point to the following body parts (on enumerator's body) one by one and ask child to name the different body parts.*

		Correct	Incorrect	Child says I do not know or no response
	Correct answer	<b>1</b>	<b>0</b>	<b>99</b>
<b>Tell me the name of this body part.</b> <i>(point to eye)</i>	Eye/eye lash /eyeball/eyelid	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Tell me the name of this body part.</b> <i>(point to ear)</i>	Ear	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Tell me the name of this body part.</b> <i>(point to elbow)</i> <i>Prompt: If child says arm, prompt once.</i>	Elbow	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

EXPRESSIVE VOCABULARY – NOUNS, VERBS, TELL A STORY

Materials: Water Hole Scene Picture and Recorder

STOP RULES: none

Turn on the recorder and face it towards the child. Pass the child the picture.  
**Now please hold this picture and look at it carefully. Then I will ask you some questions about it. Give the child 30 seconds to look at the picture.**  
**Tell me the names of different things you can see in the picture.**  
 Do not count repeated responses eg, naming more than one man.  
 Prompt: if child stops at less than 10 things, prompt:  
**What else can you see in the picture?**  
**Stop at 10 things or if the child does not respond to the prompt**



Record child's response on digital recorder.

Correctly names 6-10 things

1  
○

Correctly names 1-5 things

0  
○

Child says I do not know or no response

99  
○



Skill	Oral language task: Water Hole Picture Prompt - nouns, verbs, story								
	88	99	0	1	2	3	4	5	
<b>Structure &amp; detail</b>	Missing-child not given the task	Missing-child given the task but did not/refuse to respond	Minimal (1-3 words) or incomprehensible response.	Child mainly uses single words that reflect a very limited vocabulary of familiar words e.g. 'man', 'dog', 'water'.	Child mainly uses very short phrases of generally 2-3 words, e.g. 'girl bike', 'man read book', 'they play', 'plane fly' that reflect a range of familiar vocabulary.	Child mainly uses simple, extended responses (4 words or more) that convey meaning AND reflect a range of familiar vocabulary AND often contain some grammatical errors, e.g. 'splashing, the boy is doing that', 'dog eating cake is naughty', 'boy is sad swing is broke'.	Child mainly uses simple, extended responses (4 words or more) that convey meaning AND reflect a range of familiar vocabulary with NO grammatical errors e.g. 'This boy is splashing that one.' 'That naughty dog is eating the cake.' 'The boy is sad because the swing is broken.'	Sentences may be joined with simple conjunctions e.g. 'and', 'then', 'but'.	Child mainly uses many complex sentences that explain relationships between ideas AND reflect a wide vocabulary with NO grammatical errors e.g. 'the girl is wondering how she is going to get home, because the tyre on her bike is flat', 'the man is so busy reading his book that he has not noticed the dog is eating the cake', 'the man is watching the boys to make sure they do not drown'.
<b>Prompts required for nouns</b>	Missing-child not given the task	Missing-child given the task but did not/refuse	3 or more prompts required	1-2 prompts required	No prompts required				

NAME WRITING

*Materials: Pencil for child, hard surface for writing (for example, table, book), back of the letter identification sheet.*

*STOP RULES: If the child does not write for one minute after your instructions or if the child takes longer than 2 minutes to write, stop and say: **We're going to move on to our next activity now.***

Instructions

*Place the letter identification sheet in front of the child (face down so the child sees a blank piece of paper). Place the pencil next to the paper.*

	Correct	Incorrect	Child says I do not know or no response
	1	0	99
<b>Say: Now we are going to write. Write your name at the top of this piece of paper.</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<i>Point to the blank space at the top of the paper. Be encouraging but do not help the child.</i>			

1: Correct answer: Child writes name correctly including:  
 -all letters of name (correct orientation)  
 -in correct order

	Has letters in name, but name is not correct- letters are out of order; other mistakes are made	Non-name letters	Symbol-like marks	Scribbles, no discernible symbols
	4	3	2	1
<i>If incorrect, describe what the child wrote. Check one circle.</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Check if stop rule was used at one minute.	<input type="radio"/>			
Check if stop rule was used at two minutes.	<input type="radio"/>			



Materials:  
STOP RULES: None

**Now we're going to find small words in big words. I'll show you what I mean. Listen to the word 'popcorn. First I say 'pop' (make a fist as you say this and keep this position), then I say 'corn, (make a fist with the other hand as you say this and move fists so that they are side-by-side. Move the two fists together to touch and say) popcorn!**

