

# Including the excluded: Young language learners and fair assessment in school and early childhood

Denise Angelo<sup>1</sup>

Australian National University

Catherine Hudson

Australian National University

Julie Luxton

Evaluation Associates Ltd | Te Huinga Kākākura Mātauranga

Assessment plays a key role in recognising the multilingual capacities of young (early childhood, primary and secondary) language learners and in revealing and supporting their language learning needs. This special issue of *Studies in Language Assessment* (SiLA) focuses on identifying and responding to invisibilities and gaps in assessment for young multilingual learners from diverse language backgrounds. This is a complex and exciting area for language assessment with demonstrable social justice and policy dimensions. When assessment goes wrong, it can lead to a lack of fairness in outcomes in multiple ways, such as deficit positioning of learners, information for educators that is misleading, inappropriate interventions, truncated educational opportunities and unfulfilled student, family and community aspirations. In contrast, the creative body of work presented here illuminates such assessment mismatches and exemplifies developments in assessment tools that show first language (L1) strengths and enable second language (L2)<sup>2</sup> learners. Significantly, all contributions involve at least one author who is, or was, embedded in the education assessment context addressed in

---

<sup>1</sup> Email address for correspondence: [denise.angelo@anu.edu.au](mailto:denise.angelo@anu.edu.au)

<sup>2</sup> The term 'L2 learners' will be used to refer to all additional language learning contexts, including foreign/classroom only contexts, language revival or revitalisation settings etc. 'L2 English learners' covers terms such as ELLs (English Language Learners) in New Zealand; and EAL/D learners (English as an Additional Language or Dialect) learners, the term used currently in the Australian school sector, except for the states of Victoria and Tasmania where the term EAL learners (English as an Additional Language) is used.

their paper and who consequently brings a keen appreciation of the impact of assessment and an urge for improvement.

Recognition and responses to exclusions from fair assessments run along a continuum of activity, from the issue of identification and awareness raising at one end, through to practical alternative practitioner processes, and tool development and implementation at the other. This special issue encompasses the full gamut of such endeavours. Following Vogt and Tzagari (2022), there is a need to widen the discussion of assessment development from the concerns of standardised testing and associated stakeholders such as test developers, test publishers and researchers. A wider discussion, and a wider inclusion of stakeholders, reveal the gaps in assessment/testing development, the areas where cohorts have been excluded, and importantly ‘on the ground’ innovatory responses which provide more inclusive assessment for young multilingual learners, L1 speakers and/or L2 learners.

The contributions to this special issue thus include research papers and reports by practitioners with direct experience of the young people and the linguistic and educational contexts in which the discussed assessment takes place. Such first-hand experiences provide authentic observations and grounded motivations for the articles. The authors are practising teachers, educators in support or advisory positions, and Indigenous community members, as well as academics and researchers in the fields of applied linguistics and language assessment. They all bring fresh, real world perspectives to the field of language assessment. Each contribution is concerned with “including the excluded” which is fundamental to fair assessment of young language learners. This extends the gaze of the professional language assessment and testing community beyond the familiar and into new language situations such as language revitalisation, new learning contexts such as language and cultural knowledge, and new cohorts such as early childhood assessment in multilingual settings.

This special issue is a mark of real commitment by the Association for Language Testing and Assessment of Australia and New Zealand (ALTAANZ) to engaging with school and early childhood matters. The guest editors believe ALTAANZ efforts deserve special mention here, as it could be a model for consideration of other language assessment, linguistics and applied linguistics organisations. In order to pursue an organisational aim of becoming more actively and deeply involved with the

education sector and with educators, ALTAANZ has for some years officially included the role of teacher representative amongst its elected office bearers. Monthly ALTAANZ committee meetings allocate time for ‘teacher rep’ reports. As a result, committee members are briefed, on a monthly basis, on current language assessment issues faced by educators in Australia and New Zealand. One outcome has been that ALTAANZ and/or ALTAANZ members have contributed to submissions to government enquiries that relate to language assessment in school and early childhood settings, raising the profile of the language testing/assessment field, as well as providing specialist advice. ALTAANZ has also committed funding to educator oriented activities, such as funding rounds to support travel for teachers presenting at ALTAANZ conferences and to support language assessment activities (see Browne in this issue).

