NSW Department of Education

Revised methodology for the English language proficiency (ELP) funding model

Centre for Education Statistics and Evaluation

October 2021

NSW GOVERNMENT

Centre for Education Statistics and Evaluation

The Centre for Education Statistics and Evaluation (CESE) undertakes in-depth analysis of education programs and outcomes across early childhood, school, training, and higher education to inform whole-of-government, evidence-based decision making. Put simply, it seeks to find out what works best. Its focus is all education in NSW. CESE's main responsibilities within the department are:

- qualitative and quantitative research, including data analysis, evidence papers and case studies that build understanding and uptake of evidence-based practice
- robust data collection to enable research and statistics for the education and training sector
- evaluating key policies and programs to strengthen quality delivery and student outcomes
- national engagement on research agenda and data strategy
- driving capability uplift in the use of data and evidence as part of everyday practice
- trialling innovative initiatives to improve student outcomes.

Authors

Lucy Lu, Wai-Yin Wan and Stephen Hao in collaboration with the Multicultural Education team
Centre for Education Statistics and Evaluation, October 2021, Sydney, NSW

Please cite this publication as:
Centre for Education Statistics and Evaluation (2021), **Revised methodology for the English language proficiency (ELP) funding model**, NSW Department of Education, education.nsw.gov.au/cese

For more information about this report, please contact:

Centre for Education Statistics and Evaluation
Department of Education
GPO Box 33
Sydney NSW 2001

info@cese.nsw.gov.au

+61 2 7814 1527

education.nsw.gov.au/cese

We acknowledge the homelands of all Aboriginal people and pay our respect to Country.

Table of contents

Executive summary 5

Introduction 8

Current methodology 9

- English language proficiency phase 9
- Refugee status 10
- Length of time in Australian schools 10

Why do we need to review the ELP funding model? 12

Revised methodology 15

- Risk factor 1 – at or below NMS in NAPLAN reading and/or writing 15
- Risk factor 2 – not progressing to the Consolidating phase or higher 17
- Combining 2 risk factors to inform weightings 18

Conclusion 24

Appendix 1: Relevant research in 2014 that underpinned the current ELP loading 25

Appendix 2: Technical details of multilevel modelling to predict the likelihood of falling at or below NMS in NAPLAN reading and/or writing tests for EAL/D students 30

Appendix 3: Technical details of accelerated failure time model with shared frailty to estimate the probability of not progressing to the Consolidating phase or higher for EAL/D students 34

References 36

List of figures

Figure 1
Illustration of funding calculation under the current ELP funding model 11

Figure 2
Average LoT in Australian schools and LoT weighting across scholastic grades 13

Figure 3
Base weighting by phase and scholastic years 20

Figure 4
Steps of funding calculation under new methodology 23

Figure A1
Predicted probability of performing at or below NMS 26

Figure A2
Predicted probability for achieving below NMS – EAL/D primary and secondary students 27

Figure A3
Effect sizes of various factors driving NAPLAN performance – primary and secondary students separately 28

List of tables

Table 1
Base weighting for phase and scholastic year combinations 19

Table 2
Refugee status weighting 21

Table 3
Length of time weighting 22

Table A1
Regression coefficients of the multilevel logistic regression model for the primary year cohort 31

Table A2
Regression coefficients of the multilevel logistic regression model for the secondary year cohort 32

Executive summary

Students who are learning English as an additional language or dialect (EAL/D) need support to develop proficiency in academic and conversational English in order to access the school curriculum and engage effectively in learning. The equity loading for English language proficiency (ELP) provides schools with a teaching allocation (full-time equivalent) and/or flexible funding to provide specialist English language support to meet the needs of these students.

EAL/D students are a diverse cohort. They may enrol in school at any time and in any scholastic year. They may be Australian or overseas born and have varying levels of proficiency in English, from very limited to more developed English language skills. On average, it takes EAL/D students 5 to 7 years with specialist EAL/D teaching support to master the academic English language required for success at school. EAL/D students from refugee backgrounds who have experienced trauma and interrupted or limited prior schooling may take up to 11 years to do so. For this reason, and in recognition of the importance of the development of academic English language proficiency in the achievement of successful schooling outcomes, ELP equity loading resources are available to support students as long as they are assessed as learning English at any one of the 4 phases of English language proficiency.

Generally speaking, the lower the English language level proficiency of an EAL/D student and the older they are at the time of enrolment in an Australian school, the greater the risk they will not develop sufficient English proficiency to succeed at school. This is due to the length of time it takes for students to acquire academic English and the increasing English language demands of the curriculum as students move through scholastic years and stages of learning into more specialised courses of study. Students from refugee backgrounds, in particular those with low levels of English language proficiency and who first enrol in the higher scholastic years, are at greater risk of not developing sufficient academic English by the time they complete their schooling. This may be due to previous experiences such as trauma and interrupted schooling.

The current methodology used to allocate ELP is weighted in line with EAL/D students' needs. The current model takes into account each student's English language proficiency level, length of time (LoT) in an Australian school and refugee status, and aggregates this data across all students in the school to derive the total funding amount for each school. The English language proficiency levels used to determine funding are those assessed by teachers using the national EAL/D learning progression – Beginning (lowest language proficiency phase); Emerging; Developing; or Consolidating (highest language proficiency phase) – and are reported through the annual EAL/D census. The 4 phases in the EAL/D learning progression represent broad levels of ELP used for the purposes of system identification and reporting of EAL/D learners.

The current methodology was developed in 2014. It applies an 8:4:2:1 weighting for students at the Beginning, Emerging, Developing and Consolidating phases respectively, based on educational disadvantage determined through an analysis of 2014 NAPLAN results. It also applies higher weightings for refugee students at 1.3:1 (refugee to non-refugee students) and moderates needs by a student's LoT calculated as 1 divided by the number of years a student is enrolled in Australian schools. The LoT moderation is applied multiplicatively with ELP phase and refugee status weightings to derive the final weighting for each student.

This methodology was reviewed over 2019 and 2020 to ensure that the ELP was being allocated as equitably as possible to best meet the needs of EAL/D students at each ELP phase and stage of learning. The review considered additional data to test if the patterns of relative disadvantage across ELP phases still hold true as well as multiple years of ELP phase data across 4 language modes (listening, speaking, reading and writing) to estimate the rates of students' English language acquisition and relative educational disadvantage. It also analysed evidence to suggest that the current LoT moderation was adversely impacting on the equitable distribution of funding across scholastic years and phases.

The review confirmed that, in general, more recent arrivals had a greater level of educational disadvantage compared with students at the same grade, same ELP phase and same refugee status who had been in Australian schools for longer. However, the review found no evidence that supported either the **magnitude** of the current LoT moderation or **how** it is applied. It found that the current LoT moderation does not consider the positive relationship between scholastic grade and LoT resulting in a steep decline in weighting from primary to secondary years. This means that students in higher grades are currently receiving much lower levels of funding than those in lower grades even if they have been assessed at the same ELP phase and come from similar backgrounds.

The steepness in the decline of LoT weighting from Kindergarten and Year 1 to Year 2, and beyond, is currently resulting in about half of the total ELP funding being allocated to support EAL/D students in Kindergarten and Year 1 even though they only comprise about 25% of the total EAL/D student population. In addition, since Beginning students on average are enrolled in Australian schools for fewer years than their counterparts, EAL/D students at the Beginning phase are currently receiving about 40% of the total ELP funding despite comprising only 10% of the total EAL/D population.

In order to more accurately estimate the educational risks of EAL/D students, the review confirmed the need for:

- an updated evidence base for the relativity of funding across the 4 ELP phases using the past 6 years of data
- consideration of a new risk factor on the rate of progress for different types of EAL/D students
- the collection of evidence for a new way of moderation for more recent arrivals.

The review proposes new methodology based on available data which takes into account 2 risk factors – the risk of an EAL/D student achieving at or below National Minimum Standard (NMS) in NAPLAN reading or writing tests; and the risk of not progressing to or beyond the Consolidating phase (ELP phase).

Using this evidence base, new and more nuanced weightings have been developed for 3 key components of the ELP funding model:

- ELP phase by scholastic year
- refugee status
- length of time enrolled in Australian schools relative to that expected of the student's grade.

These weightings aim to deliver more equitable and transparent distribution of ELP funding across schools and different EAL/D student groups.

It should be noted that the review also revealed that more investigation into the best model of support for EAL/D students in schools for specific purposes (SSPs) was required. Further investigation into the best ways to support EAL/D students with special needs is being undertaken.

Implications for schools

The results of this review may also be used to inform school-based decisions about the targeted use of funds to support EAL/D students.

Using the findings of this review, schools may decide to prioritise resources for EAL/D students who are at the highest risk of not meeting NAPLAN minimum standards and not progressing to the Consolidating phase of English language proficiency. This review has identified that in general, these students were those who were in the early phases of English language proficiency who enrolled in an Australian school later than their age-equivalent cohort. Students with the highest risk were newly arrived older students in the higher grades. In addition, these risks were amplified for students from refugee backgrounds.

