aitsl
Australian Institute for Teaching and School Leadership Limited

Australian Teacher Workforce Data

# Technical Report

## Supplement to the National Initial Teacher Education Pipeline: Australian Teacher Workforce Data Report 1

Informing the future of the teaching profession

# Table of contents

# 1 Data sources

The Australian Teacher Workforce Data (ATWD) initiative is a national data linkage project designed to help better understand the teacher workforce and teacher education in Australia.

The first report to be produced from the ATWD initiative is the *National Initial Teacher Education Pipeline: Australian Teacher Workforce Data Report 1* (the Pipeline Report).

This supplement to the Pipeline Report provides technical information about the data transformation and analysis processes undertaken to produce the aggregated data included in the Pipeline Report.

**This Technical Report is not designed to be a stand-alone document. It is designed to supplement the information provided in the Pipeline Report and should be read in conjunction with it.**

The Pipeline Report is based on initial teacher education (ITE) data at the unit record level from the Australian Government Department of Education, Skills and Employment. These data now form part of the ATWD and are drawn from the following data sources:

- Higher Education Student Data Collection (HESDC)
- Quality Indicators for Learning and Teaching (QILT).

Each of these data sources are described below.

## 1.1 Higher Education Student Data Collection (HESDC)

The HESDC[1] is a census of administrative and statistical information on higher education students in Australia that is collected twice a year, based on data from all higher education providers (HEPs) as at census dates.[2] It therefore represents a snapshot of students enrolled in higher education programs on those dates.

For the Pipeline Report, data for all students who enrolled in an ITE program at an Australian tertiary institution between 2005 and 2017 (inclusive) were analysed. Data from 2005 was used in the analysis to determine values for 2006 onwards and the Pipeline Report covers the period 2006–2017.

The HESDC included three separate data files which contained the following information:

- Enrolment file—a record for every student enrolled in an ITE program in every year between 2005 and 2017.
- Load file—details of the units of study that each student was enrolled in between 2005 and 2017.
- Course file—details of the programs that each student was enrolled in between 2005 and 2017.

Data are collected by HEPs and submitted to the Australian Government Department of Education, Skills and Employment through the Higher Education Information Management System (HEIMS) under the *Higher Education Support Act 2003* (Cth).

1 The data is deidentified and includes data on enrolments, units of study, courses, and completions for students attending higher education that are eligible for Commonwealth assistance.

2 The census dates are usually at the end of March and the end of August in each year.

## 1.2 Quality Indicators for Learning and Teaching (QILT)

The QILT surveys include the following:

- The Student Experience Survey (SES). This survey collects information about student perceptions of overall quality of the educational experience, teaching quality, learner engagement, learning resources, student support and skills development.
- The Graduate Outcomes Survey (GOS). This survey collects information on the labour market outcomes and further study activities of graduates and includes a course experience questionnaire covering perceptions about overall satisfaction, good teaching and generic skills.
- The Employer Satisfaction Survey (ESS). This survey seeks feedback from employers of the graduates of higher education about the generic skills, technical skills and work readiness of graduates employed in their workplace.

All the QILT surveys are conducted by the Social Research Centre on behalf of the Australian Government Department of Education, Skills and Employment. Participation in the surveys is voluntary and students are offered incentives to participate through prize draws. Details about the surveys and survey methodology are documented on the QILT website.

Note that only data from the Graduate Outcomes Survey was included in the Pipeline Report.

## 1.3 Variables available in the data

Full and detailed specifications for all HESDC data fields are available on the Department of Education, Skills and Employment HEIMSHELP Data Element Dictionary website. **Table 1-1** provides information on the data fields analysed to prepare data for the Pipeline Report.

*Table 1-1: Variables available in the data*

| Items | Variables | HESDC Field Name |
|---|---|---|
| **Year of Birth** | Year | e314 |
| **Residential post code** | Postal address - postcode | e409 |
| **CHESSN** | Commonwealth Higher Education Student Support Number | e488 |
| **Regionality** | SES status (PC measure) | SES_PC |
| | SES status (SA1 measure) | SES_SA1 |
| | first SES status (SA1 measure) | first_SES_SA1 |
| | metro | metro |
| | regional | regional |
| | remote | remote |
| | first_metro | first_metro |
| | first_regional | first_regional |
| | first_remote | first_remote |
| **ATSI Status** | Aboriginal and Torres Strait Islander code | e316 |
| **Country of Birth** | Country of birth code | e346 |
| **Disability** | Disability | e386 |
| **Gender** | Gender | e315 |
| **Citizenship** | Citizenship/resident indicator | e358 |
| **Completion year** | Year of completion | comp_year |
| **Year data collected** | Reference year | e550 |
| **Prior academic achievement** | Highest educational participation prior to commencement | e493 (also 12) |
| **Basis of admission** | Basis of admission | e327 |
| **ATAR** | Tertiary entrance score | e369 |
| **Unique student identifier** | Student ID (Institution) | e313 |
| **Higher education provider** | Institution code | e306 |
| **Course Detail** | Course code (Institution) | e307 |
| | Course name | e308 |
| | Course of study type code | e310 |
| | Field of education code | e461 |
| | Field of education supplementary code | e462 |
| | Course load | e350 |
| | Discipline code | e464 |
| | Equivalent full-time student load | e339 |
| **Course ITE accredited** | Special course type code | e312 |
| **Attendance mode and type** | Mode of attendance code | e329 |
| | Type of attendance code | e330 |

| Items | Variables | HESDC Field Name |
|---|---|---|
| Financial Status – i.e. HECS paid upfront or deferred | Student status code | e490 |
| Subjects studied | Unit of study code | e354 |
| | Unit of study census date | e489 |
| Highest prior qualification | Highest educational participation prior to commencement | e493 (also 2) |
| Curriculum Area | Discipline code | e464 |
| Location of workplace (State) | Location of workplace - state | EMP_STATE |
| Sector of employment | Sector of employment | SECTOR |
| Industry Code | Industry code | ANZSIC |
| Full-time employment availability indicator | Full-time employment availability indicator | AVAILFT |
| Qualification preparedness | Qualification preparedness | CRSPREP |
| Importance of qualification for current job | Importance of qualification for current job | QUALIMP |

# 2 Data processing

The data used in the Pipeline Report reflects the counting of distinct individuals over time. This means that the totals may vary from totals reported in other official statistics based on Higher Education Student Data Collection (HESDC) and Quality Indicators for Learning and Teaching (QILT) data. For example, individuals may have multiple records in a single year within the HEDSC and QILT data. Australian Teacher Workforce Data (ATWD) will only count these individuals once, whereas official statistics based on the HEDSC and QILT may record those multiple occurrences.

## 2.1 Determining distinct individuals

Distinct individuals in the data set were determined based on unique combinations of institution code, student ID and reference year once the data cleansing (described in Section 2.4) process was completed.

Each student was then allocated to different categories including program level (i.e. undergraduate or postgraduate), detailed field of education (i.e. early childhood, primary or secondary education), type of qualification, commencement year, completion year, and enrolment year.

While most of the categories used for aggregating the data were identified from the data itself for most individuals, this was not true for all. As described in the sections below, the data was not always internally consistent or complete. The following sections describe the data transformations that were necessary to ensure that the reporting of individuals within each year of enrolment was as accurate as possible.

The approach to determine distinct individuals in this report will not be required in future Pipeline Reports as the data will include a unique identifier (a personalised project number) to distinguish individuals, and that has been allocated by the Australian Institute of Health and Welfare (AIHW) in the data linkage and preparation stages. This may mean that data in future reports may vary in comparison to the data provided in this inaugural Pipeline Report.

## 2.2 Variables with incorrect labels

The HEIMS load file included a variable that was labelled ‘358’, which is the label for the variable ‘citizenship status’, but the data contained in the field was instead a numeric field which matched the content of reference year information. A further check compared the contents of this field with the year component of the date variable e498 (unit of study census date). Therefore, this field was renamed ‘reference year’. The data for ‘citizenship status’ was instead found from the HEIMS enrolment file.

## 2.3 Data merging process

The purpose of data merging is to combine information from different data sets (enrolment, survey etc). Given that the HESDC data were provided in three separate data files, a means to link these files for each student was devised as follows.

- Firstly, the three HESDC data files were merged. For this purpose, a combination of three variables was used to link data belonging to the same individual: institution code (e306), student ID (institution) (e313), and reference year (e550).
- For Graduate Outcomes Survey (GOS) data, it was found that some individuals had multiple survey records, which might be due to them completing a number of ITE programs in different institutions. For each student with multiple survey responses, only the most recent survey record was selected and the other survey responses were excluded. The data was matched to individuals in the HESDC based on institution code (e306) and CHESSN (Commonwealth Higher Education Student Support Number).
- Out of 13,304 individual records in the QILT-GOS file, 41 records did not match, meaning that HESDC records were not available for these 41 individuals. Therefore, the responses from these 41 individuals were excluded from the analysis in the Pipeline Report.

## 2.4 Data transformation process

Several data limitations were identified and had to be adjusted through the transformation process before analysis for the Pipeline Report could commence. These included the following items, each of which is discussed in detail below.

- Correcting for missing data
- Replacing value for completion year
- Probabilistic data cleansing process for demographic fields
- Standardising values in course description variables
- Derived variables used for classification and calculations
- Reclassified and/or derived variables used for reporting
- Correcting commencement data

### 2.4.1 Correcting for missing data

For a small number of students, the HEIMS course file was missing their record in a particular year, although their record was available in both the enrolment and course load files in that year.

For example, a student did not have a record in the course file for 2005, but from the enrolment file it was clear that the student was enrolled in the same course in 2005 and 2006, so the information about the course was populated for 2005 based on the 2006 information.

Records for a total of 308 students were corrected in this way.