This special issue has grown out of ALTAANZ teacher and education-oriented initiatives. The guest editors are currently serving ALTAANZ teacher reps and contributors are from far-flung educator networks. The topic was developed from an ALTAANZ-supported colloquium, “The invisibility of language in ‘mainstream’ assessment tools” presented at the ALANZ<sup>3</sup>/ALAA<sup>4</sup>/ALTAANZ Applied Linguistics Conference in November 2022 in Wellington, New Zealand. Papers delivered at that conference by Angelo and Hudson, Erlam, Crosby and Franken are all represented in this full special issue devoted to language assessment in early childhood, primary and high school settings.

## **Section 1. Language learners and mismatched mainstream and standardised assessments**

This section investigates the use of mainstream and standardised assessments that are mismatched to L2 learners. The papers in this section present evidence that aims to influence education policy, so that assessment becomes more inclusive of an invisible or partly invisible L2 learner cohort.

---

<sup>3</sup> Applied Linguistics in Aotearoa New Zealand

<sup>4</sup> Applied Linguistics Association of Australia

Denise Angelo and Catherine Hudson report on how young L2 English learners have only ‘phantom’ representation in the Australian Early Development Census (AEDC), a mainstream assessment tool implemented nationally every 3 years in the first year of schooling. The AEDC has high status in early years assessment discourse in Australia and is taken up in early years policy on account of its perceived legitimacy. However, despite the efforts it makes to include L2 English learners, the authors argue that due to the inherently monolingual orientation of the questions that teachers answer – especially in domains highly relevant to language use – the AEDC lacks validity for the cohort of L2 learners with early levels of English language proficiency. L2 English learners remain shadowy figures eluding clear identification and drift into deficit reporting which is most unhelpful for informing and guiding policy. The study shows how this problematic AEDC data can be accepted unquestioned, and inappropriate and inaccurate perspectives of young L2 English learners can then be promulgated, with potential impact on this cohort. The authors propose that reporting on children’s L2 English proficiency levels, and their languages other than English, is likely to focus attention on this cohort’s language strengths and L2 language learning needs, in contrast to the current potential for deficit interpretations.

Susan Poetsch provides a case study which examines the difficulties classroom educators face when interpreting the learning achievements of L2 English learners in mainstream curriculum areas which are taught in the English medium, when advice is not provided in the Australian Curriculum resources. The study takes the example of a remote Aboriginal community in central Australia where children speak a traditional language as their L1 and the main everyday vernacular of family and community interactions, so in practice English is a foreign language, employed only in the classroom. Written and oral data exemplifying the students’ curriculum learning and their early levels of L2 English proficiency demonstrates the lack of fit with the available assessment guidance offered by the Australian Curriculum. The mainstream curriculum Achievement Standards do not speak to the classroom outputs of students at early levels of L2 English acquisition in any straightforward way, leaving teachers “unguided” when making assessment judgements. What is more, Poetsch argues, the absence of any assessed L2 learner work samples within each Curriculum Area on the Australian Curriculum website renders L2 English language learners effectively invisible. In effect, the intensive language support needed for early level L2 English

learners to engage with English-medium mainstream curriculum content is nowhere represented, again leaving their teachers “unguided” as the title of this paper asserts.

Margaret Franken’s paper focuses on the Common Assessment Activities (CAAs), the national standardised tests in literacy for post primary learners in Aotearoa New Zealand. The standards are to be mandatory by 2026 and passing the tests will be a co-requisite for New Zealand’s senior years’ qualification, effectively supplanting other school achievements. Franken discusses the trialling of the two literacy tests in 2021-22 and issues with the underlying constructs, process and design of the tests that she posits may have played a part in influencing the alarmingly low results, especially for students who are L2 learners of English, Pasifika (other Pacific Island heritages), Māori, and those in the Realm nations (Cook Islands and Niue). The study explores how the lack of clarity about the construct of a ‘base level’ of foundational literacy influences the fairness of the tests. For example, Franken’s analysis finds a lack of fit between the level of text difficulty for the 2022 Reading CAAs and the reading age of the 13 and 14 year olds who take the test; and blindness to the language demands for varying learner groups such as L2 English learners. Franken also raises pertinent questions about the Writing CAAs. For example, regarding the rubric for accuracy in a longer writing text, she asks, in terms of foundational literacy and sociolinguistic variation, what does “sufficient accuracy to communicate ideas(s)/information clearly” look like? The author finds the tests fall short in terms of their inclusion and equity. In their current position as a co-requisite, she warns that for those not achieving the new standards the tests will potentially impede their education pathways and schooling success in gaining a senior years’ qualification.

## **Section 2. Improving ways of working and assessing for language learners**

This section explores improvements in ways of working and assessing for language learners, locally in teacher training or in schools, in the realm of flexibility outside the mandates of national and state assessment policies and assessment regimes.