Introduction

The equity loading for English language proficiency provides schools with teaching allocation (full time equivalent) and/or flexible funding to support the needs of EAL/D students. EAL/D students may enrol in school at any time and in any scholastic year. They may be Australian or overseas born and have varying levels of proficiency in English, from very limited to more developed English language skills. These students need specialist EAL/D teacher support to acquire sufficient proficiency in academic and conversational English in order to access the school curriculum and engage in learning. On average, it takes EAL/D students 5 to 7 years with specialist English language support to master the academic English language required for success at school and may take up to 11 years for students from refugee backgrounds who have experienced trauma and interrupted or no prior schooling (Cummins 1984; Garcia 2000).

The following section describes the current model used to determine English language proficiency funding for each school.

Current methodology

The current model has 3 components that together determine the ELP equity funding for each school:

1. student English language proficiency phase
2. refugee status
3. length of time in an Australian school.

Funding is calculated for each student first based on student data for the 3 factors above and then aggregated across all students in a school to derive the funding amount for each school.

English language proficiency phase

The ELP phases used to determine funding are those as assessed by teachers using the national EAL/D learning progression and collected through the annual EAL/D census. The EAL/D census commenced from 2014. The learning progression, developed by the Australian Curriculum, Assessment and Reporting Authority (ACARA), describes English language proficiency in 4 ordered phases:[1]

1. Beginning (lowest language proficiency phase)
2. Emerging
3. Developing
4. Consolidating (highest language proficiency phase).

The level of funding associated with each phase is determined based on statistical analysis performed by the Centre for Education Statistics and Evaluation (CESE) in 2014,[2] after the trial of the EAL/D learning progression in NSW Government schools. At the time, the only academic results available centrally to estimate educational disadvantage for EAL/D students were 2014 Year 3, 5, 7 and 9 NAPLAN results. Guided by a departmental working group, CESE examined the relationship between students' ELP phases (and other student demographic background factors such as socio-educational advantage (SEA), Aboriginal and Torres Strait Islander status) and their NAPLAN reading results as a way of estimating disadvantage for students at different ELP phases. Whether students performed at or below National Minimum Standard was used as the outcomes measure of interest, as students performing at or below NMS were considered to be at risk of being unable to progress satisfactorily at school without targeted intervention.

1 For more information on EAL/D learning progressions as well as the EAL/D census, refer to the Department of Education EAL/D advice for schools [PDF 5.4MB] or the ACARA EAL/D learning progression [PDF 317KB].

2 Refer to CESE 2014 analysis described in Appendix 1.

The 2014 analysis shows that the ratio of the probability of students achieving at or below NMS was 8:4:2:1 for students at the Beginning, Emerging, Developing and Consolidating phases respectively. These ratios were taken as indicative relativities in the level of support required for students at different phases to acquire English language proficiency to access school curriculum, and therefore used as a basis of the funding allocation.

Separate models were also run for the primary (Years 3 and 5) and secondary (Years 7 and 9) cohorts to test whether primary and secondary students, with the same assessed ELP phases, had different levels of educational disadvantage, given secondary students may experience a higher level of demand on their English skills as they access a more complex curriculum. Analysis confirmed that secondary students had a higher probability of performing at or below NMS than primary students assessed at the same relative proficiency level and of the same demographic backgrounds. The ratio of the probability of secondary students performing at or below NMS versus primary students, across all EAL/D phases, was used as a basis to calculate the learning needs of (and hence the funding required for) secondary students relative to primary students.

Refugee status

Another factor considered in the ELP funding model is the refugee status because studies have presented consistent evidence that students who were from a refugee background generally required more support to acquire English language proficiency, due to disruption to previous schooling and experience of trauma. The current weighting for students with an active or historical refugee visa is 1.3 and for those without ('non-refugee' students) is 1. This means refugee students attract funding 1.3 times higher than non-refugee students, everything else being equal.

It is noted that the current weighting was based on 2014 data only, so new analysis was needed to examine the relative disadvantage of this group of students now that we have 5 more years' worth of data.

Length of time in Australian schools

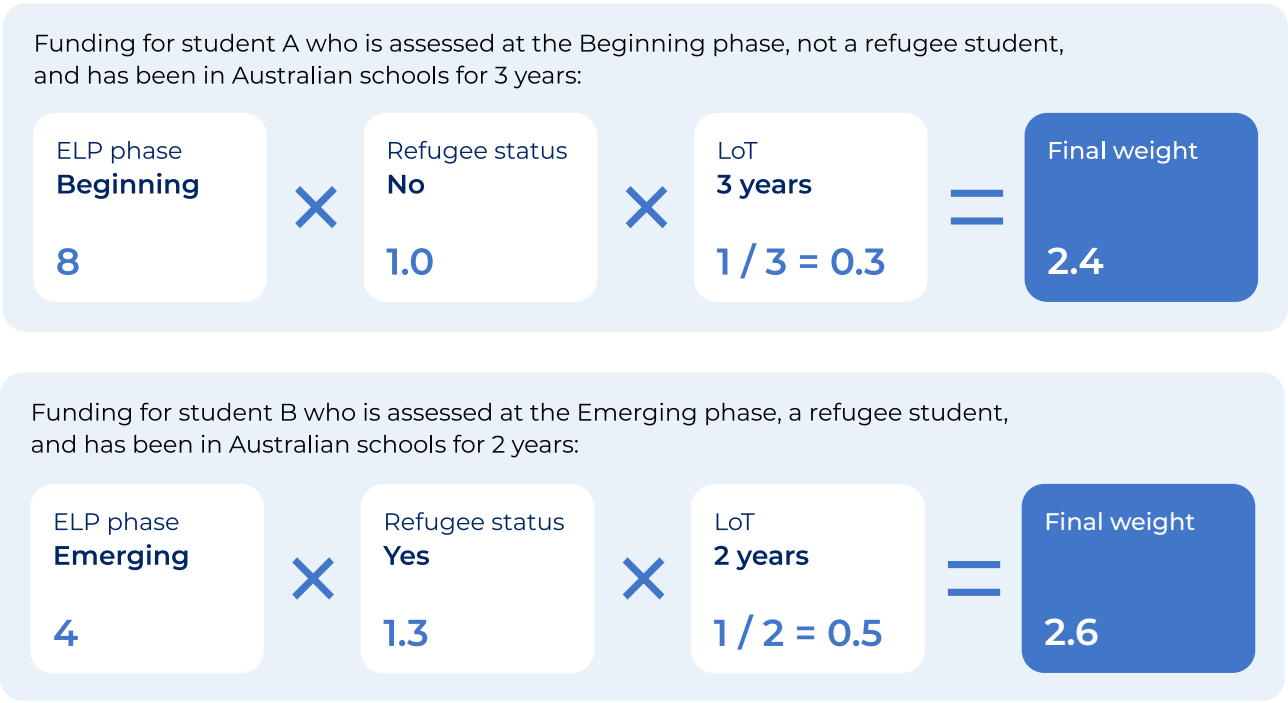
The third factor considered in the ELP funding model is the length of time enrolled in Australian schools for each EAL/D student. This factor is used to moderate the initial funding level calculated based on student's ELP phase and refugee status. The need for moderation was based on earlier analysis that, everything else being equal (such as ELP phase, scholastic year), students who were newer to the Australian education system (that is, enrolled for less time in Australian schools) had a slightly higher level of educational disadvantage and needed greater support to adjust to the teaching and learning in an Australian school.[3] The LoT factor is simply calculated as 1 divided by the number of years a student is enrolled in Australian schools and applied multiplicatively with weightings from ELP phases and refugee status to derive the final weighting for each student.

3 Refer to CESE 2014 analysis described in Appendix 1 (Figure A3 on page 28).

Figure 1 is a diagram to illustrate the basic steps involved in calculating the funding under the current methodology.[4]

Figure 1

Illustration of funding calculation under the current ELP funding model



Appendix 1 provides more details of the 2014 analysis that underpinned the current ELP funding model.

4 Note that for international students their ELP phases were adjusted such that students in the Beginning and Emerging phases were modified as Developing phase. This adjustment aimed to allocate less funding to international students as schools have already received other sources of financial support for these students.

Why do we need to review the ELP funding model?

There are 3 main factors that have led to the review of the current funding model.

Firstly, the current model was based on 2014 ELP and NAPLAN results data. There is a need to review and update the evidence base by using data accumulated over the last 6 years to test if the patterns of relative disadvantage across phases estimated using the NAPLAN tests still hold true.

Secondly, we now have multiple years of ELP phase data for the same students. This means we can estimate the rates of progress of students on the English language proficiency developmental scale and use this information to improve the estimation of the needs of different types of EAL/D students. Teacher judgements of students' English language proficiency phases, which are holistic assessments across 4 language modes (listening, speaking, reading and writing), provide a rich data source to estimate educational disadvantage. The inclusion of this new information, operationalised as an additional risk factor in the new funding model, will also reduce the model reliance on the NAPLAN tests results, which are narrower constructs as they do not assess listening and speaking aspects of language proficiency.

Thirdly, there is evidence to suggest that the current length of time moderation is adversely impacting on the equitable distribution of funding across scholastic years and phases. CESE's analysis in 2014 showed that, in general, more recent arrivals had a greater level of educational disadvantage than those who were of the same backgrounds (for example same grade, same ELP phase and same refugee status) but had been in Australian schools for longer. However, this review has found no evidence that supports either the **magnitude** of the LoT moderation or **how** it is currently applied.