### 2.4.2 Replacing value for completion year

For a total of 121 individuals between 2005 and 2017, completion year had to be corrected to make sense of the sequence of events relating to course commencement and completion over time.

In the example provided in **Table 2-1**, a student was enrolled in two programs—the first one commenced in 2007 and completed in 2010, and the second one commenced in 2010 and completed in 2012. Note that the student's enrolment record showed that in 2010 they completed the first program AND commenced the second program. In addition, the following year's data showed that they were a continuing student in the second program. Therefore, it was determined that the completion year for the first program was 2009 instead.

*Table 2-1: Replacement of completion year (example using dummy data)*

| 306 (institution) | 313 (Student ID) | 307 (Course code) | Reference Year | 922 (Commencement indicator) | Completion Year | Replaced Completion Year |
|---|---|---|---|---|---|---|
| 9999 | ABC1234 | 111AA01 | 2007 | 1 (Commencing) | . | . |
| 9999 | ABC1234 | 111AA01 | 2008 | 2 (Continuing) | . | . |
| 9999 | ABC1234 | 111AA01 | 2009 | 2 | . | 2009 |
| 9999 | ABC1234 | 333BB01 | 2010 | 1 | 2010 | . |
| 9999 | ABC1234 | 333BB01 | 2011 | 2 | . | . |
| 9999 | ABC1234 | 333BB01 | 2012 | 2 | 2012 | 2012 |

### 2.4.3 Probabilistic data cleansing process

Some information that should remain constant for individuals appeared to change over time—for example, year of birth, gender, Aboriginal and Torres Strait Islander status, country of birth, CHESSN or student ID.

For the example provided in **Table 2-2**, the individual (identified by unique CHESSN, institution code and student ID) in all years of enrolment between 2005 and 2008 showed year of birth as '1980' but, in the 2009 enrolment year only, the student's year of birth was recorded as '1981'. It is likely that this was a result of data entry error at the source (i.e. the student incorrectly input the data or there was an administrational error).

*Table 2-2: Data cleansing (example using dummy data)*

| 306 (Institution ID) | Collection Period | Reference Year | 313 (Student ID) | CHESSN | YOB (Year of birth) |
|---|---|---|---|---|---|
| 9999 | 2018 | 2005 | ABC1234 | 1234567890 | 1980 |
| 9999 | 2018 | 2006 | ABC1234 | 1234567890 | 1980 |
| 9999 | 2018 | 2007 | ABC1234 | 1234567890 | 1980 |
| 9999 | 2018 | 2008 | ABC1234 | 1234567890 | 1980 |
| 9999 | 2018 | 2009 | ABC1234 | 1234567890 | 1981 |

## Variables and sequence of change

A process was developed to identify the student records that had variables with values changing over time, e.g. for year of birth, gender, Aboriginal and Torres Strait Islander status, country of birth, CHESSN and/or student ID. The probabilistic approach described below was then applied iteratively to provide recommended values for each variable in the records selected.

The sequence of standardising the values for these variables was as follows:

- First year of birth, gender, Aboriginal and Torres Strait Islander status and country of birth were standardised across time for each unique student based on the same provider code and student ID.
- Then CHESSN was standardised for each unique student based on the student having the same provider code and student ID, and year of birth, gender, Aboriginal and Torres Strait Islander status and country of birth.
- All students with CHESSN values of 'ZZZZZZZ' were changed to 'NULL'.
- Finally, all student IDs with the same CHESSN (for all values other than 'NULL') and matching values for year of birth, gender, Aboriginal and Torres Strait Islander status and country of birth were standardised.

## Criteria used to weight values

Two criteria were applied to the data to determine the weightings attached to any given value for a variable:

- The age of the data record—older records had less weighting than more recent records.
- The number of times the same value was observed for the individual over time—the more often the value was observed, the higher the weighting.

The results of the two criteria were then averaged to determine an overall weighting for each value. Finally, the value of the variable was changed to the recommended value depending on the weighting value.

The number of records affected by this process are listed by variable type in **Table 2-3**.

*Table 2-3: Number of records changed using probabilistic data cleansing*

| Variable | Number of Records Changed |
|---|---|
| Year of Birth | 36 |
| Gender | 62 |
| Aboriginal and Torres Strait Islander status | 1,578 |
| Country of Birth | 1,856 |
| CHESSN | 3,237 |
| Student ID | 4,045 |

## Worked example

The following process applies to how the two-stage criteria were applied to the example illustrated in **Table 2-4**.

### Criterion 1: How old is the data?

It was assumed that older data would be less reliable, and therefore older data had a lower weighting than the more recent one. For example, the 2005 data had a lower weighting than the 2009 data. The following formulae were developed and applied in sequence to determine the weighting for each record based on the age of the data.

Formula 1: $\log_{10}\left[\frac{1}{((\text{Maximum Reference Year}+1)-\text{ReferenceYear})} \times \frac{100}{1}\right]$

Formula 2: $\left(\frac{\text{Answer 1}}{\text{sum of Answer 1 for all occurrences}}\right) \times \frac{100}{1}$

Formula 1 includes a reciprocal function to take into account that the more recent data is more reliable than the older one. Further, more recent data would produce a very large result, and therefore, a logarithmic scale was used to dampen the loading effect. Formula 2 ensures that the sum of possible values is one (i.e. it is assumed that the possible correct value is within the data).

**Table 2-4** illustrates the application of the formulae.

*Table 2-4: Steps to determine weighting for criterion 1*

| Reference Year | Formula 1 | Answer 1 | Formula 2 | Answer 2 (Weighting 1) |
|---|---|---|---|---|
| 2005 | $\log_{10}\left[\frac{1}{2018-2005} \times \frac{100}{1}\right]$ | 0.886056648 | (0.886056648/4.811240207)*100 | 18.41638766 |
| 2006 | $\log_{10}\left[\frac{1}{2018-2006} \times \frac{100}{1}\right]$ | 0.920818754 | (0.920818754/4.811240207)*100 | 19.13890628 |
| 2007 | $\log_{10}\left[\frac{1}{2018-2007} \times \frac{100}{1}\right]$ | 0.958607315 | (0.958607315/4.811240207)*100 | 19.92432873 |
| 2008 | $\log_{10}\left[\frac{1}{2018-2008} \times \frac{100}{1}\right]$ | 1 | (1/4.811240207)*100 | 20.78466169 |
| 2009 | $\log_{10}\left[\frac{1}{2018-2009} \times \frac{100}{1}\right]$ | 1.045757491 | (1.045757491/4.811240207)*100 | 21.73571565 |

### Criterion 2: Frequency of occurrence

The more frequently the value was observed, the higher the weighting. Therefore, 1980 had a higher percentage weighting than 1981, due to its frequency.

Formula 3: $\left(\frac{\text{Frequency of Occurence}}{\text{Total Number of Occurrences}}\right)$

Formula 4: $\left(\frac{\text{Answer 3}}{\text{sum of Answer 3 for all occurrences}}\right) \times \frac{100}{1}$

*Table 2-5: Determination of weighting for criterion 2*

| YOB | Occurrence | Formula 3 (divide by sum of Occurrence) | Answer 3 | Formula 4 | Answer 4 (% Weighting 2) |
|---|---|---|---|---|---|
| 1980 | 4 | 4/5 | 0.8 | 0.8/3.4*100 | 23.5294118 |
| 1980 | 4 | 4/5 | 0.8 | 0.8/3.4*100 | 23.5294118 |
| 1980 | 4 | 4/5 | 0.8 | 0.8/3.4*100 | 23.5294118 |
| 1980 | 4 | 4/5 | 0.8 | 0.8/3.4*100 | 23.5294118 |
| 1981 | 1 | 1/5 | 0.2 | 0.2/3.4*100 | 5.8823529 |

### Overall weighting and determination of correct value

The two weightings resulting from criterion 1 and 2 were averaged (i.e. assigned equal weighting) to determine final weighting for each value. The value with the highest average weighting was then deemed the 'correct' value for the year of birth as shown in **Table 2-6**.

*Table 2-6: Final weighting determined as average of Weighting 1 and Weighting 2*

| Weighting 1 | Weighting 2 | Average weight | YOB | YOB corresponding to max weight will be chosen and replace |
|---|---|---|---|---|
| 18.41638766 | 23.5294118 | 20.9729 | 1980 | 1980 |
| 19.13890628 | 23.5294118 | 21.33416 | 1980 | 1980 |
| 19.92432873 | 23.5294118 | 21.72687 | 1980 | 1980 |
| 20.78466169 | 23.5294118 | 22.15704 | 1980 | 1980 |
| 21.73571565 | 5.8823529 | 13.80903 | 1981 | 1980 |

### Duplicate records

After the cleansing process was applied, if multiple records were found for the same individual in any given reference year, the duplicate entries were flagged to be excluded from the aggregation and analysis process so that each student was counted only once in any single year.

The number of records excluded from the aggregation and reporting process are listed in **Table 2-7**.

*Table 2-7: Final data (records) included in the analysis*

| Type of record | Before cleanse | After cleanse | Records excluded |
|---|---|---|---|
| Enrolment | 983,624 | 983,444 | 180 |
| Load | 6,501,205 | 6,500,884 | 321 |

### 2.4.4 Standardising values in course description variables

Some course description variables had inconsistent values or formatting, even though they were meant to be the same. For example, variable 350 (course of study load) contained both values 4.0 and 40, when the data dictionary specified the value should be a number between 00 and 99. **Table 2-8** shows the modifications that were made to the listed variables.

