

# Assessing English language proficiency

Centre for Education Statistics and Evaluation

A trial of the *English as an Additional Language or Dialect (EAL/D) Learning Progression instrument* in NSW schools has demonstrated strong reliability and validity for identifying students' English language proficiency.

The instrument is able to better identify and discriminate EAL/D needs for resource allocation at the school level, across NSW and potentially Australia-wide.

Teachers were able to make more discriminating judgements about each student than is possible with the current NSW ESL phase tool, and they found the new instrument easy to use.

The new school funding model under the National Plan for School Improvement includes a loading for students with limited English language proficiency. However, there is currently no nationally consistent measure of English language proficiency available as the basis for the distribution of this funding.

The NSW Department of Education and Communities (DEC) is now preparing a cost-benefit analysis for a national trial of the EAL/D instrument for the Standing Council on School Education and Early Childhood (SCSEEC).



*A more reliable and valid instrument for identifying English language proficiency for resource allocation.*

## What is the *EAL/D Learning Progression*?

The Australian Curriculum Assessment and Reporting Authority (ACARA) developed the *English as an Additional Language or Dialect (EAL/D) Learning Progression* to help teachers who are implementing the Australian curriculum. It was developed primarily for teachers who are not EAL/D specialists and who need to identify the English language levels of EAL/D students in their classrooms to address their specific learning requirements.

The *EAL/D Learning Progression* describes the development of English language typical of students learning English as an additional language or dialect. It includes broad descriptions of the characteristics of learner groups at each of four phases of learning (beginning, emerging, developing and consolidating) in each of four different language modes (listening, speaking, reading and writing). The phase descriptors for each mode are further differentiated by three broad stages of schooling: Years K-2, 3-6, 7-10.

## NSW and national contexts

In 2012 in NSW government schools, approximately 230,000 students were from language backgrounds other than English (LBOTE). These students made up around 30 per cent of total enrolments. More than 136,000 students (18 per cent) were learning English as an additional language or dialect.

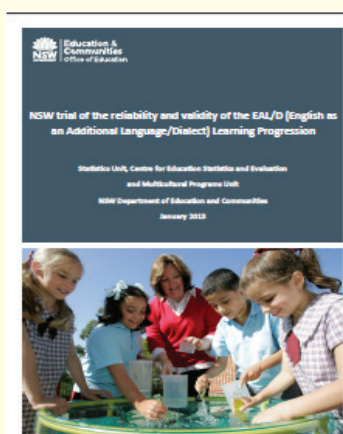
Many EAL/D students commence schooling in Australia with little or no proficiency in English, requiring targeted resources to overcome that disadvantage. In NSW, students' English proficiency is assessed with the three phase ESL tool. Data collected using this tool provides the basis for ESL support and funding to schools.

The NSW three phase tool classifies students in three levels, but lacks discrete descriptors that are age appropriate and specific to different aspects of proficiency. Teachers cannot identify different levels of a student's listening, speaking, reading, and writing skills in English, even though many students have varying proficiency across these modes. While another assessment tool – the ESL scales – allows for more refined assessments of English language skills, it requires significant ESL expertise and time to complete.

The new school funding model under the National Plan for School Improvement includes a loading for students with limited English language proficiency. In the absence of a nationally consistent measure of English language proficiency the new funding model uses a proxy measure, 'Disadvantaged LBOTE (language background other than English)', to distribute the funding to individual jurisdictions and schools. However, prior NSW analysis indicated that the 'Disadvantaged LBOTE' measure was a poor approximation for limited English proficiency. The NSW analysis highlighted the need for a more accurate and consistent measure for the equitable distribution of funding for limited English proficiency across jurisdictions.

The *EAL/D Learning Progression* was recognised as a possible alternative measure. Prior to the NSW trial however, there had been no large-scale trials to collect evidence and test its validity. In particular, it is important to know whether teachers can use the EAL/D instrument to produce sufficiently consistent assessments.