A fundamental issue with the current LoT moderation method is that it does not consider the positive relationship between scholastic grade and length of time in Australian schools. This leads to the ELP funding being confounded by the grade a student is in, rather than based purely on needs.

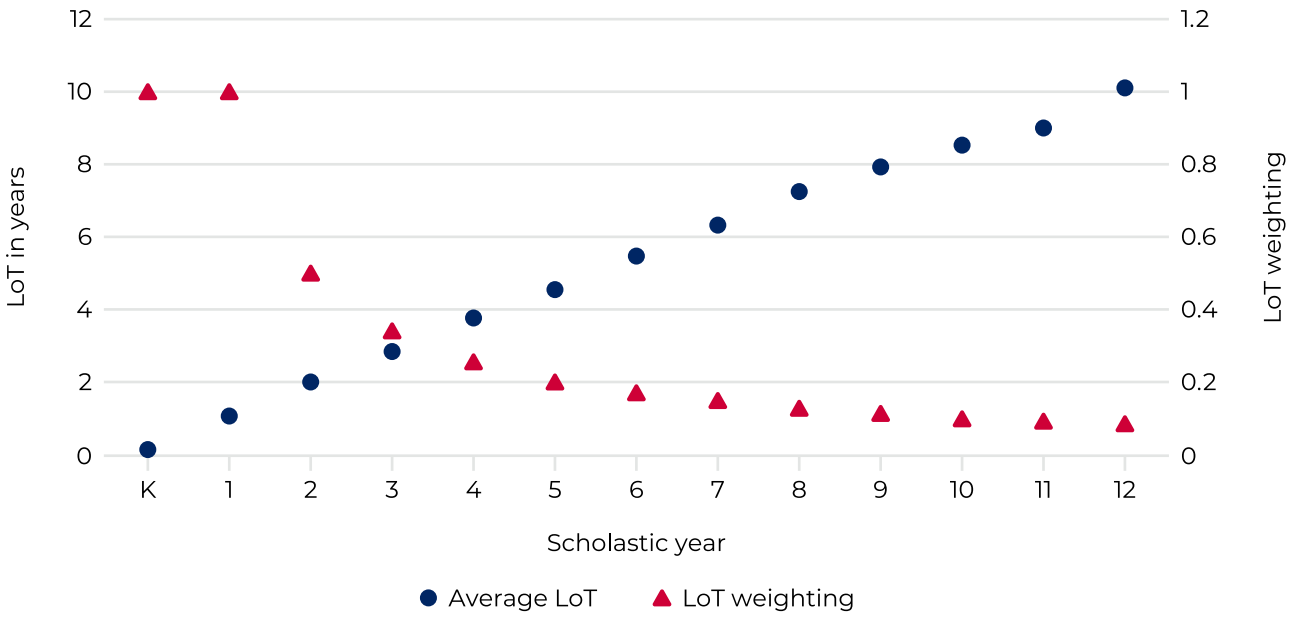
Figure 2 illustrates this point by showing the average length of time in Australian schools for each scholastic year, based on data extracted from the enrolment system for the 2019 EAL/D cohort. The blue dots demonstrate that, on average, EAL/D students of higher grades have been enrolled in Australian schools for longer than those of lower grades, which is as expected.

As mentioned above, the weighting for the LoT factor is simply calculated as 1 / LoT (in years). For example, for a typical Year 2 EAL/D student, the LoT (in years) is 2 years, therefore the LoT weighting is 1 / 2 = 0.5. For a typical Year 4 EAL/D student, the LoT is 4 years, therefore the LoT weighting is 1 / 4 = 0.25, which is half of that assigned to a typical Year 2 EAL/D student.

These LoT weightings, calculated for a typical student in each grade (corresponding to the average LoT for that grade), are displayed as red triangles on the graph.

Figure 2

Average LoT in Australian schools and LoT weighting across scholastic grades



Source. ERN data for 2019

The graph shows the steep decline in weighting from primary to secondary. The flow-on funding implication is that higher grades students would receive much lower funding than lower grades students, even if they may be assessed at the same language proficiency phase and of similar backgrounds (for example, refugee status) so the needs may be similar. For example, on average, a Year 4 student at the Emerging phase would receive only half of the funding allocated to a Year 2 student at the same ELP phase even if their ELP needs may be similar, since the corresponding LoT weightings are 0.5 versus 0.25. Similarly, a typical Year 10 Beginning phase student would receive funding about 30% less than a typical Year 7 Beginning phase student because the corresponding LoT weightings are 0.1 versus 0.14, even if an argument can be made that the Year 10 student needs more support to develop language proficiency more quickly to access a more complex curriculum.

While there was some evidence to justify newer arrivals receiving more funding than students who have been in Australian schools longer, the current method of moderation is too blunt, confounded by an external factor (grades) not related to needs, and the magnitude of moderation is not supported by any existing evidence.

The steepness in the decline of LoT weighting from Kindergarten and Year 1 to Year 2 and beyond results in Kindergarten and Year 1 EAL/D students receiving 50% of total ELP funding even if they only comprise 25% of the total EAL/D population. It also contributes to the pattern that Beginning phase students receive 40% of total funding despite them comprising only 10% of the total EAL/D population, since Beginning students on average are enrolled in Australian schools for fewer years than their counterparts.

In summary, the review of the current funding model has confirmed the need to:

- update the evidence base for the relativity of funding across the 4 ELP phases using the past 6 years of data
- add a new risk factor that considers the rate of progress for different types of EAL/D students to more accurately estimate educational risks
- collect evidence to inform a new way of moderation for more recent arrivals.

Revised methodology

With the availability of the annual EAL/D phase data and NAPLAN data over the last 6 years, the new methodology considers 2 risk factors for each student given their ELP phases and background:

1. risk of at or below NMS in NAPLAN reading or writing tests (risk factor 1)
2. risk of not progressing to the Consolidating phase (risk factor 2).

The 2 risk factors for each student were estimated using 2 different statistical models and are both expressed as probabilities so that they can be combined and placed on the same scale in order for us to calculate a composite risk index for each student.

Both models consider a full range of factors that might influence students' academic literacy achievement and/or rate of progress on the EAL/D learning progression, to estimate the unique effect associated with each factor. Factors included in both models include student Aboriginal and Torres Strait Islander status, refugee status, SEA quartile group, scholastic year, English language proficiency phase, student type (mainstream or students in support classes) and a new 'length of time enrolled in Australian schools' measure.

The new length of time measure (referred to as 'LoT relative to that expected of the student's grade') is constructed in a way to remove the apparent association between grade and the number of years enrolled in Australian schools mentioned in the preceding section. The new measure is of a relative nature as it is calculated as the number of years enrolled in Australian schools minus the typical number of years expected given the grade a student is enrolled in.[5]

Risk factor 1 – at or below NMS in NAPLAN reading and/or writing

In the current methodology, the statistical model applied to calculate the phase weighting did not use any information from the NAPLAN writing test, due to the reliability issue associated with 2014 writing results. The new methodology considers NAPLAN writing as well as reading results because:

1. Writing is an important aspect of English language proficiency.
2. We now have 6 years of writing results to help derive a more robust estimate of educational risks.

The second improvement made to the existing model is to take into account the hierarchical nature of the NAPLAN data where students are nested within school. If the hierarchical nature of the data is not accounted for, the weightings derived from the modelling for each factor would be biased. In short, in the new model, 6 years of NAPLAN and ELP phase data (2014 to 2019) were used to estimate this particular risk, to ensure the model parameters were not affected by any aberrant trend in any particular calendar years.

5 For students who enrolled before July, the number of years enrolled in Australian schools is calculated as the current calendar year minus the year the student first enrolled in an Australian school, otherwise the number of years enrolled is equal to current calendar year minus the year the student first enrolled in an Australian school minus 1. The expected number of years enrolled depends on the scholastic year (grade). For Kindergarten students, the expected number of years is 0; for Year 1 students, the expected number of years is 1 and so on. For example, Year 3 students are expected to have enrolled in Australian schools for 3 years. Year 3 2019 students enrolled in Kindergarten in Term 1 2016 had a relative length of time equal to 0, whereas Year 3 2019 students first enrolled in Term 1 2017 had a relative length of time equal to −1.

The statistical model implemented is known as the 'multilevel logistic regression model'. It takes into account the nesting structure of the data where students' outcomes within school were correlated. Separate models were fitted for the primary (Year 3 and Year 5) and secondary (Year 7 and Year 9) cohorts. In the multilevel logistic regression model, the binary outcome variable is defined as whether or not an EAL/D student is performing at or below NMS in NAPLAN reading and/or writing tests. Controlling variables included in the model are:

- Aboriginal and Torres Strait Islander status
- refugee status
- SEA status
- scholastic year
- EAL/D learning progression phase
- number of years spent in the current EAL/D phase
- whether enrolled in Australian schools for less than a year
- length of time enrolled in Australian schools relative to that expected of the student's grade and student type.[6]

The mean-variance adaptive Gauss-Hermite quadrature estimation method was used to estimate the regression coefficients associated with each controlling variable. Using the coefficient estimates and given the controlling variables, the probability of performing at or below NMS in NAPLAN reading and/or writing tests was calculated for each EAL/D student, at a particular phase and of a particular background (for example refugee status, length of time in an Australian school, scholastic year).