*Table 2-8: Standardising values*

| Variables (code) | Variable definition | Modification | Before (Example of value) | After (Example of value) | File name |
|---|---|---|---|---|---|
| 307[1], 308[2] | Course code1 / name2 | Converted to upper-case | Master of Teaching | MASTER OF TEACHING | HEIMS_COURSE[1,2], HEIM_ENROLMENT[1] HEIM_LOAD[1] |
| 310 | Course of study type | 2 digits (zero-padded) | 3 | 03 | HEIMS_COURSE |
| 313 | Student Identification code | 10 digits (zero-padded) | xxyyzz | 0000xxyyzz | HEIM_ENROLMENT, HEIM_LOAD |
| 339 | Equivalent full-time student load | Standardised by multiplication (*100) | 0.125000000 | 12.5 | HEIM_LOAD |
| 350 | Course of study load | Removal of decimal point (".") | 4.0 | 40 | HEIMS_COURSE |
| 354 | Unit of study code | 12 digits (zero-padded), removed special characters and spaces | ED520. 306 | 0000ED520306 | HEIM_LOAD |
| 369 | Tertiary entrance score | 3 digits (zero-padded) | 85 | 085 | HEIM_ENROLMENT |
| 461, 462 | Field of education | 6 digits (zero-padded) | 70100 | 070100 | HIEMS_COURSE |
| 464 | Discipline | 6 digits (zero-padded) | 10101 | 010101 | HEIM_LOAD |
| 493 | Highest educational participation prior to commencement | 6 digits (zero-padded) | 32010 | 032010 | HEIM_ENROLMENT |

### 2.4.5 Derived variables used for classification or in calculations

The variables listed in **Table 2-9** were derived and used as specified.

*Table 2-9: Derived variables*

| Variables (code) | Variable definition | Details for use of variable | |
|---|---|---|---|
| Age | Age of student in each reference year | As only year of birth was available for each student, the age of each student was determined at 31 December in each year's enrolment record as follows: **(ReferenceYear – Year of Birth)+1** | |
| FYS | First year of commencement (by ReferenceYear) throughout the history of initial teacher education (ITE) program(s) | FYS was set to the same value as the **ReferenceYear** only when commencement indicator (variable e922) had a value of '1' in a student's enrolment record. | |
| FYS_COU | First year student appeared in a particular program | Derived from **ReferenceYear** for any student who commenced any program. | |
| FYS_UG | First year student appeared in an undergraduate program | Derived from **ReferenceYear** for any student who commenced an undergraduate program. | |
| FYS _PG | First year student appeared in postgraduate program | Derived from **ReferenceYear** for any student who commenced a postgraduate program. | |
| FYS_PERSON | First year student appeared in HESDC between 2005 and 2017 | Derived from **ReferenceYear**. A student will only have one FYS_PERSON populated between 2005 and 2017. | |
| LYC | Last year of completion | LYC was set to the value of variable **'comp_year'** when it was the last course the individual had undertaken. | |
| LYC_COU | Last year of completion for the course | Derived from **comp_year** for last program completed by student between 2005 and 2017. A student may have multiple LYC_COU fields populated if they completed more than one program between 2005 and 2017. | |
| LYC_UG | Last year of completion for the undergraduate program | Derived from **comp_year** for last UG program completed by student between 2005 and 2017. A student may have multiple LYC_UG fields populated if they completed more than one UG program between 2005 and 2017. | |
| LYC_PG | Last year of completion for the postgraduate program | Derived from **comp_year** for last PG program completed by student between 2005 and 2017. A student may have multiple LYC_PG fields populated if they completed more than one PG program between 2005 and 2017. | |
| LYC_PERSON | Last year student completed any program between 2005 and 2017 | Derived from **comp_year** for last completion by student between 2005 and 2017. A student may have only one LYC_PERSON populated between 2005 and 2017. | |
| G_LEVEL | | Variable '310' (Course of study type code) was used to classify graduate levels. | |
| | Graduate level (undergraduate or postgraduate) | For undergraduate:<br>**08:** Bachelors Graduate Entry<br>**09:** Bachelors Honours<br>**10:** Bachelors Pass<br>**13:** Associate Degree<br>**20:** Advanced Diploma<br>**21:** Diploma<br>**22:** Other undergraduate award course | For postgraduate:<br>**04:** Masters (Coursework)<br>**06:** Graduate Diploma/Postgraduate Diploma (pass or honours) involving new academic, professional or vocational area<br>**07:** Graduate Diploma/Postgraduate Diploma (pass or honours) extending skills and knowledge in a professional area previously studied<br>**11:** Graduate Certificate<br>**12:** Doctorate by coursework<br>**14:** Masters (Extended) |

### 2.4.6 Reclassified and/or derived variables for reporting

**Table 2-10** lists all variables that were derived or reclassified for reporting purposes.

*Table 2-10: Reclassified and/or derived variables*

| Variables (code) | Variable definition | Modification | Values before | Values after | Details of how variable was used |
|---|---|---|---|---|---|
| METRO (derived) | Region | Simplified three variables into one variable | Metro: 0<br>regional: 0.6813615<br>remote: 0.3186385 | METRO: 2 | The largest number dictates whether it is Metro ('1') or not ('2'). When no information was available, '99' was assigned. |
| 316 (reclassified) | Aboriginal and Torres Strait Islander | Simplified into Non-Indigenous, Indigenous and unknown | 2 (Non-Indigenous)<br>3 (Of Aboriginal origin but not Torres Strait Islander)<br>4 (Of Torres Strait Islander origin but not Aboriginal)<br>5 (Both Aboriginal and Torres Strait Islander origin)<br>9 (Unknown) | 2 (Non-Indigenous)<br>3 (Indigenous)<br>9 (Unknown | If 316 was either '3', '4' or '5', it was classified as '3'. |
| 330 (reclassified) | Type of attendance | Reclassified based on the entire course, rather than by each year | 1 (Full-time)<br>2 (Part-time) | 1 (Full-time)<br>2 (Part-time)<br>3 (Mixed-type) | If 330 was '1' throughout the course duration, it remained as '1' (and same for '2').<br>If a mix of '1' and '2' were present during the course (e.g. 2005='1', 2006='2', 2007='1', 2008='1'), it was classified as '3'. |
| 358 (reclassified) | Citizen/resident indicator | Simplified into Australian resident and Non-resident | 1 (Australian citizen)<br>2 (New Zealand)<br>3 & 8 (Permanent visa)<br>4 (Temporary visa)<br>5 (None of the above/ international) | 1 (Australian resident)<br>2 (Non-resident) | If 358 was either '1', '2', '3' or '8', it was classified as '1'.<br>If 358 was either '4', or '5', it was classified as '2'. |
| 386 (reclassified) | Disability | Simplified into Yes, No or NA | 10010111 | 1 (Yes) | Each digit represents different disability category. First digit indicates the existence of a disability, impairment or long-term medical condition. |

### 2.4.7 Correcting commencement data

The 'commencing student indicator' (variable 922) shows '1' for a commencing student and '2' for a student, who is enrolled as a continuing student in a course of study. Data interrogation showed that some students will appear in the longitudinal data between 2005 and 2017 without having a commencing student indicator of '1' at all throughout their studies during this time. It is estimated that between 4% and 9% of all enrolments in any year between 2009 and 2017 represented students who commenced a course but had a commencing student indicator showing that they were a continuing student all the way through. The proportion was much higher for enrolment data prior to 2009.

The following logic was developed to correct the commencing student indicator only for students who had completed one program of study and then enrolled in a second or further program of study.

For students who were enrolled in more than one program of study, the commencing student indicator value of '2' was replaced with the value '1' in the year the student first appeared enrolled in the second program (indicated by the completion year). This logic is illustrated in **Table 2-11**.

This logic affected a total of 363 cases in the data.

*Table 2-11: Replacing commencement student indicator value*

| 306 (institution) | 313 (Student ID) | Reference Year | Completion Year | 922 (Commencement indicator) | Replaced 922 |
|---|---|---|---|---|---|
| 9999 | ABC1234 | 2007 | . | 1 | 1 |
| 9999 | ABC1234 | 2008 | . | 2 | 2 |
| 9999 | ABC1234 | 2009 | 2009 | 2 | 2 |
| 9999 | ABC1234 | 2010 | . | 2 | 1 |
| 9999 | ABC1234 | 2011 | . | 2 | 2 |

## 2.5 Transforming unit code data into curriculum areas

The student load file includes a variable, 'discipline code' (464), for each unit of study that a student is enrolled in. The code assigned to each unit is based on the field of education classification values in the *Australian Standard Classification of Education* (ASCED) by the Australian Bureau of Statistics (ABS).

For example, a mathematics unit will have a six-digit code '010101', where the first two digits represent the broad field of 'Natural and Physical Sciences', the first two digits together with the middle two digits represent the narrow field of 'Mathematical Sciences', and the entire six digits provide the detailed field of education, 'Mathematics'.

When the discipline codes in the student load file (HEIMS_LOAD) were investigated, it was identified that most ITE program units in the file were designated only with a general education classification code. For example, 'Curriculum Studies', 'Education Studies' or 'Teacher Education: Secondary'. As a result, this field could not be used to determine areas of specialisation for ITE students and graduates.

If students were undertaking combined degrees, for example, Bachelor of Arts and Bachelor of Education, then the units undertaken in the Bachelor of Arts component were more likely to be coded meaningfully according to the field of education, while only general education codes were applied to the units undertaken in the Bachelor of Education component. This meant that information about teacher content knowledge could only be gleaned from analysis of the discipline code for those undertaking combined degrees. This data could not be used to determine exactly which areas of specialisation were studied. In addition, ITE students undertaking combined degrees comprise only a small subset of the entire ITE population, so it was determined that there would be little value in the analysis of the discipline code as provided in the HESDC file.

Therefore, to understand the curriculum areas that ITE students were preparing to teach, it was important to investigate each of the units/subjects that the students undertook over the entirety of their program of study.