As a result of the trial initiative, DEC has been commissioned by a national working group to prepare a cost-benefit analysis for a possible national trial of the EAL/D instrument for resource allocation. The national trial would test the suitability of the EAL/D instrument across all jurisdictions, with a greater number of teachers and students from diverse cultural and linguistic backgrounds.



### The full report:

**NSW trial of the reliability and validity of the *EAL/D (English as an Additional Language/Dialect) Learning Progression* is available on the CESE website at:**

**[www.cese.nsw.gov.au/publications/research-reports](http://www.cese.nsw.gov.au/publications/research-reports)**

## Methodology

The NSW trial included 97 teachers, both specialist ESL teachers and classroom teachers with diverse ESL teaching and assessment backgrounds. Fifty-six schools were included – primary, secondary, central schools, and Intensive English Centres. Most were metropolitan, with a few in provincial centres.

A sample of 944 students was selected by the participating teachers in the trial. The sample was stratified by target grades (Kindergarten, Years 3, 5, 7 and 9), gender groups, sub-demographic groups (i.e., Aboriginal, international student, refugee), and represented the range of English proficiency levels using the NSW ESL phase assessment tool. A professional learning program prepared teachers for the trial.

Each student was assessed on four language modes – listening, speaking, reading, and writing. Teachers submitted assessment results of the four language modes for every student onto a purpose-built website.

Double-marking was used to investigate the reliability of assessments. Of the sample of 944 students, 639 were assessed by two teachers on every language mode using the EAL/D instrument. Teachers' assessments for each student were compared using various reliability statistics.

The trial focused on assessing the instrument's suitability for resource allocation in NSW government schools. Four aspects of validity were analysed – concurrent, discriminant, structural and measurement.

Following the trial, DEC surveyed teachers to assess the usefulness of the instrument and the quality of support provided and needed for the trial.

## Analysis and results

### Reliability

The trial results showed that the EAL/D instrument enabled teachers to make consistent judgements of English language proficiency across the four language modes (listening, speaking, reading, writing).

**Table 1** illustrates that the rates of exact agreement between two teachers assessing the same student were high, averaging 80 per cent. Even where two teachers' judgements differed, almost all differences were by only one level.

Statistical analysis showed that inter-rater reliability exceeded the required level for high stakes tests.

However, some variations in the consistency of teacher judgements were also observed. For example, teacher judgements were less consistent with some student cohorts (e.g., boys and ESL Phase 3 students) and in the speaking and listening modes. Teacher judgements were more consistent with

other groups of students, such as those at Intensive English Centres.

Factors that may explain the variations included:

- differences in teachers' prior experience and training in similar assessment tasks, in particular in assessing informal student interactions which dominate the speaking and listening mode
- teachers' knowledge of students
- time available to collect and assess work samples.

The trial did not find evidence that teachers with a direct teaching relationship with a student assessed the student differently from teachers who did not have a direct teaching relationship with the student. This suggests that the use of the instrument in a school may not need to be limited to a particular group of teachers, provided that teachers have expertise in using the instrument and follow the assessment guidelines in collecting work samples over time to inform phase judgements.

**Table 1 Rates of agreement between two teachers assessing the same student.**

Student Groups	Exact Agreement			
	Listening (%)	Speaking (%)	Reading (%)	Writing (%)
All Students	80.5	78.7	81.8	82.4
Girls	81.5	80.5	81.5	83.6
Boys	79.5	77.0	82.2	81.3
Aboriginal	100	83.3	88.9	100
ESL Phase 1	88.1	84.4	90.4	87.2
ESL Phase 2	78.7	75.1	81.5	80.3
ESL Phase 3	69.4	75.5	68.5	76.4
Kindergarten	84.4	77.1	86.3	84.4
Year 3	87.4	72.7	80.2	88.2
Year 5	85.2	80.0	80.9	83.5
Year 7	78.7	81.3	82.0	82.7
Year 9	71.5	80.3	80.9	75.9
Refugee Students	82.1	78.8	85.4	88.1
International Students	57.1	71.4	92.9	71.4
Intensive English Centre Students	89.7	85.0	96.3	88.8

Note: Total number of students included in the double-marking process: 639. Agreement rates are calculated based on teachers' judgements on four broad EAL/D phases for each mode.