Rosemary Erlam adds to the evidence for a positive story about teacher assessment literacy when teachers are given autonomy and support. Erlam’s study focuses on the

inquiries made by three New Zealand teachers into the validity of assessment practices from their teaching contexts when used with L2 English language learners. This study is based on the teachers' 'inquiry' assignments, which were undertaken partway through an intensive TESOL assessment course. Erlam examines the processes and decisions the teachers make in their inquiry critiques through the lenses of two theoretical frameworks, both of which encompass concepts about fairness. Erlam uses Hill's (2017) teacher assessment literacy framework to map out the teachers' reflections and subsequent questioning of the appropriacy of a classroom assessment practice. Additionally, she uses an Assessment Use Argument framework to show the steps the teachers followed which led them to challenge the usefulness of the focus practices for L2 learners of English. In this light, Erlam finds that the teachers are able to make informed decisions about assessments. They are able to critically evaluate assessment practices in their local contexts, to determine the extent these are appropriate when used with L2 English learners, and to propose fairer adaptations to improve assessment validity.

In the New Zealand setting again, Simon Crosby's paper examines how an asset-based, translanguaging approach to assessment enables bi/multilingual early secondary school students to demonstrate their learning of subject content. Crosby contrasts this with assessment processes that have a monolingual bias that preclude such students from showing that they have met curriculum learning criteria. The study is set in a highly language diverse junior college in urban Auckland where project-based learning is being implemented using peer assessment practices. Crosby analyses power-related peer assessment transcripts to show instances of where assessment goes wrong. Underachieving, bi/multilingual students appear in a deficit light, or are frustrated, and as one student states it, "come out feeling weaker". He confronts the complexities of school policy change, but traces the efficacy of translanguaging assessment principles of flexibility, integration, collaboration and criticality that could afford useful guidance in curriculum assessment practices. Crosby shows that enabling students to draw on their full linguistic repertoires leads to more equitable assessment outcomes, but points out current constraints on implementation such as a predominantly monolingual English workforce and school policy which do not reflect the student population.

In the Northern Territory of Australia, Emma Browne also explores the full use of students' language repertoires in her report on school assessment practices in a remote Warlpiri speaking Aboriginal community. She provides on-the-ground documentation of collaborative exercises for educators on planning assessment in a linguistically complex education situation. While incomer teachers are English speakers, the students and local teachers in the school are L1 speakers of Warlpiri, the traditional Aboriginal language of the area, and the community supports Warlpiri-English bilingual schooling. The school delivers the Australian Curriculum, which is designed for English speaking students and teachers, differentiating for bilingual delivery using the Northern Territory Indigenous Languages and Culture bilingual pathway document (which is generic, i.e. not specifically for Warlpiri) and for local language and culture with the Warlpiri Theme Cycle. Browne reports that although the school aspires to deliver bilingual education, it is a constant challenge at many levels and students' Warlpiri language achievement is not currently reported on. One assessment intervention described in this paper was the development and workshopping of a language and literacy scope and sequence for the Warlpiri Theme Cycle by year level teaching teams, with Warlpiri and English speaking teachers working together. The second assessment intervention described in this paper is the trialling of bilingual assessment for science curriculum content learning with a bilingual teaching team. A key message in the paper is the need for system awareness and support for the role of languages in learning. A key learning from the workshops and the unit of teaching is the importance of educators learning together in cross-linguistic teaching teams to generate and disseminate effective, linguistically inclusive and shared assessment practices.

In an Aboriginal language revival context in south-eastern Australia, Jasmine Seymour, a Dharug woman and language teacher and researcher of Dharug, the Aboriginal language traditionally spoken in urban Sydney, and Denise Angelo, a non-Indigenous Aboriginal language teacher educator, explain the current context of assessment in school Aboriginal language revival programs in the state of New South Wales. The authors describe how the generic (not language specific) state and national Aboriginal language curriculum documents give little guidance for developing and delivering Aboriginal language programs or for assessing student learning. Seymour and Angelo make proposals about how to progress assessment practices in school

Aboriginal language revival programs positively. They navigate the tension between the “potential unease” among some Aboriginal community members about assessment and their languages, and the role of assessment in course programming and academic credit for students’ language learning. In view of the fact that Aboriginal languages have been excluded from education until relatively recently and Aboriginal language teacher training and professional pathways remain under-developed, the authors recommend that accessible information and tangible examples of assessing the language taught in class be made available: there has been hitherto little if any prior experience of school-based language assessment items. In Seymour’s experience, community members approve of the aims of well-constructed language assessment tasks as they desire supportive and respectful teaching of their languages. The authors also propose that the current Aboriginal language curriculum documents should have a bank of language assessment tasks suitable for adaptation, a process they illustrate via two sample tasks. Further, they suggest there be a support mechanism to develop language-specific curriculum, including assessment, should a language community decide to pursue a school language program.