A method was developed to obtain additional information about the subjects studied using subject codes (variable 354) to understand the curriculum areas being studied. University handbooks, from current and past programs, were sourced. The information collected was used to recode the discipline code for each subject code using ASCED classification, and to determine if the subjects were about the discipline (content knowledge) or about how to teach the discipline (pedagogical knowledge).

### 2.5.1 Recoding the unit data

The HESDC data in the ATWD includes generic course codes, specific to each ITE program, which are not usable in providing information about the units ITE students are studying, and therefore what they will become qualified to teach.

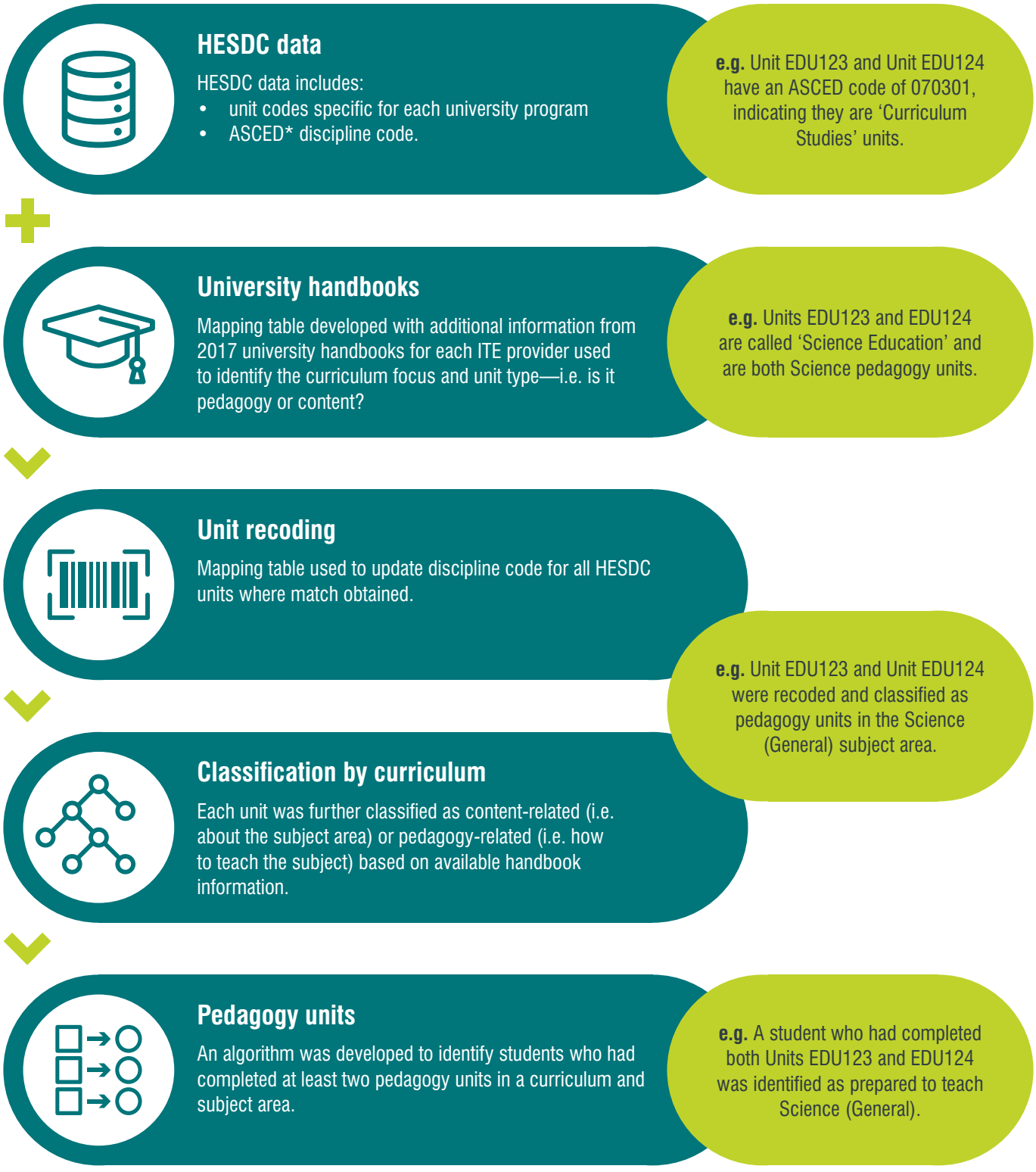
Unit codes studied by undergraduate and postgraduate students enrolled in secondary education programs in 2017 were therefore recoded from a generic Education ASCED discipline code, to a curriculum-based discipline code. While data for all students who were enrolled in 2017 was recoded, the analysis was based on those who completed their studies in 2017.

The HESDC contains a classification code for each unit a student is enrolled in based on the ABS *Australian Standard Classification of Education*. ITE units are most often classified as generic 'education' subjects, and generally do not provide useful information on the curriculum area students are able to teach unless the unit is an undergraduate unit offered outside of the education field.

To specifically identify the subjects each ITE student has studied, it was necessary to recode the HESDC data in the ATWD in a four-stage process.

**Figure 1** shows the transformation process undertaken to recode generic course code data into a usable form.

*Figure 1: Process for recoding the unit data*




*ASCED is the Australian Standard Classification of Education maintained by the Australian Bureau of Statistics

**Stage 1:** A lookup table was prepared including all the unit codes found in the 2017 HESDC file. Additional information from university program handbooks was used to reclassify as many units in the lookup table as possible to indicate:

- the unit content (i.e. Science, Mathematics, Creative Arts, History, Politics, Humanities and so on)
- if the unit focused on content knowledge or pedagogical content knowledge
- if the unit included a practicum component
- if the unit included a literacy or numeracy component.

**Stage 2:** Unit information was updated for all students enrolled in an ITE program in 2017 by comparing the lookup table with the HESDC data and applying the new classification code to matching unit codes.

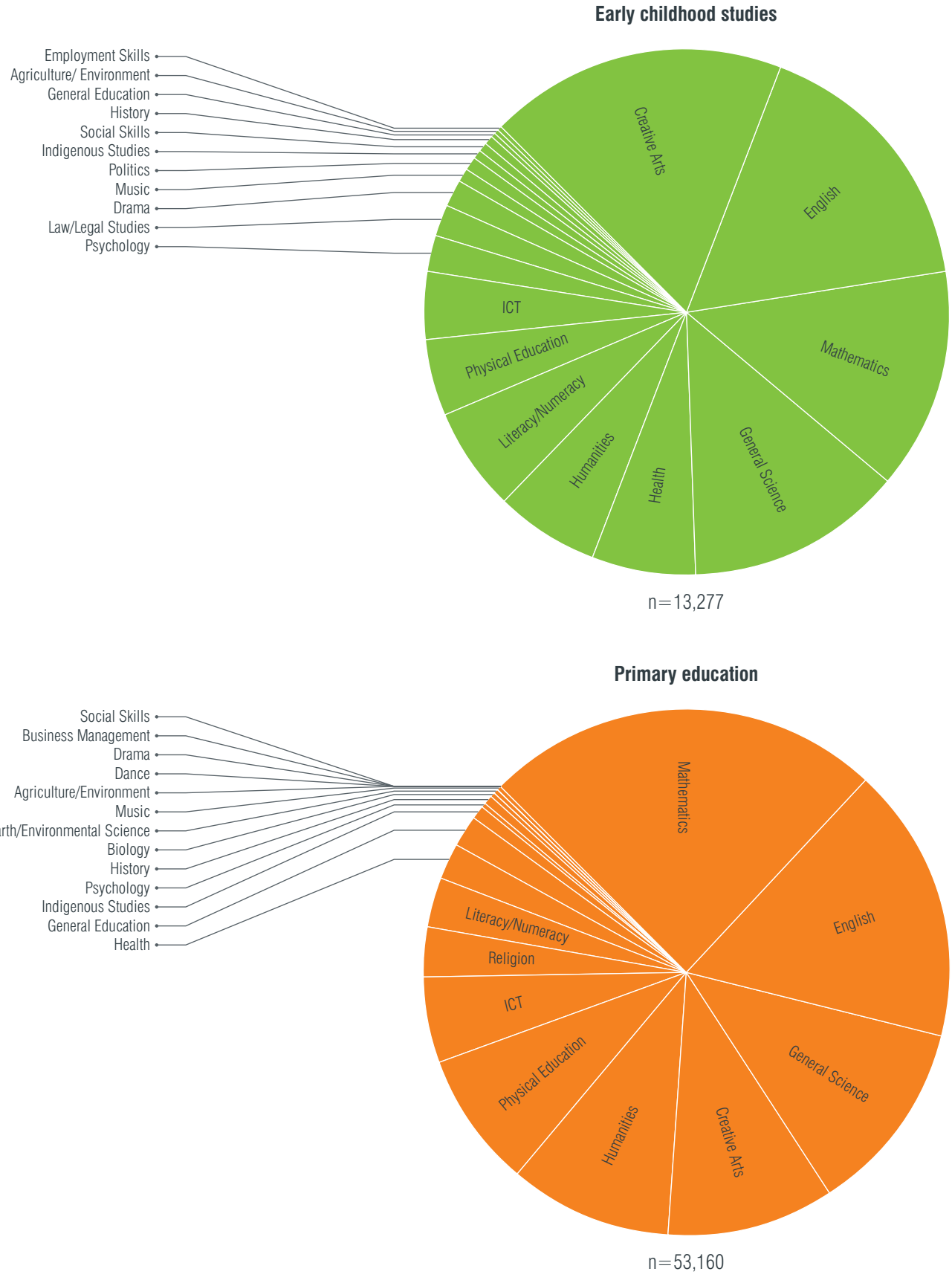
**Stage 3:** The units studied by each student were summarised by detailed subject group (i.e. General Science, Mathematics, Creative Arts and so on).