## Validity

Validity of the EAL/D assessments is demonstrated by analysis of teachers' assessments in comparison to other related measures of the same students; analysis of whether teachers could effectively discriminate between language modes; and analysis of the consistency in the use of rating scales and in relation to existing theories.

### a) Alignment with related measures

*Teachers' judgements were analysed to see whether they aligned with other related measures of proficiency for the same students: ESL phase assessments and NAPLAN reading and writing results.*

### ESL phases

The trial found teachers' assessments using the EAL/D instrument and assessments based on another theoretically similar construct (the current NSW ESL phase tool) were generally consistent.

The ESL phase tool identifies three learner phases:

- Phase 1: basic English language proficiency – elementary functioning in a classroom
- Phase 2: transitional English language proficiency – partial and variable functioning
- Phase 3: fluent – competent functioning with some specific assistance needs.

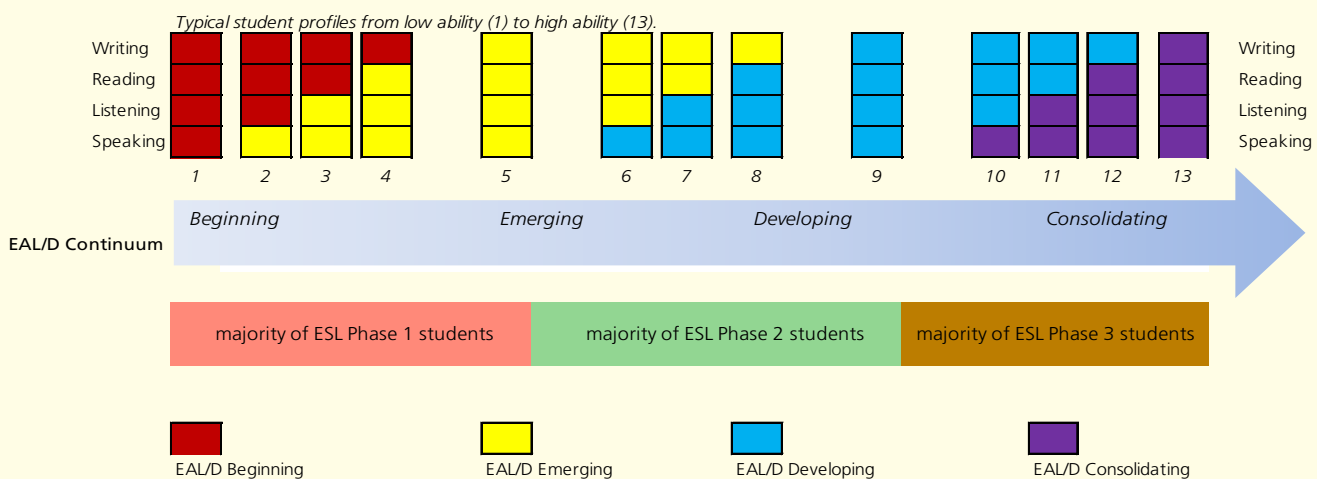
Figure 1 illustrates the typical developmental profiles of students assessed using the EAL/D instrument, and the overall alignment of these profiles with the students' ESL phase assessments (see technical note below Figure 1).

As expected, students in higher ESL phases are generally assessed to have greater proficiency across language modes on the EAL/D continuum than students in lower ESL phases. This indicates that the EAL/D instrument is working as intended.

Figure 1 also demonstrates that the EAL/D instrument allows for more refined and discriminating judgements between students than the current ESL phase tool. This is especially important for determining resources needed for students at the low end of the English language proficiency continuum.