Full technical details of the multilevel logistic regression model and associated results are provided in Appendix 2.

6 Length of time enrolled in Australian schools relative to that expected of the student's grade' is calculated as the number of years enrolled in Australian schools minus the expected number of years enrolled in Australian schools for students in the same grade. For students enrolled before July, the number of years enrolled in Australian schools is calculated as the current calendar year minus the year the student first enrolled in an Australian school, otherwise the number of years enrolled is equal to current calendar year minus the year the student first enrolled in an Australian school minus 1. The expected number of years enrolled depends on the scholastic year (grade). For Kindergarten students, the expected number of years is 0; for Year 1 students, the expected number of years is 1 and so on. For example, Year 3 students are expected to have enrolled in Australian schools for 3 years. Year 3 2019 students enrolled in Kindergarten in Term 1 2016 had a relative length of time equal to 0, whereas Year 3 2019 students first enrolled in Term 1 2017 had a relative length of time equal to −1.

Risk factor 2 – not progressing to the Consolidating phase or higher

Since 2014, the department has collected teacher judgements of EAL/D students' English language proficiency phases using the EAL/D learning progression tool in June every year through an EAL/D census. The EAL/D learning progression tool allows teachers to assess their students' English language proficiency in one of 4 phases – Beginning, Emerging, Developing or Consolidating. EAL/D students at different starting phases and of different backgrounds progress along the learning progression at different rates. Students who take longer to progress to the Consolidating phase or higher can be perceived as at a higher risk of being unable to develop English language proficiency requisite for schooling and beyond.

The availability of the census data from 2014 to 2019 has allowed us to follow a student's progress on the EAL/D learning progression for up to 6 years. The accelerated failure time (AFT) model with shared frailty was used to estimate the probability (risk) of EAL/D students not progressing to the Consolidating phase on the learning progression given the student-level and school-level background and characteristics. In this model, the outcome variable is the time taken for an EAL/D student to progress to the Consolidating phase or higher (event of interest).[7] The AFT model specifies a direct linear relationship between the log of event time and contextual factors. These factors included:

- Aboriginal and Torres Strait Islander status
- gender
- refugee status
- New Arrivals Program status
- student socio-educational advantage
- scholastic year
- lowest EAL/D phase
- student type (mainstream versus additional learning needs)
- whether previously or currently enrolled in Intensive English Centres (IECs)
- the length of time enrolled in Australian schools relative to that expected of the student's grade
- school-level mobility rate (that is, rate of student movements in a given year within a school).

The shared frailty incorporated in the AFT model accounts for the lack of independence among students in the same school. In other words, the progress of students along the EAL/D learning progression are more similar within school in clustered data.

Maximum likelihood estimation method was used to estimate the coefficients associated with each contextual factor and other scale parameters. Using the model parameter estimates and given the contextual factors, the probability of not progressing to the Consolidating phase or higher was calculated for each EAL/D student.

Full technical details of the AFT model with shared frailty and associated modelling results are provided in Appendix 3.

7 Some EAL/D students may skip the Consolidating phase and progress to 'Not required' status. For example, some students progress from the Developing phase in the previous census to 'Not required' status after one year. Therefore 'Not required' needs to be considered as an event so that the risk of not progressing to Consolidating can be accurately estimated.

Combining 2 risk factors to inform weightings

The current methodology only uses the probability of performing at or below NMS in NAPLAN reading tests to inform the phase weighting. Using additional EAL/D phase history data, the revised methodology estimates 2 distinct risk factors and then combines them as mentioned above to inform the weightings for the 3 key components of the new ELP funding model:

- phase by scholastic year
- refugee status
- length of time enrolled in Australian schools relative to that expected of the student's grade.

The 2 risk factors are combined by multiplying the 2 probabilities to form a composite risk index for each student. We can combine them in this way because both risk factors are expressed in probabilities. The weightings for each component variable are then generated sequentially as discussed below.

Before the steps to generate the weightings are discussed, it is noted that this review did also attempt to examine any evidence that would support EAL/D students with significant additional learning needs receiving different levels of funding than other EAL/D students. Additional modelling and analysis undertaken during this review has suggested that the EAL/D learning progression may not be the right tool to assess the needs of this cohort (constituting only 2% of the total EAL/D cohort), and that a further review of support for SSPs is required because of the highly complex needs of EAL/D students with additional learning needs. The ELP review working group has recommended that additional investigation into the best model of support for this student cohort be undertaken over the course of 2021 and 2022 for implementation in 2023.

On the advice from the working group, the new modelling described here includes all students but does not attempt to separately estimate a weighting for students with significant additional learning needs which requires further investigation.

Phase by scholastic year weighting (base weighting for all EAL/D students)

The first step in the sequential approach is to calculate the phase by scholastic year weighting for non-refugee students. This weighting is set as the base weighting for all students.

For risk factor 1, to calculate the phase by scholastic year weighting, we first calculated the average probability of performing at or below NMS in NAPLAN reading or writing tests for non-refugee students in each phase and scholastic year combination (column 3 of Table 1). Note that as NAPLAN tests are only applicable to Year 3, 5, 7 and 9 cohorts, the average probabilities are only available for NAPLAN grades. We extrapolated the average probabilities for NAPLAN grades to non-NAPLAN grades based on recommendations by the EAL/D experts: estimated probability of performing at or below NMS for Year 3 students is applied to students in Kindergarten to Year 2, probability for Year 5 applied to Years 3 to 6, probability for Year 7 to Years 7 to 10 and probability for Year 9 to Years 11 and 12.

Similarly, for risk factor 2 we calculated the average probability of not progressing to the Consolidating phase for non-refugee students in each phase and school level combination (column 4 of Table 1). Then we computed the composite risks (probabilities) by multiplying the 2 risk factors (column 5 of Table 1). The weighting is then calculated as the ratio of composite risk for each EAL/D phase and scholastic year combination relative to the reference group (that is, the K to 2 Consolidating cohort), which has the smallest calculated composite risk.

The base weighting is reported in the last column of Table 1 and is applied to each EAL/D student according to their phase and scholastic year. For example, the weighting of 24.3 for the Beginning phase students in the K to 2 cohort indicates that the funding for this group should be 24.3 times more than that for the Consolidating phase students in the K to 2 cohort (reference group).

Table 1

Base weighting for phase and scholastic year combinations

Scholastic year	EAL/D phase	Risk factor 1	Risk factor 2	Composite risk (risk 1 × risk 2)	Weighting (composite risk/reference)
K-2	Beginning	0.85	0.75	0.64	24.3
	Emerging	0.39	0.67	0.26	9.9
	Developing	0.16	0.53	0.08	3.2
	Consolidating	0.06	0.47	0.03 (reference)	1.0
3-6	Beginning	0.96	0.75	0.72	27.5
	Emerging	0.73	0.67	0.48	18.4
	Developing	0.36	0.53	0.19	7.3
	Consolidating	0.16	0.47	0.07	2.8
7-10	Beginning	0.98	0.70	0.69	26.1
	Emerging	0.80	0.66	0.52	19.9
	Developing	0.57	0.51	0.29	11.1
	Consolidating	0.31	0.42	0.13	4.9
11-12	Beginning	1.00	0.70	0.70	26.5
	Emerging	0.88	0.66	0.58	21.9
	Developing	0.71	0.51	0.36	13.8
	Consolidating	0.46	0.42	0.19	7.4