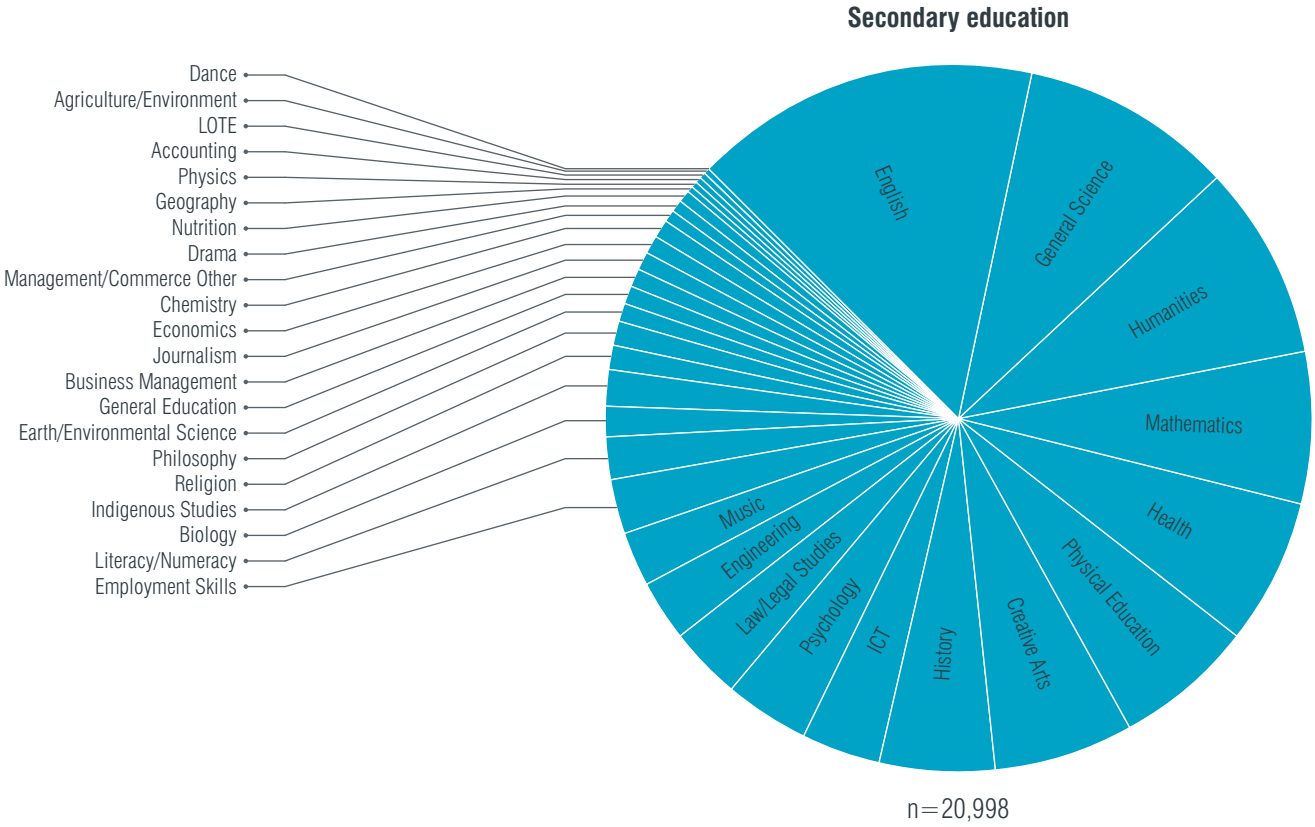
**Stage 4:** An algorithm was applied to each student's data to determine whether they had completed at least two pedagogy units in each subject, preparing them to teach the subject.

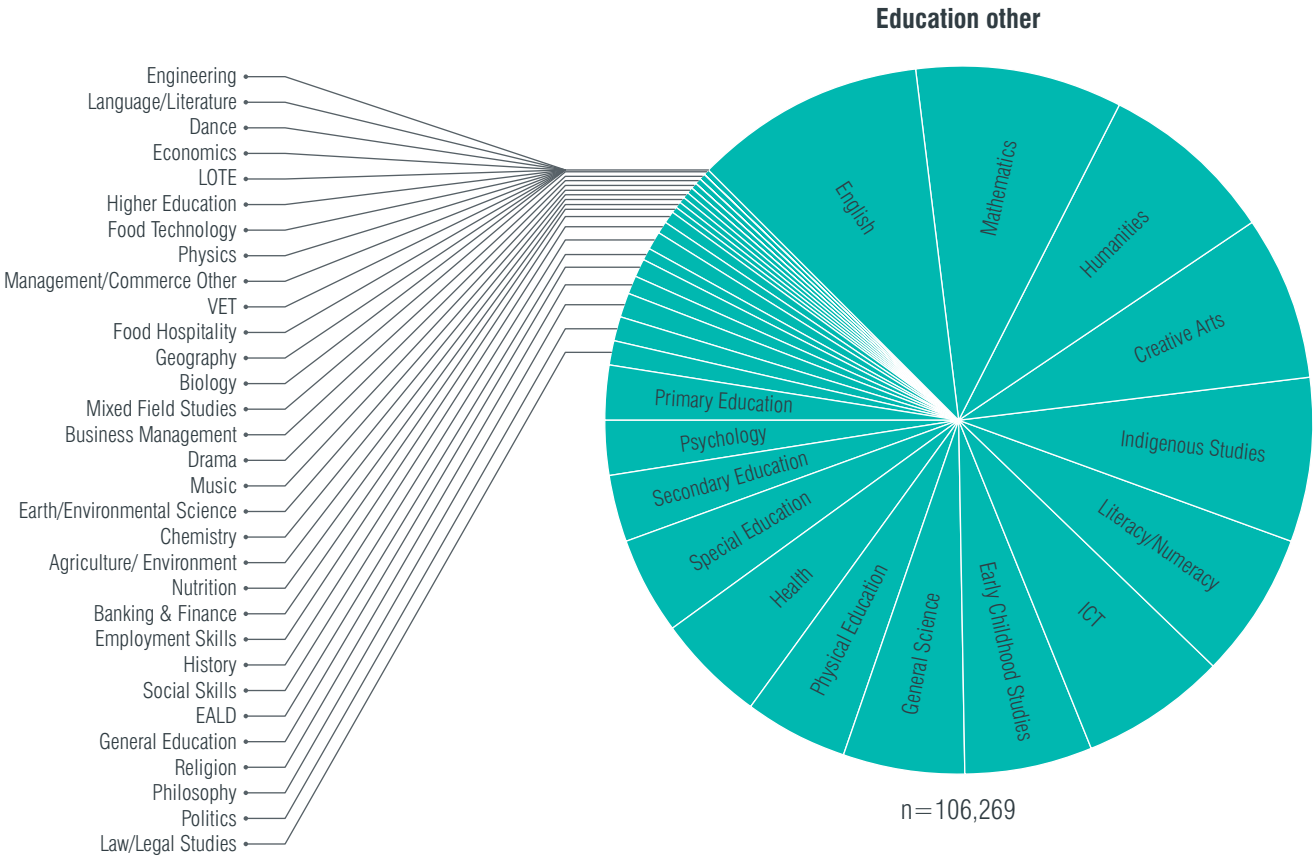
This recoding was completed for 2017 only, and was achieved for 75% of all ITE students enrolled in 2017. Seventy-nine percent of all students who commenced ITE studies, and 64% of all students who completed their ITE studies in 2017, had their subject data updated through this recoding process.

**Figure 2** illustrates how these units were transformed through the recoding process. Each segment of the pie chart represents a different subject.

*Figure 2: Data available before and after subject recoding process*




Secondary education
Dance
Agriculture/Environment
LOTE
Accounting
Physics
Geography
Nutrition
Drama
Management/Commerce Other
Chemistry
Economics
Journalism
Business Management
General Education
Earth/Environmental Science
Philosophy
Religion
Indigenous Studies
Biology
Literacy/Numeracy
Employment Skills
English
General Science
Humanities
Mathematics
Health
Physical Education
Creative Arts
History
ICT
Psychology
Law/Legal Studies
Engineering
Music
n=20,998


Education other
Engineering
Language/Literature
Dance
Economics
LOTE
Higher Education
Food Technology
Physics
Management/Commerce Other
VET
Food Hospitality
Geography
Biology
Mixed Field Studies
Business Management
Drama
Music
Earth/Environmental Science
Chemistry
Agriculture/ Environment
Nutrition
Banking & Finance
Employment Skills
History
Social Skills
EALD
General Education
Religion
Philosophy
Politics
Law/Legal Studies
English
Mathematics
Humanities
Creative Arts
Indigenous Studies
Literacy/Numeracy
ICT
Early Childhood Studies
General Science
Physical Education
Health
Special Education
Secondary Education
Psychology
Primary Education
n=106,269

The data includes the subjects being studied by ITE students who were enrolled in 2017 including commencing, continuing and completing students.

All subjects studied by these students over the course of their ITE program were counted, including subjects undertaken in years prior to 2017 (although only 2017 subjects were recoded as described above).

The data does not differentiate between students enrolled in single or multiple areas of study. Further analysis will be required to investigate the combinations of subjects undertaken by individual students.

The 'education general' category includes units that could not be classified under specific curriculum subject areas (e.g. subjects under STEM, humanities and so on). Some of the units included in this category were called, for example, 'Education studies—contextualising primary education' or 'Assessment and reporting'.

The proportion is higher for students who commenced than for those who completed, because the subject classification was based on 2017 units. Students who completed studies in 2017 would have had fewer 2017 units than those who commenced.