**If I take away the 'pop' (move first fist behind your back or drop to your lap), then the word left is 'corn.**

*Note, if you are sitting across from the child use your right fist to represent the first word (form their perspective it will look like the leftmost representation of the word) This is a practice item.*

Instructions and Questions

**Now you try one.** (present both fists lined up together and as you say the whole word with a natural pace)

**Here's the big word (pause) rainbow.**

**If I take away 'bow' (move fist away so only one fist remains) what word is left?**

**Correct answer: Rain**

correct	incorrect	I don't know or no response
<b>1</b>	<b>0</b>	<b>99</b>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Instructions and Questions

**Now we're going to break a word into parts. I'll show you what I mean.**

**'Robot' is a word. I can break it into parts: 'ro' (clap), 'bot' (clap), 'ro' (clap), 'bot' (clap).**

**You try.** Child attempts to clap and say the syllables in robot.

*This is a practice item.*

**Now you try one.**

**Say 'baby'.** Child says 'baby'.

**Now break 'baby' into parts.**

*Correct answer: 'ba'...'by' with a break in between – with or without clapping*

Correct	Incorrect	No attempt
<b>1</b>	<b>0</b>	<b>99</b>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Instructions and Questions

**Listen carefully and tell me, do these words have the same first sound?**

*(pause after each pair of words for response and scoring)*

	correct	incorrect	No attempt
<b>Pot ... pen</b>	1	0	99
1: <u>Correct answer: yes</u>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Pot ... fan</b>	1	0	99
1: <u>Correct answer: no</u>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Pot ... top</b>	1	0	99
1: <u>Correct answer: no</u>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Instructions and Questions

**Now we're going to take the *first* sound out of a word, to make a new smaller word. I'll show you what I mean. The word is bend. You say the word 'bend'. Child says bend.**

**If we take out /b/ we make a new word - end.**

**Now say it again but don't say /b/. Child says 'end.'**

*This is a practice item.*

	Correct	Incorrect	No attempt
<b>Now you try taking the <i>first</i> sound out of hand.</b>	1	0	99
<b>Say 'hand'. Child says hand.</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Now take out /h/.</b>			
<b>What is the new word?</b>			
<i>The new word is 'and'.</i>			

	Correct	Incorrect	No attempt
<b>Now you try taking the <i>first</i> sound out of fall.</b>	1	0	99
<b>Say 'fall'. Child says fall.</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Now take out /f/.</b>			
<b>What is the new word?</b>			
<i>The new word is 'all'.</i>			

Materials: *Picture Story Book – Wombat Trouble*  
 STOP RULES: None

Instructions and Questions

Show me which way to go when I'm reading.

Indicates  
 left to right  
 (does not  
 matter  
 which line  
 of text child  
 indicates)

Other

No  
 attempt

1

0

99





"Zip up the tent, we'll go for a walk" said Aunty.  
 "Wait for me!" called Banjo.

2

Instructions and Questions

Open the book to pages 1 & 2.

Let's talk more about the story.

Banjo and Aunty have put up the tent. What are they doing now?

Going for a  
 walk; Banjo  
 running to  
 catch up to  
 Aunty

Other

No  
 attempt

1

0

99

READ A SENTENCE

Materials: Koala Reading Sentence  
STOP RULES: none

Instructions and Questions

**Here is a picture and sentence about koalas.**

*Place the sheet in front of the child.*

	Reads all words correctly	Reads most words correctly	Reads at least 4 words correctly	Other	No attempt
<b>Read the sentence for me?</b>	<b>3</b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>99</b>
<i>Run your finger under the sentence from beginning to end.</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*If the child is unable to read the sentence ask, **Can you read any words in the sentence?***



"A little koala is on top of the big koala" said Tarni.

CONNECTOR SCRIPT

**Well done!**

INDEPENDENT READING TASKS

Materials: Independent reading administration sheet

STOP RULES: none

Read the sentence and then answer the questions by shading the bubble next to the correct answer.

Banjo stirs the soup with a wooden spoon.  
What is Banjo doing?

washing

cleaning

cooking

shopping

---

Helen needs to put the autumn leaves in the bin.  
What can she use to clean up the leaves?

a rake

a bush

a sock

a pencil

CONNECTOR SCRIPT

**You did such a great job. Thank you so much for all those great answers. Give sticker.**

*Assist the child to reintegrate into the normal activities in the classroom.*

[www.acer.org](http://www.acer.org)

**Australian Council for Educational Research**

*Creating and promoting research-based knowledge,  
products and services to improve learning.*



**ALNF.ORG**