### **Section 3. Development and implementation of purpose-built assessment tools**

In this section, the contributions describe the development and implementations of purpose-built language assessment tools. These language assessment initiatives fill gaps: gaps in assessment tools available for specific languages, in design criteria for language assessment tools, in processes for implementing language assessment and in assessment tools for understudied multilingual settings.

The paper by Rosalie Grant, Gayle Arnaqulluk Miller and Gary Cook discusses the *Yupiit Piciryaraitnek Qaneryaranek-llu Cuqyun--Yup’ik Culture and Language Measurement* project, the first school-based Indigenous language and culture assessment tool of its kind in Alaska. The paper outlines the inception of the assessment project, which was initiated through Yup’ik community members and their Expert Group in order to respond to the expressed aspirations of Yup’ik communities for children to be maintaining their language and culture. The Expert Group sought the assistance of language assessment researchers to engage in a community-based,



co-designed process founded on participatory and collaborative research principles to develop a linguistically and culturally sustaining assessment tool. The tool for kindergarten to grade 6 students, administered using the Yugtun language, is also designed to meet the relevant requirements for school-based assessment systems of the United States federal government. Grant, Miller and Cook's paper focusses on the Yup'ik Cultural Awareness subtest, a pivotal component of the overall language assessment tool as it illustrates how cultural knowledges selected by the appropriate community experts can be brought into the design of school assessment. This is of special significance for many Indigenous groups in settler colonial countries as their knowledges and practices have historically not been represented as core curriculum. This paper adds significantly to what has been the limited development of assessment tools for Indigenous languages and cultures.

Catherine Hudson, Denise Angelo and Sue Creagh's paper lays out considerations for "policy responsible" (Elder, 2021) L2 English proficiency assessment, by examining characteristics instantiated in an assessment tool developed in the Australian state of Queensland. The authors aim to address a potential national policy change for L2 English learners who might be recognised as an 'equity group', thereby requiring assessment and reporting measures to be put in place: Australia, unlike some English dominant countries, does not have federal or state laws enforcing provisions for L2 English learners in schools. The paper draws on McNamara and Ryan's (2011) notion of justice applied to school L2 proficiency assessment, encompassing the values implicit in the constructs and the social uses of the tool. In a context of a possible policy development requiring nationally consistent L2 English proficiency measurement, the paper expands on the hitherto very limited discussion in the academic literature of the concepts underlying L2 English proficiency tools used for school-aged populations in Australia. The authors outline principles and features of L2 English assessment tools that best serve fairness, justice and inclusion by providing for the pedagogical, administrative and policy visibility for the *full* cohort of L2 English learners. The paper takes as an exemplar the *Bandscales State Schools (Qld)*, which derives from the original National Languages and Literacy Institute of Australia (NLLIA) ESL Bandscales parent document, to illustrate the characteristics that are required and how this L2 English proficiency tool fulfils them, keeping in mind generalist classroom

teacher users, and complex student language backgrounds, and previously unserved and un-/under-recognised L2 English learner groups.

In the report “EAL/D or an Additional Need?”, Bernadette Barker discusses how young L2 English learners are commonly misidentified as having a learning difficulty, due to the widespread use of monolingual English assessment tools designed for monitoring reading and writing in schools. These tools range from standardised reading testing products to locally designed progress maps and frameworks and are a prominent and commonplace feature in Australian schools for showing learner improvement. L2 English learners who are assessed as not on par through such tools can be referred without delay to Speech Pathology or Psychological standardised testing, followed by unsuitable interventions. Based on the work of Hall et al. (2001), Barker reports on the development of a classroom-based, formative assessment tool for L2 English learners to address this equity issue. The tool consists of a Personalised Learning Plan focussing on a cycle of teaching for language growth, and which assesses L2 language development with an L2 proficiency tool. An essential part of the learning plan is the recording of the teaching and learning that occurs and the learner’s responses. The cycle therefore provides time for language teaching and language learning and a mechanism for monitoring both. It also provides comprehensive information about L2 learners not produced through standardised L1 English literacy and/or L1 special needs testing.