### 2.5.2 Recoding results

#### Units studied in 2017

Of the units undertaken by students completing secondary education ITE programs in 2017, 22% (n=3,297) of undergraduate and 16% (n=4,039) of postgraduate units could be accurately recoded as curriculum area subjects by using program handbook and website information from ITE providers. The remaining units could not be recoded, either due to insufficient or unavailable information from university handbooks, or because the information available was not sufficient to accurately identify if a unit was specifically related to a particular curriculum area.

#### Units studied prior to 2017

Only 9% (n=16,395) of undergraduate and 6% (n=17,909) of postgraduate units undertaken prior to 2017 were recoded as curriculum area subjects. As the subject information from university handbooks was obtained for 2017 only, many of the units undertaken before 2017 did not match the mapping profile.

### 2.5.3 Algorithm applied to determine potential curriculum areas being prepared to teach

#### Undergraduates

Consideration was given to the application of a content–pedagogy algorithm based on the *ITE Accreditation Standards and Procedures*[3] criteria for undergraduate students being qualified to teach in a curriculum area. These standards and procedures require the completion of at least four content units and two pedagogy units to be qualified to teach in a curriculum area.

#### Postgraduates

Consideration was also given to apply the *ITE Accreditation Standards and Procedures* criteria to postgraduate students. The standards and procedures require that postgraduates must complete at least two pedagogy units in a given curriculum area to be qualified to teach in that area.

3 Accreditation of initial teacher education programs in Australia – Standards and Procedures

## The algorithm

An algorithm was developed to apply a content test (that the student had completed at least four content units throughout their program), and a pedagogy test (that the student had completed at least two pedagogy units throughout their program), in a given curriculum area.

To identify students who were prepared to teach in a given curriculum area, both the content and pedagogy tests were applied to units undertaken by undergraduate students, and only the pedagogy test was applied to units undertaken by postgraduate students.

When only the pedagogy test was applied to the data, of the 6,947 students completing a secondary education program in 2017, 51% (n=3,541) could be identified has having completed at least two pedagogy units in a curriculum area during their studies—this incorporated both undergraduate and postgraduate students. These were therefore available to be included in the analysis of subjects being prepared to teach.

When both the pedagogy and content test was applied to the data, the number was reduced to 34% (n=2,333) because only a third of the undergraduate students that met the pedagogy test also met the content test. Given the low rates of being able to recode pre-2017 data, this result was not unexpected.

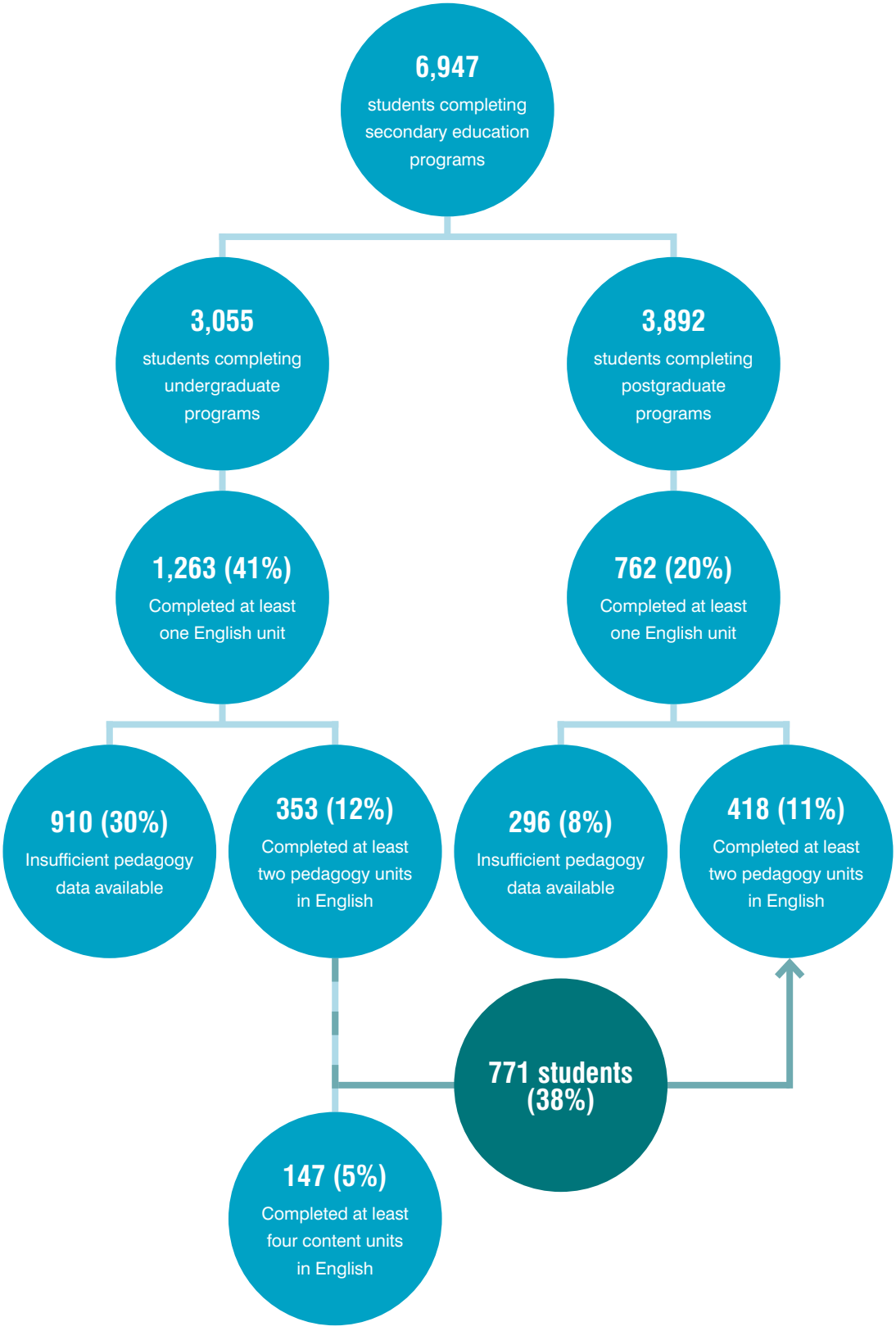
In order to progress this analysis, it was assumed that the pedagogy test alone would suffice to identify undergraduate students who were prepared to teach in particular curriculum areas. It is not unreasonable to assume that, as completing ITE undergraduates, these students will also have met their content requirements, even though this analysis is unable to show that with the data. Therefore, the algorithm that was applied to the data, and the subsequent analysis, is based on the pedagogy test alone to determine if students were prepared to teach in a particular curriculum area.

## Example of the process applied to identify 'prepared to teach English'

**Figure 3** provides an example of the process of applying the analysis that was used to identify students who have been prepared to teach English in 2017.

- The potential pool of students eligible to teach English was identified by selecting all students with at least one content or pedagogy unit in English. Of all students, 28% (both postgraduate and undergraduate; n=2,025) completed at least one unit in English.
- Applying the pedagogy algorithm: Of all those potentially studying to teach English (n=2,025), 771 (made up of 353 undergraduate students and 418 postgraduate students) had completed at least two English pedagogy units during their program. This represents 11% of all students completing secondary education ITE programs.
- Applying a content and pedagogy algorithm: Of the 353 undergraduate students, only 147 had four content units in English, reducing the proportion of undergraduate students who are clearly prepared to teach English to 5%.
- A higher proportion of postgraduate students (55%, or 418) had completed two pedagogy units in English, compared with undergraduate students (28%, or 353).

*Figure 3: Process to identify ITE prepared to teach English from completing secondary education cohort 2017*



### 2.5.4 Extraction of unit information from university handbooks (2017)

A list of the 14,344 unique subject codes for each institution from the 2017 course load file was used as the basis for sourcing additional information about each subject code. Data for the 2017 subject codes was collected in a reference table to be used to recode the HESDC data including the following information:

- the unit name
- the curriculum area that the unit was focussed on (for example Science, Mathematics, Creative Arts, History, Politics, Humanities etc.). This field was coded using the ASCED classification system. The field was used to update the discipline code in the load file as described in the next section.
- if the unit focused on content knowledge or pedagogical content knowledge
- if the unit included a practicum component
- if the unit included a literacy or numeracy component.

While for some institutions, the subject codes are consistent for the years before and after 2017, this is not always the case. Therefore, the 2017 subjects undertaken by ITE students are the most complete and will provide a baseline of information going forward.

### 2.5.5 Recoding discipline code using unit information

Unit information in the student load file was then updated by comparing the ASCED code value assigned to each unit code in the reference table to the discipline code value in variable 464 for all records with matching unit codes (354).

Logic was developed to determine whether the original value for variable 464 (discipline code) should be recoded by comparing the corresponding unit code (354) with the lookup table. A series of case logics were applied to the data to determine whether the value should be recoded as follows:

- All discipline code values in the HESDC file were ranked, with general education code values having lower ranking than other codes.
- The reference table ASCED codes were also ranked using the same ranking schema.
- For matching subject codes in the HESDC student load file and the reference table, the discipline code (464) in the former was updated only when the ranking of the ASCED code value in the reference table was higher than the ranking of the original discipline code value.

In general, the original value was recoded with the value provided in the reference table when the value in the reference table was not ‘Education other’. Statistical analysis software (SAS) code was developed to run this logic and can be provided upon request.