Figure 3

Base weighting by phase and scholastic years

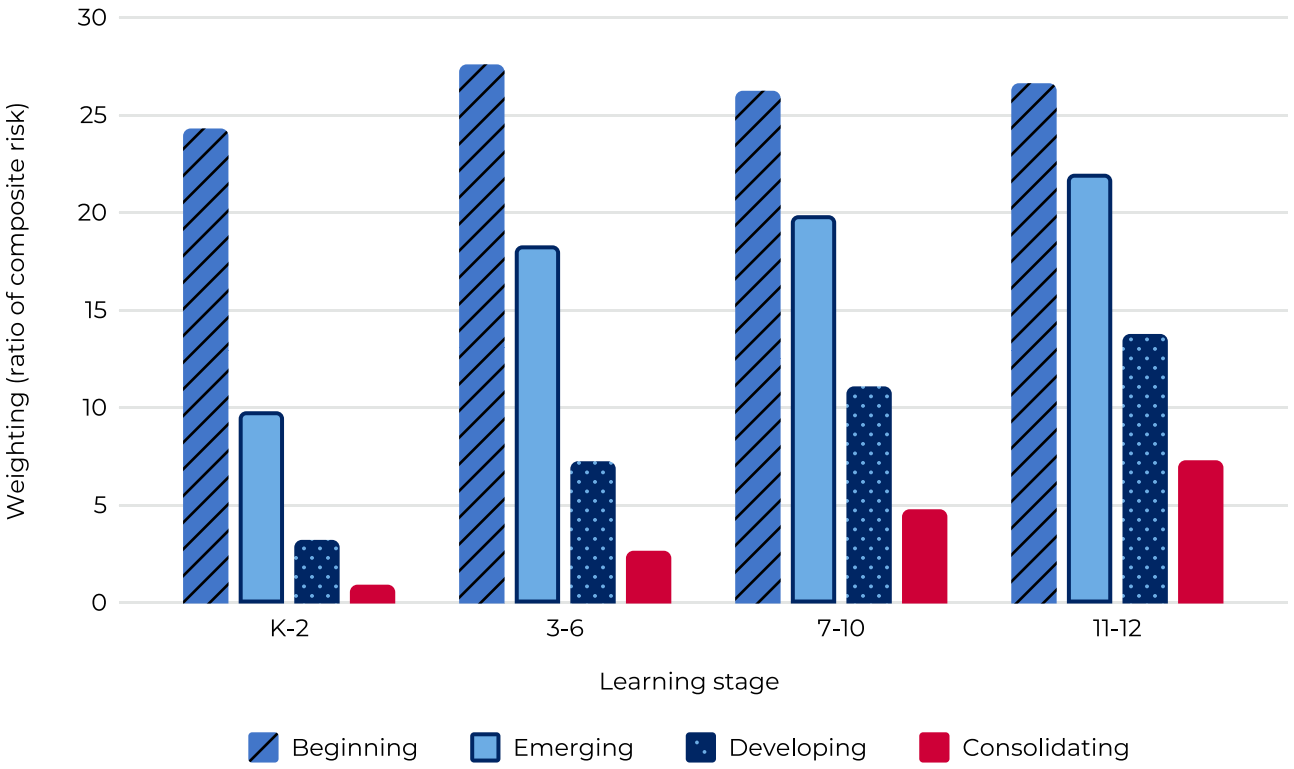


Figure 3 plots the weighting by phase and learning stage. Within each learning stage, weightings are always in descending order from the Beginning phase to the Consolidating phase, as expected. In general, the weighting for each phase increases across learning stages indicating that students in the same phase require more resources and support if they are in higher learning stages.

Refugee status weighting

After the base weightings are determined, the second step is to calculate the weighting for refugee students. In the revised methodology, refugee students are defined as students who are currently on a refugee visa and students who were previously on a refugee visa. This is the same definition used under the current methodology. Only one weighting for refugee students is estimated because this cohort of students only constituted about 3.5% of the NSW EAL/D population, so any further disaggregation could result in volatile estimates.

Similar to the calculation of the phase by scholastic year weighting, the probability of performing at or below NMS in NAPLAN reading or writing tests is first calculated for refugee students and non-refugee students respectively, followed by the calculation of the probability of not progressing to the Consolidating phase or higher for both groups. After that, the 2 probabilities are multiplied to generate the composite risk for each group. The weighting is computed as a ratio of the composite risk of the refugee students to non-refugee students. Table 2 shows the calculation of the refugee status weighting. The resulted weighting of 2.5 indicates that refugee students should receive funding 2.5 times more than non-refugee students. This new weighting is nearly twice that of the current weighting indicating that the funding towards refugee students was underestimated in the current ELP funding model.

Table 2

Refugee status weighting

Refugee status	Risk factor 1	Risk factor 2	Composite risk (risk 1 × risk 2)	Weighting (composite risk/reference)
Non-refugee	0.29	0.55	0.16 (reference)	1.0
Refugee	0.62	0.65	0.40	2.5

Length of time weighting

The current LoT weighting aims to address the fact that, where everything else is the same (grade, ELP phase and refugee status), students who have been enrolled in an Australian school for longer periods of time are comparatively less educationally disadvantaged than recently arrived students. However, in so doing, it unintentionally pushes ELP funding towards Kindergarten and Year 1 students because the length of time enrolled in Australian schools is highly correlated with scholastic year (grade). To appropriately allocate funding to better cater for newer arrivals to the Australian system, we need to adjust for their grade as well as their length of time enrolled in Australian schools when we determine the weighting.

To that end, and as previously stated, we construct a new LoT variable that is a relative measure to remove the association between grades and time enrolled in Australian schools. This variable ranges from −8 to 0 where a value of −8 means that the student has been in Australian schools 8 years fewer than what can be expected given the grade of the student. A value of 0 means that the length of time in an Australian school for a student is as expected given the grade the student is in.

With this range of values for the revised LoT variable, we calculate the 2 risk factors for each value (expressed in years) and combine them as a composite index. The weighting is then calculated as the ratio of composite index for each group relative to the reference group. We estimate 2 sets of weighting for the primary and secondary cohorts respectively.

Table 3 summarises the estimated LoT weightings. Note that the LoT weighting for the '<=−5' group in the primary cohort is smoothed in order to reduce the noise associated with the small cohort sizes. Overall the LoT weighting increases as the length of time relative to expected becomes more negative (that is, more recent arrivals) and the weighting for the primary cohort is higher than that for the secondary cohort. This means that EAL/D students who have been enrolled for a shorter time in an Australian school ('more recent arrivals') than what is normally expected for students in the same scholastic year, are at greater educational risk than their peers. This is particularly so for primary students. Under this new LoT weighting, this cohort of students would receive more funding relative to students in the same grade who were enrolled for a typical length of time.

Table 3

Length of time weighting

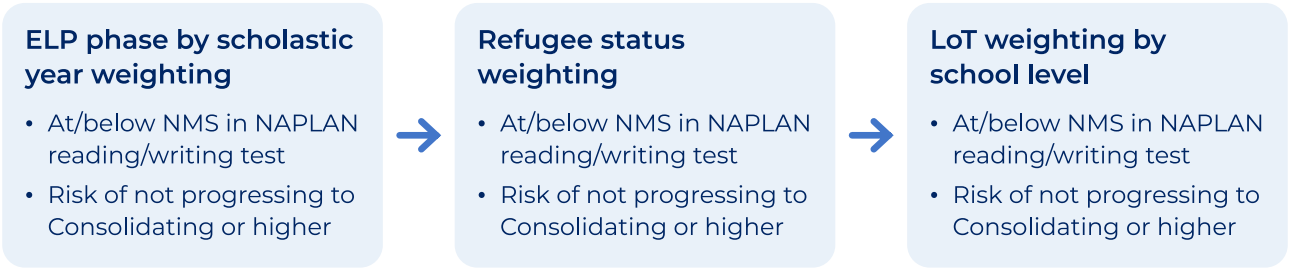
School level	LoT relative to expected	Risk factor 1	Risk factor 2	Composite risk (risk 1 × risk 2)	Weighting (composite risk/reference)
Primary	<=−5	0.78	0.85	0.66	4.00
	−4	0.60	0.72	0.44	3.51
	−3	0.55	0.67	0.36	2.93
	−2	0.38	0.62	0.23	1.89
	−1	0.25	0.60	0.15	1.23
	0	0.21	0.58	0.12 (reference)	1.00
Secondary	<=−7	0.78	0.64	0.50	2.57
	−6	0.72	0.61	0.44	2.24
	−5	0.61	0.58	0.35	1.81
	−4	0.55	0.56	0.31	1.81
	−3	0.48	0.53	0.25	1.30
	−2	0.42	0.49	0.21	1.07
	−1	0.41	0.48	0.20 (reference)	1.00
	0	0.43	0.46	0.20	1.02

It is worthwhile noting that the impact of the LoT weightings on the funding allocation is much smaller than that of the current LoT weighting, which is based on a crude measure and subject to undue influence from grades.

Figure 4 shows the sequence of the steps involved in the funding calculation under the new methodology.

Figure 4

Steps of funding calculation under new methodology



Conclusion

The review of the ELP funding model identifies a number of ways the current model can be improved. With the availability of more years of ELP phase and NAPLAN data, the new methodology expands the current model to consider differential rates of progress on the English language proficiency scale for students of different characteristics, to more accurately estimate the educational risks associated with different EAL/D student groups. The new model also proposes a new way to estimate the risks associated with being a more recent arrival to the Australian system, to better align funding to needs.

The review also attempted to examine any evidence that would support EAL/D students with significant additional learning needs. Modelling undertaken for this review, however, suggests that the EAL/D learning progression may not be the right tool to assess the needs of this cohort, and that a further review of support for these students, especially those in SSPs, is required. As the next step, we will undertake the additional investigation into what is the best way to measure the needs of this cohort as well as the best model of support for this cohort, in the near future.

In summary, the new weightings for the 3 key components of the new funding model are defensible and evidence based. The implementation of the new funding model should lead to more equitable and transparent distribution of funding across schools and different EAL/D student groups.

Appendix 1:
Relevant research in 2014 that underpinned the current ELP loading

Aim of the research

In order to assist the development of the ELP loading, CESE conducted a research project in 2014 which focused on the following 3 questions:

- What is the relationship between the English language proficiency phases of the EAL/D students (measured as EAL/D phases) and student outcomes, taking into account other known risk factors?
- Does this relationship vary across primary and secondary years of learning – that is, should the funding for each phase be different for a primary student than for a secondary student?
- Does this relationship vary across schools with different levels of concentration of students with limited English ability – that is, should concentration of low ELP students in a school be an additional factor in the funding formula?