Students assessed as Phase 1 with the ESL phase tool include a range of proficiencies according to the EAL/D instrument, from those assessed as beginning on all language modes through to those assessed as emerging on all language modes. The capacity of the EAL/D instrument to provide more precise judgements about student's abilities is beneficial to teachers in planning teaching and learning programs for students with low English language proficiency.

**Figure 1 Alignment of student profiles on the EAL/D continuum with the three ESL Phases**



Technical note: The sequence of EAL/D profiles in Figure 1 represents the expected developmental sequence across the four language modes. However, each profile actually represents all possible combinations of modes across the EAL/D phases shown (eg: Profile 3 with 'two beginning' and 'two emerging' represents any two language modes at the beginning phase with the other two modes at the emerging phase).

The alignment of students' ESL phases with the EAL/D profiles is based on the majority of students in each ESL phase. While there is strong evidence of alignment between the ESL phases and the EAL/D profiles at the overall level, there are some students

in each ESL phase who have a different profile to the majority. For clarity of interpretation, these students have been omitted in Figure 1.

The EAL/D continuum (represented by the blue arrow in the middle) is constructed from EAL/D mode assessments using the Rasch partial credit model (Wright & Masters, 1982).

The gaps between different profiles on the continuum reflect the relative differences in the underlying language ability from one profile to the next.

## NAPLAN

A similar analysis was also carried out to examine the relationship between results from the EAL/D instrument and NAPLAN.

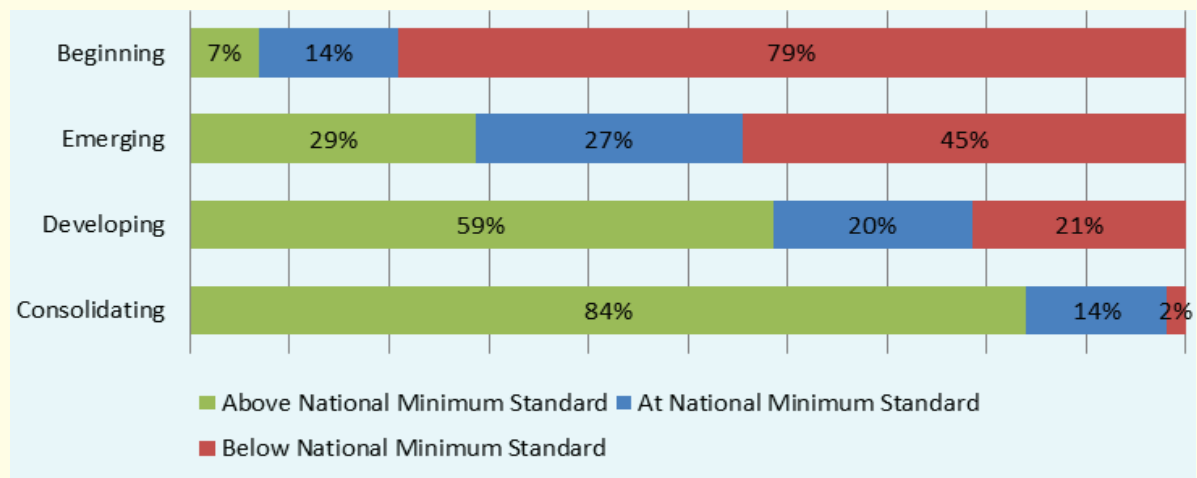
Results showed that higher language proficiency levels, as assessed by teachers using the EAL/D instrument, corresponded to higher levels of achievement in NAPLAN reading and writing tests for the same students assessed.

Figure 2 shows the proportion of matched Years 3, 5, 7, and 9 students at each EAL/D phase who are either below, at or above National Minimum Standards for reading.

While nearly 80 per cent of students assessed at the Beginning phase on the reading mode were below the National Minimum Standards in the NAPLAN reading tests, only 21 per cent of the students at the Developing phase, and 2 per cent of the students at the Consolidating phase were below the National Minimum Standards.

These findings show that the EAL/D instrument is highly consistent with related measures of student language abilities. This provides supporting evidence for the validity of the EAL/D instrument.