*Table 2-12: Examples of unit information recoding*

| 306 (institution) | 354 (unit code) | 464 (discipline code) | |
|---|---|---|---|
| | | Original Value | Recoded |
| Institution_A | EDU124 | 070100 (Teacher Education) | 090311 – Pedagogy (Indigenous Studies) |
| Institution_B | ESH106 | 070101 (Primary Education) | 091599 – Content/Literacy (English) |
| Institution_C | CUR4212 | 070105 (Secondary Education) | 029999 – Pedagogy/Professional Practice (ICT) |
| Institution_D | AEB4169 | 079900 (Education Other) | 010101 – Pedagogy / Numeracy (Mathematics) |

The match rate between HESDC unit data and the reference table is shown below. The match rate was of course highest in 2017, as the reference table was created using 2017 university handbooks.

*Table 2-13: Recoding match rates for unit information*

| ReferenceYear | Match | % | Non-Match | % |
|---|---|---|---|---|
| 2005 | 40444 | 9% | 412679 | 91% |
| 2006 | 47648 | 11% | 396735 | 89% |
| 2007 | 53314 | 12% | 406371 | 88% |
| 2008 | 61824 | 14% | 394983 | 86% |
| 2009 | 78078 | 17% | 386067 | 83% |
| 2010 | 110744 | 23% | 379963 | 77% |
| 2011 | 150798 | 30% | 348757 | 70% |
| 2012 | 182498 | 35% | 340095 | 65% |
| 2013 | 215652 | 41% | 308381 | 59% |
| 2014 | 245109 | 46% | 287899 | 54% |
| 2015 | 315924 | 59% | 223953 | 41% |
| 2016 | 353584 | 65% | 189669 | 35% |
| 2017 | 420750 | 74% | 148965 | 26% |

### 2.5.6 Number of records affected by recoding of discipline code

Students enrolled in secondary education programs during 2017 were used as the reference group for recoding of discipline code.

The recoded discipline areas were summarised by detailed subject group as mapped in the modified ASCED Subject Grouping table provided at **Appendix 2**.

**Table 2-14** shows the effect of the recoding on the number of subject groupings that have been replaced from unspecified categories (for example 'education other') to specific curriculum areas (for example Chemistry) using the 2017 reference table. The percentage (%) indicates the proportion of curriculum areas students were studying.

*Table 2-14: Effect of recoding unit information*

| ReferenceYear | All HEIMS data | % | Secondary Education (Undergraduate) | % | Secondary Education (Postgraduate) | % |
|---|---|---|---|---|---|---|
| 2005 | 16549 | 16% | 3612 | 8% | 569 | 33% |
| 2006 | 19341 | 18% | 4467 | 9% | 925 | 59% |
| 2007 | 22115 | 19% | 5593 | 11% | 1149 | 56% |
| 2008 | 24355 | 22% | 6248 | 13% | 939 | 58% |
| 2009 | 28788 | 25% | 7769 | 15% | 701 | 54% |
| 2010 | 45098 | 34% | 9817 | 18% | 3353 | 86% |
| 2011 | 63851 | 43% | 13663 | 23% | 4763 | 91% |
| 2012 | 73666 | 46% | 14206 | 23% | 6263 | 93% |
| 2013 | 89687 | 52% | 16240 | 25% | 7351 | 94% |
| 2014 | 102680 | 56% | 20301 | 29% | 8714 | 95% |
| 2015 | 135683 | 63% | 24946 | 34% | 11744 | 94% |
| 2016 | 149635 | 65% | 26287 | 36% | 13889 | 95% |
| 2017 | 175253 | 68% | 32952 | 41% | 15375 | 94% |

# 2.6 Detailed field of education

## 2.6.1 Accurately identifying detailed field of education

The raw HESDC data[4] for detailed field of education includes student records from 2006–2017 classified as 'education other'—a generic education code, resulting in around 20% of students in any year for whom the field of education being studied is unknown.[5]

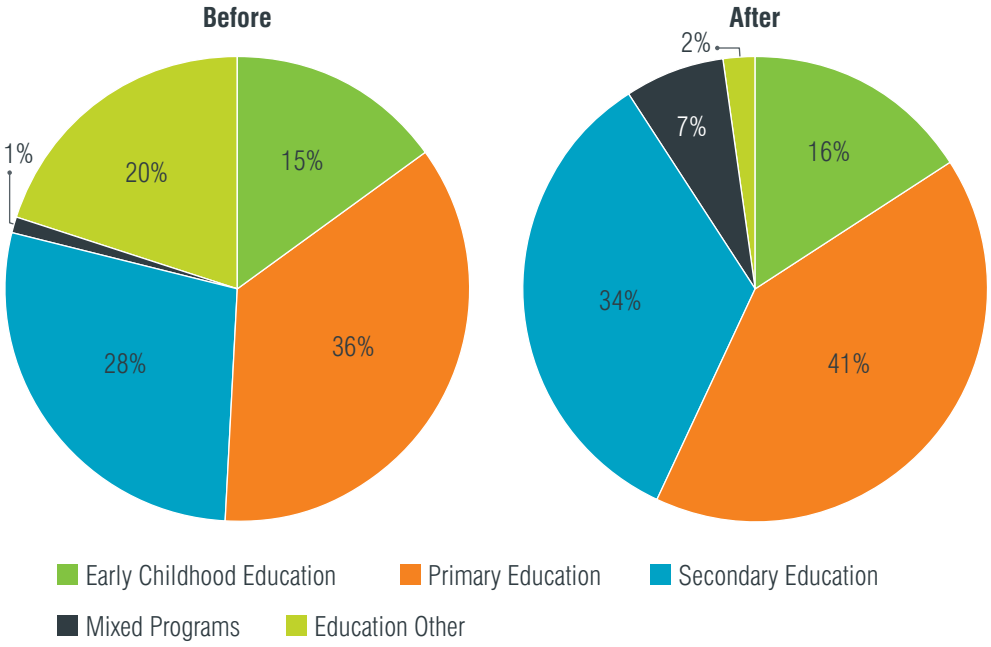
To improve the data quality for analysis of detailed field of education, the 'education other' code was recoded using subject code data for the units each student was enrolled in to inform the detailed field of education. For example, subjects indicated as 'teacher education: primary' or 'teacher education: secondary' informed the field of education recode to 'primary' or 'secondary'.

As an example, **Figure 4** shows the impact of the recoding for the 2017 data. In 2017, 88% of 'education other' codes have been reclassified into accurate detailed field of education categories, reducing the 'education other' category from 20% to 2% (n=18,520 to 2,284 students).

This process was completed for all years 2006–2017, providing greater insight into the numbers and relative proportions of ITE students being trained for early childhood, primary or secondary teachering over the last decade.

Given that the 'mixed program' and 'education other' categories constitute a relatively small proportion of the overall number of students (mixed program 7% and education other 2%), and it is not possible to determine the exact field of education a student is studying, these categories have been included only in the aggregate totals for all ITE students in this report. Detailed profiles of the characteristics of students in these categories have not been provided in the report.

*Figure 4: 'Detailed field of education', before and after recoding by subjects studied*




4 'Raw data' refers to HESDC data prior to being recoded as described in this section.

5 These are codes other than Early Childhood (070101), Primary (070103) and Secondary Education (070105).

### 2.6.2 Enhancing detailed field of education data

This section details the steps undertaken to enhance the data in the detailed field of education (dFOE) to provide more detail about the ITE qualification for individuals who studied programs classified as 'education other' in the raw HESDC data.

The logic was applied to the data in the following sequence to maximise the redistribution of students classified as 'education other' to one of the categories: 'early childhood', 'primary', 'secondary education' or a combination of these (i.e. 'mixed program'):

1. Recoding based on discipline code information at the program and/or unit level
2. Recoding based on the information provided in the name of the ITE program
3. Recoding based on the information provided in the subject information

### 2.6.3 Recoding based on discipline code at program and unit levels

The detailed field of education is derived from the value in the variable 461 (field of education) and 462 ('field of education for double major') and is used to denote students who are studying early childhood, primary or secondary education ITE programs.

It was determined that dFOE data quality could be enhanced if updated discipline codes at the course or unit level could be used to inform recoding of dFOE. The unit information used in this first pass included the raw values in the discipline code (variable 464) in the HESDC.

The discipline codes for units studied by students with a course-level discipline code of 'education other' were analysed to determine if they provided greater insight into the detailed field. As shown in **Table 2-15** below, while the course-level code for this individual was 'education other', and the units that the student had studied had a generic 'education other' value of '070100', the student completed two units with a discipline code of 'primary education' ('070103').

*Table 2-15: Recoding detailed field of education (example using dummy data)*

| 306 (institution) | Reference Year | 313 (Student ID) | 461 (Field of education) | 462 (Field of education – for double major) | 464 (Discipline) | 46x (Modified dFOE) |
|---|---|---|---|---|---|---|
| 9999 | 2007 | ABC1234 | 070199 | 000000 | 070100 | 070103 |
| 9999 | 2007 | ABC1234 | 070199 | 000000 | 070103 | 070103 |
| 9999 | 2007 | ABC1234 | 070199 | 000000 | 070103 | 070103 |
| 9999 | 2007 | ABC1234 | 070199 | 000000 | 070100 | 070103 |
| 9999 | 2007 | ABC1234 | 070199 | 000000 | 070100 | 070103 |

When the discipline code at the unit level included more than one type in the same year (e.g. 070101 and 070103), then students were categorised as a mixed program (07010x).

The same rules were applied when variable 461 (Major 1) and 462 (Major 2) had a different detailed field of education (e.g. 461 = 070103 and 462 = 070105).

If both variables 461or 462 had a value of '070101', '070103' or '070105', then detailed field of education was not recoded.