Data sources

This research used data collected from the department's annual EAL/D census conducted in June/July 2014 for all EAL/D students in Years 3, 5, 7, and 9. The specific EAL/D indicators included were EAL/D phase, length of time students were enrolled in an Australian school and refugee status. To estimate the support needs of students at different EAL/D phases, EAL/D data was then matched to students' test results from the NAPLAN reading test administered in May the same year. Student demographic data from the centralised enrolment system, including parental education level and occupation group, was also added to the matched data file to aid the investigation of the unique relationship between students' language proficiency and outcomes.

Research methods

A series of regression analyses were conducted to investigate the association between EAL/D phases and outcomes. The outcome variable used for most of the analyses was whether or not students achieved at or below the National Minimum Standard in the NAPLAN reading test. This was selected rather than test scores because many EAL/D students are exempted from the NAPLAN test due to limited English proficiency (hence they don't have test scores). As noted by ACARA, students who are **below** the NMS are at risk of being unable to progress satisfactorily at school without targeted intervention, and students who are performing **at** the NMS may also require additional assistance to enable them to achieve their potential.

Consequently, the main form of regression analysis was the multivariable logistic regression controlling for demographic factors. Multilevel analysis using NAPLAN test scores was also conducted to compare the effect sizes of EAL/D indicators to other demographic factors that impact student performance.

Findings

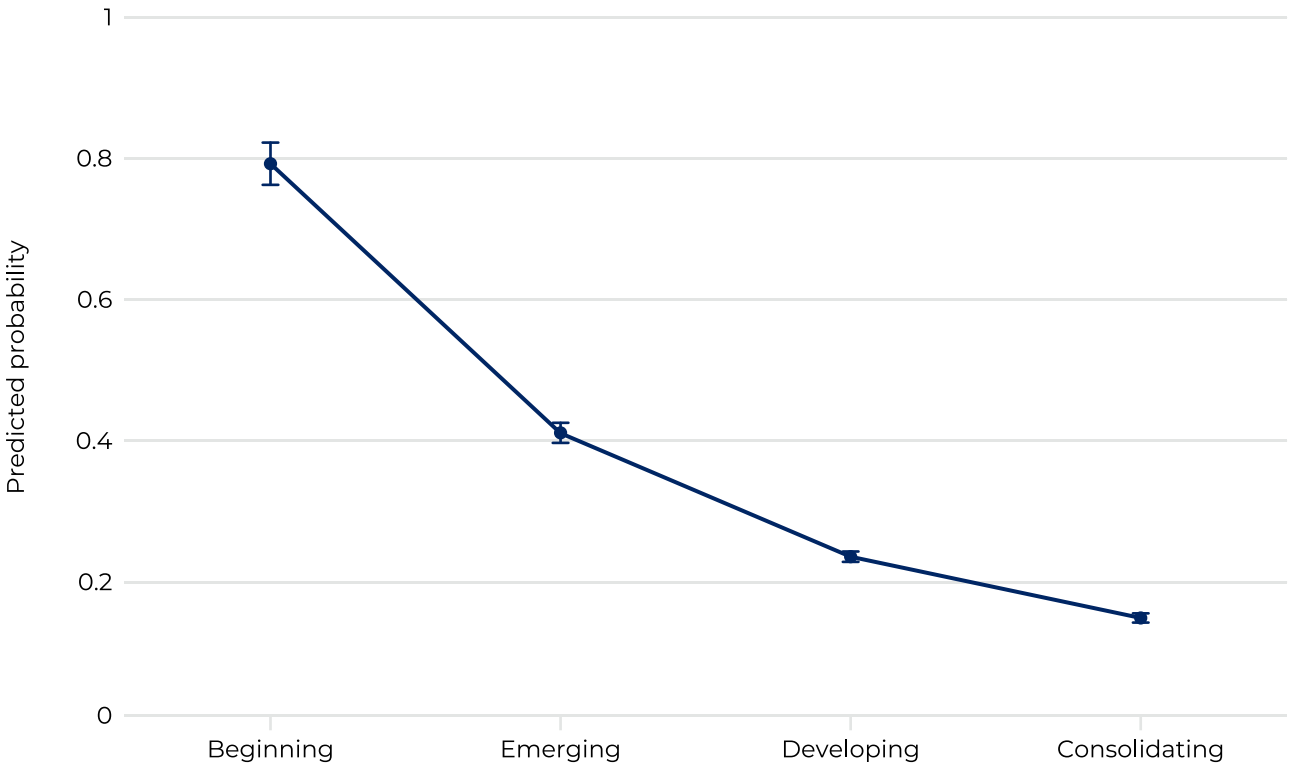
For the first research question, all language background other than English (LBOTE) students whose language proficiencies were assessed by EAL/D teachers for the annual census, and who had matched NAPLAN information, were included in the analysis (N = 40,422). Results indicated that, as expected, students of lower socioeconomic status (SES) have a higher likelihood of achieving at or below NMS. Having controlled for SES, refugee status and new arrival status, students of lower English language proficiency have a higher probability of performing at or below NMS than similar students of higher language proficiency.

Specifically, the results show that the ratios of the probability of students at the 4 EAL/D phases (Beginning, Emerging, Developing, Consolidating) performing at or below NMS are roughly 8:4:2:1. These ratios are taken as indicative relativities in the level of support required for students at different phases, and therefore used as a basis of the funding allocation.

Figure A1 shows the predicted probability of performing at or below NMS for students across the EAL/D phases, when all other variables were set to their grand means.

Figure A1

Predicted probability of performing at or below NMS

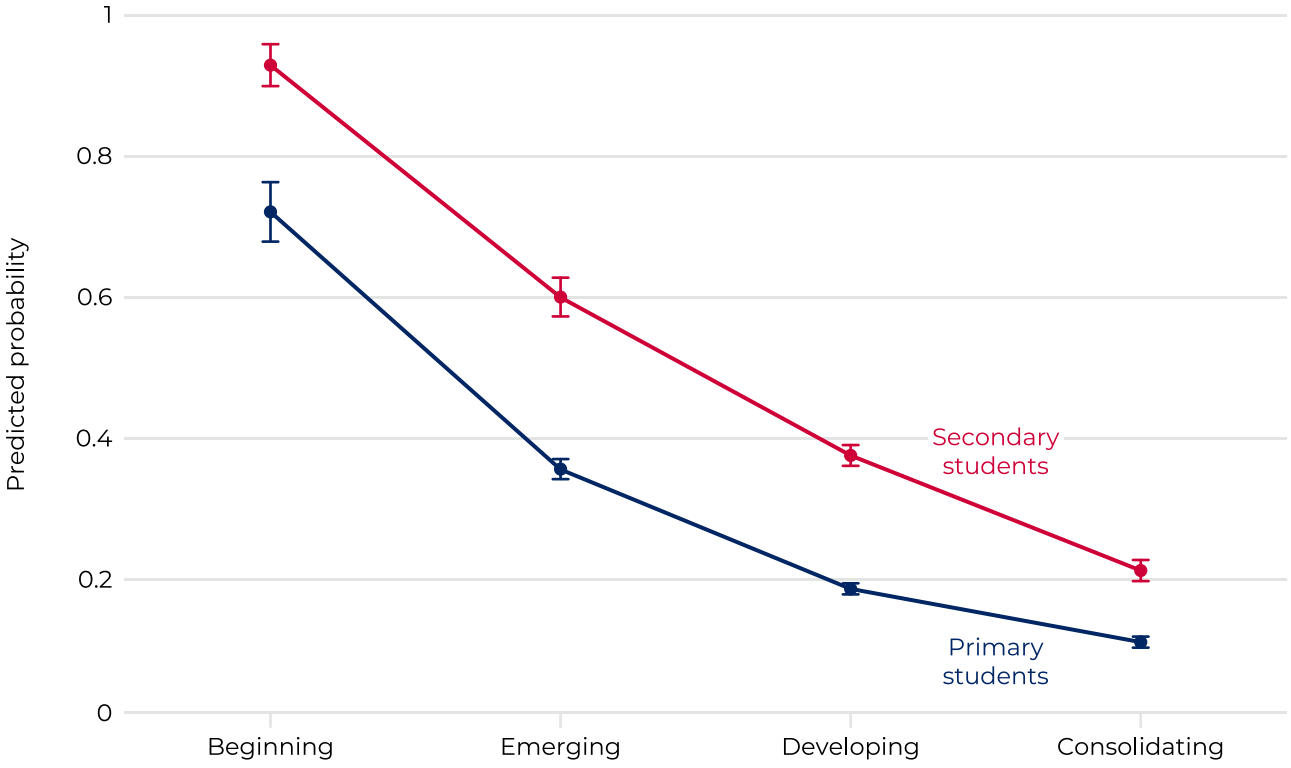


For research question 2, similar regression analyses were carried out for primary and secondary students separately, to investigate whether the relationship between EAL/D factors and student outcomes varies across primary and secondary levels of schooling.

Figure A2 shows the predicted probability of primary and secondary students performing at or below NMS across the 4 EAL/D phases, when all other variables were set to their grand means.