**Figure 2** Proportions of students at each EAL/D reading phase who were either below, at or above the National Minimum Standards in NAPLAN reading tests



Note:

Exempt students are included in the analysis as achieving below National Minimum Standards. Absent and withdrawn students are excluded from the analysis.

## b) Discrimination

*Teachers' judgements were analysed to see whether there is evidence that they could use the EAL/D instrument to effectively discriminate between the language modes.*

Table 2 shows that teachers do effectively discriminate using the instrument. The table shows each diagonal value is higher than all other values lying in its row or column. For example, Teacher 2's judgements on the listening mode correlate with Teacher 1's judgements on the same mode for the same students at 0.85. This correlation is higher than the correlations with Teacher 1's judgements on any other mode, for the same students (0.75, 0.70 and 0.68 for reading, speaking and writing, respectively).

This means that the different modes of the EAL/D instrument do measure different aspects of language proficiency and that teachers can identify the phase for one mode without judgement being clouded by student performance on another mode.

## c) Measurement and Structural Validity

Analysis of the trial results demonstrates that the EAL/D instrument has the characteristics of a high quality assessment tool:

- The four language modes (speaking, listening, reading and writing) all measure aspects of an underlying overall language proficiency, which means the instrument can be used to create a single English language proficiency rating for each student.
- The rating scales used to assess the four language modes work well and are used meaningfully and consistently by teachers.
- The mode assessments made by teachers using the instrument are consistent with theories of second language acquisition. For example, teachers judged that some EAL/D students developed reading and writing proficiency at different rates than listening and speaking. These findings are consistent with results from other ESL research studies.

**Table 2 Correlations between two teachers' judgements, for the same students, across language modes**

		Teacher 2 judgement				
		MODE	LISTENING	READING	SPEAKING	WRITING
Teacher 1 judgement	LISTENING		.85**	.69**	.69**	.68**
	READING		.75**	.87**	.70**	.75**
	SPEAKING		.70**	.70**	.84**	.67**
	WRITING		.68**	.75**	.67**	.87**

Note: \*\*Correlation is significant at the 0.01 level (2-tailed).

## Teacher feedback

The survey indicated support from teachers for use of the EAL/D instrument to plan teaching and learning. There was also considerable interest in using it for resource allocation instead of the current ESL phase tool.

[Note: use of the ESL scales which measure 28 levels across three strands of proficiency will continue in NSW as a diagnostic tool for assessing support needs of individual students.]

Figure 3 shows that the majority of teachers reported the EAL/D instrument was useful, that they felt confident to use it and that it should replace the current ESL phase tool for resource allocation purposes.

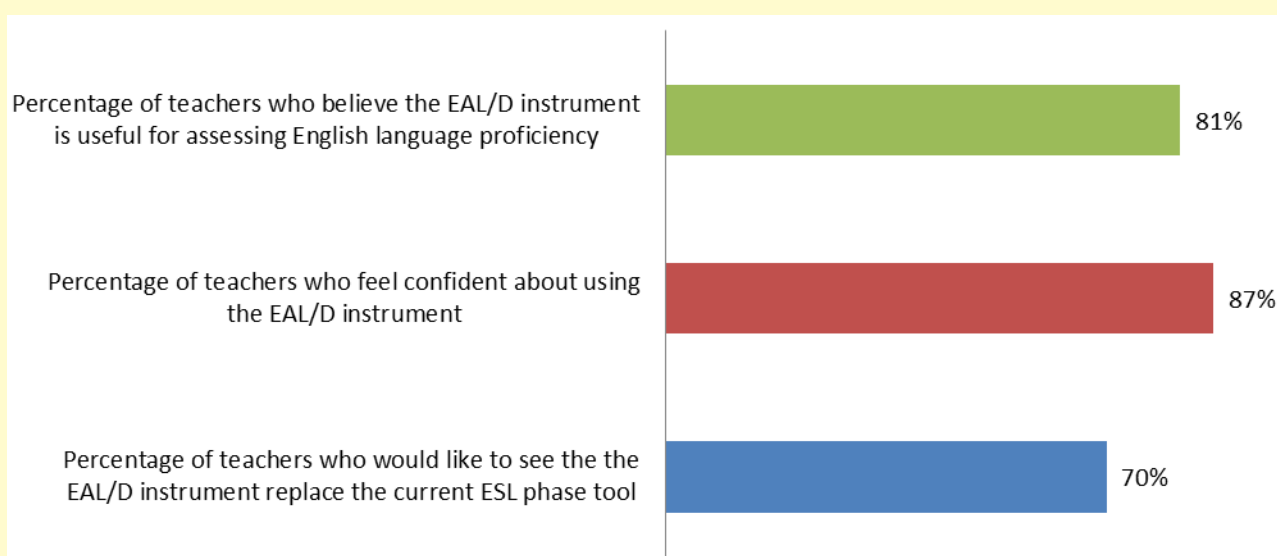
The EAL/D instrument was seen as the more favourable tool because it shows the impact of student age and stage of schooling on phase characteristics. It does not link English language proficiency to length of time learning English. It is