The paper by Carmel O’Shannessy and Aboriginal researchers, Vanessa Davis, Jessie Bartlett, Alice Nelson and Denise Foster, focusses on language assessment suited to young multilingual Aboriginal children and their families in the remote town of Alice Springs in the Northern Territory of Australia. Early childhood development is a current and increasing focus of Australian government policy and this drives the need for assessment to measure outcomes for children in this sector. For this purpose, the role of family languages in the lives of very young children cannot be under-estimated. But, as the authors note, little is known about children’s languages and their paths of multilingual development in central Australia, a situation complexified by additional variables of different configurations of languages across families, language change within language communities and language contact between language communities. The paper describes how some of the specific challenges for ascertaining young

multilingual Aboriginal children's language development in central Australia have been met through a multilingual vocabulary assessment tool, the Little Kids Word List (LKWL). This is a multilingual MacArthur Bates Communicative Development Inventory (CDI) app for four of the languages spoken by young children in Central Australia, Eastern and Central Arrernte, Western Arrarnta, Warlpiri and English as spoken locally by Aboriginal people. The paper provides a detailed discussion of the design processes and features of the LKWL, which is tailored to the linguistic and cultural context of these children and their families. The LKWL is an important initiative in the early childhood sector, as targets for child development measured by tools designed for L1 English speakers do not assess multilingual youngsters fairly. The LKWL will capture data that provides a strengths-based perspective of these children's language repertoires when they engage with preschool or other services.

### **Conclusion**

The contributions to this special issue make clear the contingent nature of fair assessment in school contexts. Fair assessment must be inclusive of the learning and achievements of young L1 speakers and/or L2 learners, especially those whose linguistic repertoires are liable to lie outside what is catered for by "mainstream" curriculum and assessment tools and processes. Young multilingual learners may not be visible as efficacious language learners in mainstream and standardised assessments. They may be dependent on teacher training that includes awareness of L1 or L2 assessment, or on a local development that includes their multilingual capacities in assessment. These young learners may need policy makers to engage with the language and assessment nexus, and in some contexts maybe community members too, so that they can be credited for their L1 or L2 language learning. There are many gaps in current language assessment toolkits for young language learners in early childhood, primary and high school settings. Including these excluded young language learners frequently depends on informed educators, community members, academics and researchers exercising agency to advocate for the need for fair language assessment, for more judicious use of existing assessment tools or for filling assessment gaps with purpose-built language assessment tools.

The authors in this special issue of SiLA describe and respond to a wide range of invisibilities and gaps in assessment for young multilingual learners from diverse language backgrounds. A recurring theme is the creative agency required of educators to add language assessment to the mainstream assessment structure where it has had limited leverage in the assessment power hierarchy (see Morita-Mullaney, 2017). The social justice and policy dimensions of language assessment for young L2 learners are clear and demonstrable. Our hope is that the papers in this issue will provide a source of understanding and motivation to inform and encourage ongoing advocacy by the language assessment community for fair assessment of young language learners in school and early childhood contexts.

## References

- Elder, C. (2021). Towards 'policy responsible' language assessment: Framing the Language Testing Research Centre's contribution in languages other than English. *Papers in Language Testing and Assessment*, 10(1), 4-29. <https://doi.org/10.58379/hqym2624>
- Hall, D., Griffiths, D, Haslam, L., & Wilkin, Y. (2001). *Assessing the needs of bilingual pupils: Living in two languages* (2<sup>nd</sup> ed.). David Fulton Publishers.
- Hill, K. (2017). Understanding classroom-based assessment practices: A precondition for teacher assessment literacy. *Papers in Language Testing and Assessment*, 6(1), 1-17. <https://doi.org/10.58379/yiwz4710>
- McNamara, T., & Ryan, K. (2011). Fairness versus justice in language testing: The place of English literacy in the Australian Citizenship Test. *Language Assessment Quarterly*, 8(2), 161-178. <https://doi.org/10.1080/15434303.2011.565438>
- Morita-Mullaney, P. (2017). Borrowing legitimacy as English learner (EL) leaders: Indiana's 14-year history with English language proficiency standards. *Language Testing*, 34(2), 241-270. <https://doi.org/10.1177/0265532216653430>
- Vogt, K., & Tsagari, D. (2022). Language assessment literacy and test validation: Highlighting the role of language teachers. *Studies in Language Assessment*, 11(1), 1-15. <https://doi.org/10.58379/uqvq7101>