Figure A2

Predicted probability for achieving below NMS – EAL/D primary and secondary students

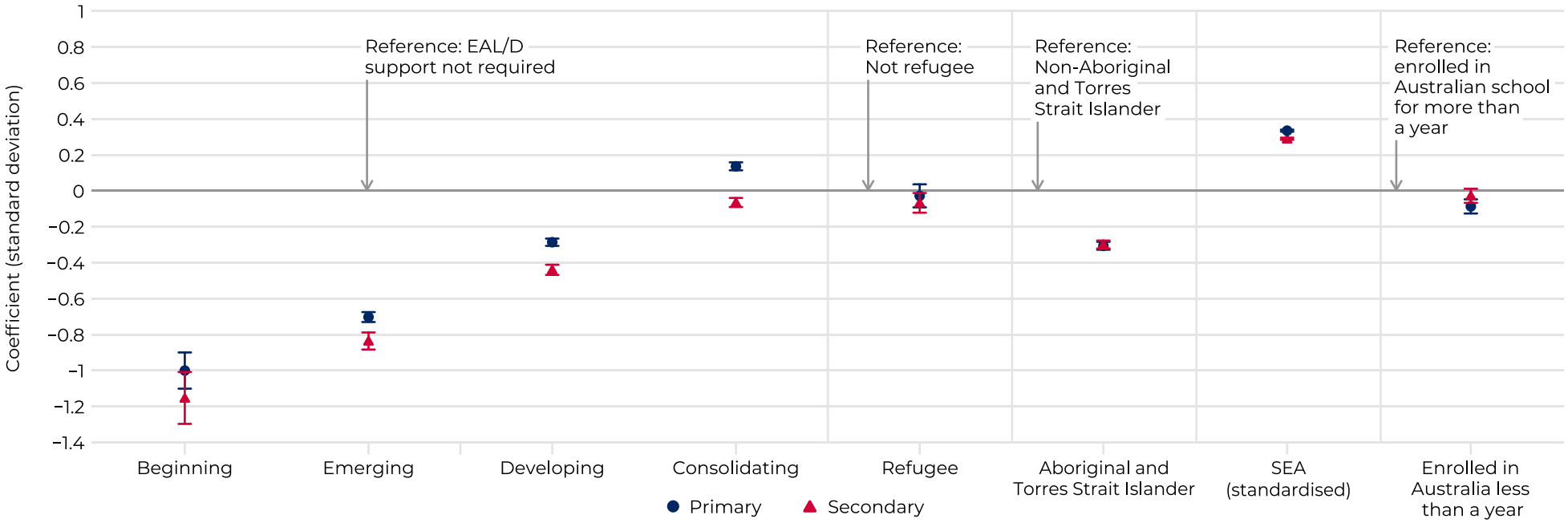


Results indicate that for secondary EAL/D students assessed at the same relative level of proficiency as primary students, the probability of performing at or below NMS is higher. This is consistent with EAL/D experts' understanding that secondary students of limited English proficiency face a greater learning challenge than primary students as they require an even higher level of English proficiency to access the curriculum. The ratio of the probability of secondary students performing at or below NMS versus primary students, across all EAL/D phases, was used as a basis to calculate the learning needs of (and hence the funding required for) secondary students relative to primary students.

An additional analysis compared the effect of limited language proficiency with the effects of other known disadvantage indicators on student academic achievement. This multilevel regression analysis used 2014 NAPLAN reading scores as the outcome variable and included all matched EAL/D and non EAL/D students. Figure A3 displays the effect sizes associated with each EAL/D phase as compared to those of the other factors that also impact student performance, for primary and secondary students separately. Results confirm that limited English language proficiency presents a significant barrier to a student's academic achievement and that the negative effect is greater for secondary students than for primary students. It is noted that the effect size associated with the lowest proficiency phase – Beginning English – is a conservative estimate since many Beginning students were exempted from the NAPLAN reading test, hence were excluded from this analysis. By the same token, the effect associated with refugee status is also a conservative estimate as many refugee students were exempted from the test and therefore not included in this analysis.

Figure A3

Effect sizes of various factors driving NAPLAN performance – primary and secondary students separately



Note. SEA is a composite score assigned to each student, derived from parents' highest education levels and occupation groups.

Finally, for research question 3, multilevel regression analysis, similar to the one performed above, was carried out with an additional variable to investigate whether a high concentration of low ELP students in a school is an additional contributing factor to low achievement for EAL/D students. The additional variable is constructed as the average percentage of students with limited English proficiency (assessed at English as a second language (ESL) phase 1 using the old ESL phase framework) in a school, over a 3 year period (2011 to 2013). Analysis showed that both the main effect of the additional variable and the interaction between this variable and the EAL/D phases were not statistically significant. The results were discussed with the stakeholders and were used as a basis to exclude the concentration of low ELP students from the funding formula.

Appendix 2:
Technical details of multilevel modelling to predict the likelihood of falling at or below NMS in NAPLAN reading and/or writing tests for EAL/D students

The multilevel logistic regression model adopted in the analysis is a 2-level hierarchical model to account for the hierarchical structure of the data with students (level 1 units) nested within schools (level 2 units) (Rasbash et al 2005).

The multilevel logistic regression model relates the log-odds to a linear function of covariates:

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_1 X_{1ij} + \cdots + \beta_n X_{nij} + \mu_{0j} \qquad [A1]$$

where p_{ij} denotes the probability of falling at or below NMS in reading and/or writing tests for an EAL/D student i in school j.

In equation [A1], the outcome variable is the log odds $\log\left(\frac{p_{ij}}{1-p_{ij}}\right)$; X_{1ij} to X_{nij} denote the n student-level controlling factors and β_0 is the average log-odds of all EAL/D students, conditional on the student-level factors. $\beta_1, \ldots, \beta_n$ are the regression coefficients associated with the controlling factors $X_1, \ldots, X_n$. μ_{0j} is the unobserved school effect which is assumed to be normal distribution with mean 0 and variance ${\sigma_{\mu 0}}^2$. Inclusion of a school effect in the model allows the students within a school to be correlated and helps to accommodate the clustering nature of the data.

Separate models are fitted to the primary year (Years 3 and 5) cohort and secondary year (Years 7 and 9) cohort because the relationship between the controlling factors and the outcome variable is different for the 2 cohorts. In this analysis, the student level controlling factors include:

- Aboriginal and Torres Strait Islander status (non-Aboriginal and Torres Strait Islander or missing versus Aboriginal and Torres Strait Islander)
- refugee status (non-refugee versus currently on a refugee visa or previously on a refugee visa)
- student socio-educational advantage (SEA)
- scholastic year (Years 3 and 5 for primary model and Years 7 and 9 for secondary model)
- EAL/D learning progression phase (Beginning, Emerging, Developing, Consolidating)
- number of years spent in the current EAL/D phase
- whether enrolled in Australian schools for less than a year (yes versus no)
- the length of time enrolled in Australian schools relative to that expected of the student's grade (in years)
- student type (mainstream versus additional learning needs).

The mean-variance adaptive Gauss-Hermite quadrature estimation method was used to estimate the model parameters. Tables A1 and A2 summarise the parameter estimates for the primary and secondary models with robust standard errors and whether the parameter estimate is statistically significantly different from zero at a 5% significance level.

Table A1

Regression coefficients of the multilevel logistic regression model for the primary year cohort

Variables	Coefficients	Robust standard error	Test statistic Z	P-value
Aboriginal and Torres Strait Islander status (ref: Non-Aboriginal and Torres Strait Islander)				
Aboriginal and Torres Strait Islander	0.761	0.123	6.200	<0.001
Refugee status (ref: Non-refugee)				
Currently/previously on a refugee visa	−0.015	0.052	−0.280	0.776
Student SEA	−0.194	0.006	−34.020	<0.001
Scholastic year (ref: Year 3)				
Year 5	1.422	0.034	41.780	<0.001
EAL/D phase (ref: Consolidating)				
Beginning	4.696	0.143	32.820	<0.001
Emerging	2.808	0.078	36.010	<0.001
Developing	1.390	0.041	34.220	<0.001
Number of years in current phase (ref: 1 year)				
2 years	−0.310	0.028	−11.190	<0.001
3 years	−0.527	0.035	−15.270	<0.001
4 years	−0.668	0.051	−13.060	<0.001
5+ years	−1.146	0.081	−14.240	<0.001
Whether enrolled in Australian schools for less than a year (ref: No)				
Yes	0.569	0.116	4.900	<0.001
LoT relative to expected (ref: 0)				
<=−5	−0.246	0.180	−1.370	0.171
−4	0.011	0.108	0.100	0.919
−3	0.088	0.086	1.020	0.305
−2	0.033	0.061	0.550	0.584
−1	−0.170	0.054	−3.140	0.002
Student type (ref: Mainstream)				
Additional learning needs	3.797	0.327	11.600	<0.001
Constant	−1.142	0.071	−16.100	<0.001
School-level variance	0.564	0.039		

Table A2

Regression coefficients of the multilevel logistic regression model for the secondary year cohort