more detailed than the ESL phase tool and therefore supports more accurate teacher judgement.

The survey indicated that professional learning provided in the trial was adequate for confident and consistent judgements using the EAL/D instrument.

Teachers were most positive about the face-to-face learning opportunities and the annotated work samples and other teaching resources provided in the trial. However, teachers also indicated they required more time to plan and assess students to make adequately informed phase judgements.

**Figure 3 Teacher feedback on the use of the EAL/D Learning Progression**





## Next steps

### NSW resource allocation

From 2015, use of the *EAL/D Learning Progression* will replace the ESL Phase tool to allocate resource for ESL programs in NSW.

Introduction of the *EAL/D Learning Progression* for 2014 will enable data to be collected and used in the 2015 resource allocation process.

The implementation plan includes:

- a program of professional learning
- resources to support teacher assessment
- guidelines for using the EAL/D instrument
- capacity in the new student administration and learning management system (SALM) to allow teachers to enter EAL/D assessments.

Professional learning is being provided for teachers in 2013 in how to use the *EAL/D Learning Progression* to prepare for implementation in 2014.

### NSW classroom support

The *EAL/D Learning Progression* will be mapped against the new NSW syllabuses and the literacy continuum. On the basis of this mapping, teaching resources will be developed where appropriate to support classroom teachers to program and plan for EAL/D learners.

### National trial of the *EAL/D Learning Progression*

The NSW trial recommended a national trial to test how well the evidence from the NSW trial can be generalised, including:

- a broader range of student and teacher demographic groups, including Aboriginal students
- examination of differential reliability in teachers' judgements across different types of schools, students and language modes
- use of the EAL/D instrument to report English language proficiency across jurisdictions.

NSW DEC is preparing a cost-benefit analysis of a prospective national trial for the Standing Council on School Education and Early Childhood (SCSEEC).

## Key References

- Australian Curriculum, Assessment and Reporting Authority (2012). English as an Additional Language or Dialect Teacher Resource — Overview and *EAL/D Learning Progression.*, ACARA, Sydney.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed.), (pp.17–64). Westport, CT: American Council on Education.
- Messick, S. (1996). Validity of performance assessments. In G. Phillips (Ed.), *Technical issues in large-scale performance assessment* (pp.1–18). Washington, DC: National Centre for Educational Statistics.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.

For a full list of references see the full report at:

<http://www.cese.nsw.gov.au/publications/research-reports>

Centre for Education Statistics and Evaluation  
Office of Education  
GPO Box 33  
Sydney NSW 2001  
Australia  
T 02 9561 1220  
E [cese@det.nsw.edu.au](mailto:cese@det.nsw.edu.au)  
[www.cese.nsw.gov.au](http://www.cese.nsw.gov.au)

NSW Department of Education and Communities