Variables	Coefficients	Robust standard error	Test statistic Z	P-value
Aboriginal and Torres Strait Islander status (ref: Non-Aboriginal and Torres Strait Islander)				
Aboriginal and Torres Strait Islander	0.658	0.206	3.200	0.001
Refugee status (ref: Non-refugee)				
Currently/previously on a refugee visa	0.112	0.076	1.470	0.140
Student SEA	−0.177	0.009	−19.240	<0.001
Scholastic year (ref: Year 7)				
Year 9	0.780	0.058	13.550	<0.001
EAL/D phase (ref: Consolidating)				
Beginning	3.868	0.263	14.720	<0.001
Emerging	2.060	0.109	18.880	<0.001
Developing	1.069	0.041	26.250	<0.001
Number of years in current phase (ref: 1 year)				
2 years	−0.202	0.055	−3.710	<0.001
3 years	−0.441	0.057	−7.720	<0.001
4 years	−0.579	0.092	−6.300	<0.001
5+ years	−0.868	0.094	−9.260	<0.001
Whether enrolled in Australian schools for less than a year (ref: No)				
Yes	−0.095	0.162	−0.590	0.558
LoT relative to expected (ref: 0)				
<=−8	−0.208	0.176	−1.180	0.239
−7	0.156	0.146	1.070	0.285
−6	0.096	0.100	0.960	0.339
−5	−0.104	0.092	−1.130	0.260
−4	−0.026	0.074	−0.360	0.722
−3	−0.173	0.068	−2.560	0.011
−2	−0.236	0.074	−3.180	0.001
−1	−0.195	0.084	−2.330	0.020
Student type (ref: Mainstream)				
Additional learning needs	5.251	0.570	9.220	<0.001
Constant	1.175	0.123	9.550	<0.001
School-level variance	0.921	0.166		

Using the parameter estimates from Tables A1 and A2, the probability (risk) of falling at or below NMS in NAPLAN reading and/or writing tests is calculated for each EAL/D student using the following formula:

$$\hat{p}_{ij} = \frac{\exp\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{1ij} + \cdots + \widehat{\beta}_n X_{nij} + \widehat{\mu}_{0j}\right)}{1 + \exp\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{1ij} + \cdots + \widehat{\beta}_n X_{nij} + \widehat{\mu}_{0j}\right)} \qquad [A2]$$

where $\hat{\beta}_0, \ldots, \hat{\beta}_n$ are the parameter estimates. These probabilities (risks) are used later to determine the weightings for the ELP funding model.

Appendix 3:
Technical details of accelerated failure time model with shared frailty to estimate the probability of not progressing to the Consolidating phase or higher for EAL/D students

Survival analysis is a statistical approach to analyse the time taken for 'an event of interest' to occur. All EAL/D students from 2014 to 2019 were tracked to identify progress to the Consolidating phase or higher. 'Event of interest' includes reaching 'Not required' status (the next state after the Consolidating phase) because it is possible that some students, who were in the Developing phase or lower in the previous year, skipped the Consolidating phase and progressed to 'Not required' status. We need to include 'Not required' as an event so that the probability of EAL/D students not progressing to the Consolidating phase can be accurately estimated. Note that in most survival analyses, the event of interest is defined as a failure or a death, which is a negative outcome. However, in this analysis the event refers to a progression from a lower phase to the Consolidating phase or higher, which is a positive outcome.

We proposed using the AFT model with shared frailty to analyse the event time because the AFT model is more flexible that the traditional Cox regression model, which only allows a constant rate of event occurrence. As a student's progression rate to the Consolidating phase is not constant, the AFT model will be more appropriate to model the non-constant or non-monotonic progression rate along the learning progression. The shared frailty specification accounts for the nesting structure of the data where students' progresses on the EAL/D learning progression are more similar within a school. Multiple events were allowed in the model because EAL/D students can travel backward along the learning progression. The risk of not progressing to the Consolidating phase can be estimated for those students who have reached the Consolidating phase at one stage but gone backward to the Developing phase in the next year and then progressed to the Consolidating phase in the following year.

EAL/D census data from 2014 to 2019 was extracted and analysed. Some students or student records were excluded from the analysis with different reasons. These included:

1. EAL/D students with only one observation as their phase progression cannot be tracked
2. EAL/D students who were in the Consolidating phase or higher throughout their observational period as the event time for these students cannot be calculated
3. EAL/D records with unknown time of origin for the first recorded phase (left-censoring) because the event time for these students cannot be calculated. For example, an EAL/D student was in the Emerging phase in 2014 and progressed to the Developing phase in 2015. His EAL/D record in 2014 would be discarded because it is uncertain when he first started on the Emerging phase. However, we can measure his event time starting from the time when he first reached the Developing phase in 2015.

Let T_{ij} denote the time of the occurrence of progression to the Consolidating phase or higher for student i in school j. The AFT model with shared frailty specifies a direct linear relationship between the logarithm of the event time and the contextual factors as well as the shared frailty term α_j for school j:

$$\log T_{ij} = \alpha_j + \beta_0 + \beta_1 X_{1ij} + \cdots + \beta_n X_{nij} + \beta_{n+1} W_{1j} + \cdots + \beta_{n+k} W_{kj} + \varepsilon_{ij} \qquad [A3]$$

where X_{1ij} to X_{nij} are the student-level contextual factors and W_{1j} to W_{kj} are the school-level contextual factors. The student-level factors included Aboriginal and Torres Strait Islander status, gender, refugee status, New Arrivals Program status, student socio-educational advantage, scholastic year, lowest EAL/D phase, student type, whether previously/currently enrolled in IECs, and the length of time enrolled in Australian schools relative to that expected of the student's grade. The school-level factors included school-level mobility rate. $\beta_1, \ldots, \beta_{n+k}$ are the regression coefficients associated with the contextual factors. α_j represents the shared frailty and is assumed to follow a gamma distribution with a shape parameter $\frac{1}{\theta}$ and scale parameter θ. ε_{ij} is the random error which is assumed to be normally distributed. This results in a lognormal regression model. A lognormal model is chosen for this analysis because the EAL/D history data exhibited a non-monotonic instantaneous rate of occurrence of event – specifically, the rate of progress initially increased and then decreased gradually as time went on.

In the lognormal AFT model, the natural logarithm of the event time T_{ij} follows a normal distribution. The lognormal density function $f(t_{ij}|\alpha_j)$ conditional on the shared frailty α_j for the event time is specified as:

$$f(t_{ij}|\alpha_j) = \frac{1}{t_{ij}\sigma\sqrt{2\pi}} \exp\left[\frac{-1}{2\sigma^2}\{\log(t_{ij}) - \mu_{ij}\}^2\right]$$

Given the lognormal distribution of the event time, the lognormal survival function $S(t_{ij}|\alpha_j)$ conditional on α_j, which is the probability of an EAL/D student not progressing to the Consolidating phase or higher given the shared frailty α_j, is expressed as:

$$S(t_{ij}|\alpha_j) = \left[1 - \Phi\left\{\frac{\log(t_{ij}) - \mu_{ij}}{\sigma}\right\}\right]^{\alpha_j} \qquad [A4]$$

where $\Phi(z)$ is the standard normal cumulative distribution function. μ_{ij} is the mean of the lognormal distribution which is linked to the contextual factors given by equation [A3] and σ is the standard deviation, an ancillary parameter to be estimated from the data.

Because the shared frailty α_j is unobservable, it must be integrated out at the group level to obtain the unconditional likelihood function for each school j. The log likelihood is then calculated as the sum of the log-likelihood contributions from each school. Maximum likelihood estimation is used to estimate the model parameters such that the log likelihood function is maximised. Refer to the Stata manual on the command 'streg' for more details (StataCorp 2017). Once all the model parameters are estimated, the probability of not progressing to the Consolidating phase for each EAL/D student conditional on the shared frailty can be estimated using equation [A4].

References

ACARA (Australian Curriculum, Assessment and Reporting Authority) (2014) *English as an additional language or dialect teacher resource, EAL/D learning progression: Foundation to Year 10*, accessed 20 July 2021.

Cummins J (1984) *Bilingualism and special education: issues in assessment and pedagogy*, Multilingual Matters, Clevedon, England.

Garcia G (2000) *Lessons from research: what is the length of time it takes limited English proficient students to acquire English and succeed in an all-English classroom? Issue Brief No. 5*, National Clearinghouse for Bilingual Education, Washington, DC.

NSW Department of Education (2020) *EAL/D advice for schools*, NSW Department of Education, accessed 22 July 2021.

Rasbash J, Steele F, Browne W and Goldstein H (2005) *A user's guide to MLwiN, version 2.32*, Centre for Multilevel Modelling, University of Bristol.

StataCorp (2017) *Stata 15 base reference manual*, Stata Press, College Station, TX.

Centre for Education Statistics and Evaluation
GPO Box 33, Sydney NSW 2001, Australia

Visit our website to subscribe to the CESE newsletter

02 7814 1527
education.nsw.gov.au/cese
info@cese.nsw.gov.au
yammer.com/det.nsw.edu.au

This work is licensed under the Creative Commons Attribution 4.0 International License

NSW GOVERNMENT

235_141021GS_v4_AA1