**Table 2-16** shows the number of enrolments in each detailed field of education before and after recoding.

*Table 2-16: Number of enrolments in each detailed field of education (step 1)*

| | Before | | | | | After | | | | | Total N |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | EC | PRI | SEC | MIX | EO | EC | PRI | SEC | MIX | EO | |
| 2005 | 7427 | 25291 | 16750 | 0 | 13338 | 8019 | 28019 | 19732 | 1122 | 5914 | 62806 |
| 2006 | 7151 | 25348 | 16363 | 0 | 15633 | 7719 | 28240 | 19463 | 1436 | 7637 | 64495 |
| 2007 | 7901 | 25563 | 16082 | 0 | 18612 | 8506 | 29303 | 19459 | 1446 | 9444 | 68158 |
| 2008 | 8067 | 25626 | 16458 | 0 | 17482 | 8599 | 28695 | 19321 | 1235 | 9783 | 67633 |
| 2009 | 8694 | 25376 | 17302 | 0 | 17905 | 9157 | 28672 | 20022 | 1133 | 10293 | 69277 |
| 2010 | 9807 | 27041 | 18752 | 0 | 17194 | 10233 | 30446 | 21562 | 1229 | 9324 | 72794 |
| 2011 | 10416 | 28067 | 19997 | 0 | 15995 | 10775 | 30998 | 22740 | 1042 | 8920 | 74475 |
| 2012 | 11854 | 29448 | 20578 | 0 | 16320 | 12396 | 32148 | 23651 | 1095 | 8910 | 78200 |
| 2013 | 12137 | 28647 | 21813 | 0 | 17019 | 12791 | 31694 | 25050 | 1156 | 8925 | 79616 |
| 2014 | 13051 | 28427 | 23938 | 0 | 15970 | 13787 | 31437 | 27062 | 1357 | 7743 | 81386 |
| 2015 | 13938 | 31307 | 24279 | 0 | 15859 | 14842 | 34469 | 27308 | 1414 | 7350 | 85383 |
| 2016 | 13456 | 29997 | 25123 | 0 | 18555 | 14813 | 34321 | 28031 | 1921 | 8045 | 87131 |
| 2017 | 13695 | 33323 | 26307 | < 5 | 18762 | 15010 | 37780 | 29710 | 2352 | 7238 | 92090 |

* EC = early childhood; PRI = primary; SEC = secondary; MIX = mixed program; EO = education other

### 2.6.4 Recoding based on the program name – variable 308

Records were recoded when the variables 461 or 462 had a value of '070101', '070103', '070105' or '07010x' and the program name (variable 308) included the following string texts:

- 'EARLY' or 'CHILD'—recoded to 070101
- 'PRI' or 'R-7'—recoded to 070103
- 'SEC'—recoded to 070105
- 'MID', 'MYS', 'M,PR', 'M,SEC','K-10','K-12','P-10','P-12' or '1-10'—flagged as 07010x

If the student ended up with more than one program type in the same reporting year, they were assigned as mixed program (07010x). **Table 2-17** shows the number of enrolments in each detailed field of education before and after recoding.

*Table 2-17: Number of enrolments in each detailed field of education before and after update*

| | Before | | | | | After | | | | | Total N |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | EC | PRI | SEC | MIX | EO | EC | PRI | SEC | MIX | EO | |
| 2005 | 8019 | 28019 | 19732 | 1122 | 5914 | 8019 | 28164 | 20203 | 2105 | 4315 | 62806 |
| 2006 | 7719 | 28240 | 19463 | 1436 | 7637 | 7719 | 28380 | 20188 | 2767 | 5441 | 64495 |
| 2007 | 8506 | 29303 | 19459 | 1446 | 9444 | 8506 | 29364 | 20058 | 2849 | 7381 | 68158 |
| 2008 | 8599 | 28695 | 19321 | 1235 | 9783 | 8599 | 28773 | 19693 | 2792 | 7776 | 67633 |
| 2009 | 9157 | 28672 | 20022 | 1133 | 10293 | 9157 | 28775 | 20469 | 2702 | 8174 | 69277 |
| 2010 | 10233 | 30446 | 21562 | 1229 | 9324 | 10281 | 30581 | 22220 | 2626 | 7086 | 72794 |
| 2011 | 10775 | 30998 | 22740 | 1042 | 8920 | 10777 | 31210 | 23688 | 2335 | 6465 | 74475 |
| 2012 | 12396 | 32148 | 23651 | 1095 | 8910 | 12402 | 32467 | 24539 | 2509 | 6283 | 78200 |
| 2013 | 12791 | 31694 | 25050 | 1156 | 8925 | 12798 | 32252 | 26154 | 2487 | 5925 | 79616 |
| 2014 | 13787 | 31437 | 27062 | 1357 | 7743 | 13804 | 32116 | 28268 | 2806 | 4392 | 81386 |
| 2015 | 14842 | 34469 | 27308 | 1414 | 7350 | 14888 | 35182 | 28733 | 3284 | 3296 | 85383 |
| 2016 | 14813 | 34321 | 28031 | 1921 | 8045 | 14861 | 35101 | 29294 | 3818 | 4057 | 87131 |
| 2017 | 15010 | 37780 | 29710 | 2352 | 7238 | 15054 | 38460 | 31082 | 3969 | 3525 | 92090 |

* EC = early childhood; PRI = primary; SEC = secondary; MIX = mixed program; EO = education other

### 2.6.5 Recoding based on updated subject information

A final step involved the use of the updated subject discipline code (as described in Section 2.6) to provide more information about the type of program being studied.

**Table 2-18** shows the number of enrolments in each of the detailed field of education before and after recoding at final stage.

*Table 2-18: Number of enrolments in each detailed field of education (final stage)*

| | Before | | | | | After | | | | | Total N |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | EC | PRI | SEC | MIX | EO | EC | PRI | SEC | MIX | EO | |
| 2005 | 7427 | 25291 | 16750 | 0 | 13338 | 8019 | 28164 | 20203 | 2105 | 4315 | 62806 |
| 2006 | 7151 | 25348 | 16363 | 0 | 15633 | 7719 | 28380 | 20188 | 2767 | 5441 | 64495 |
| 2007 | 7901 | 25563 | 16082 | 0 | 18612 | 8506 | 29364 | 20058 | 2849 | 7381 | 68158 |
| 2008 | 8067 | 25626 | 16458 | 0 | 17482 | 8599 | 28773 | 19693 | 2792 | 7776 | 67633 |
| 2009 | 8694 | 25376 | 17302 | 0 | 17905 | 9157 | 28775 | 20469 | 2702 | 8174 | 69277 |
| 2010 | 9807 | 27041 | 18752 | 0 | 17194 | 10281 | 30581 | 22220 | 2626 | 7086 | 72794 |
| 2011 | 10416 | 28067 | 19997 | 0 | 15995 | 10777 | 31210 | 23688 | 2335 | 6465 | 74475 |
| 2012 | 11854 | 29448 | 20578 | 0 | 16320 | 12402 | 32467 | 24539 | 2509 | 6283 | 78200 |
| 2013 | 12137 | 28647 | 21813 | 0 | 17019 | 12798 | 32252 | 26154 | 2487 | 5925 | 79616 |
| 2014 | 13051 | 28427 | 23938 | 0 | 15970 | 13804 | 32116 | 28268 | 2806 | 4392 | 81386 |
| 2015 | 13938 | 31307 | 24279 | 0 | 15859 | 14888 | 35182 | 28733 | 3284 | 3296 | 85383 |
| 2016 | 13456 | 29997 | 25123 | 0 | 18555 | 14861 | 35101 | 29294 | 3818 | 4057 | 87131 |
| 2017 | 13695 | 33323 | 26307 | < 5 | 18762 | 15084 | 38488 | 31283 | 3969 | 3266 | 92090 |

* EC = early childhood; PRI = primary; SEC = secondary; MIX = mixed program; EO = education other

## 2.7 Longitudinal aggregation of data

This section discusses the logic applied to determine which categories students fell into over time, as their situation changed.

### 2.7.1 Determining start or end of course

Each student's data were summarised by the course(s) that the student was enrolled in each reporting year.

The start of the course was defined by the commencement indicator (1), and the end by the completion year or the commencement indicator (1) for the next course, if present. If commencement indicator and/or completion year were not present, then year-first-appeared and year-last-appeared data were used instead to allocate each student to the correct category for the reporting period.

### 2.7.2 Determining type or mode of attendance over time

When a student's enrolment changed from full-time to part-time or vice versa throughout their studies in any program, it was necessary to devise a method for capturing this information at the aggregate level.

A new category for type of attendance was created, called 'mixed type'. If a student completed both an undergraduate and postgraduate program, and they changed from full-time to part-time during the undergraduate program but completed the postgraduate program on a full-time basis, they would be assigned to 'mixed type' of attendance for the undergraduate program but full-time for the postgraduate program.

This is illustrated in the example in **Table 2-19** below. This approach ensures that, depending on the level of aggregation, students were categorised correctly to reflect their individual circumstances.

*Table 2-19: Summarising data by course (example using dummy data)*

| 306 (institution) | 313 (Student ID) | Reference year | G_LEVEL (Graduate level) | COURSE (Courses taken; alphabetical labelling) | 330 (Type of attendance; 1 = Full-time, 2 = Part-time) | 30_COU (330 by course; 3 = Mixed type) | 330_GLVL (330 by graduate level) | 330_PERSON (330 by individual) |
|---|---|---|---|---|---|---|---|---|
| 8888 | ABC5555 | 2007 | UG | A | 3 | 3 | 3 | 20 |
| 8888 | ABC5555 | 2008 | UG | A | 3 | 3 | 3 | 21 |
| 8888 | ABC5555 | 2009 | UG | A | 3 | 3 | 3 | 22 |
| 8888 | ABC5555 | 2010 | UG | A | 3 | 3 | 3 | 23 |
| 8888 | ABC5555 | 2011 | PG | B | 1 | 3 | 1 | 24 |
| 8888 | ABC5555 | 2012 | PG | B | 1 | 3 | 1 | 25 |

The same treatment was applied for students who changed mode of attendance from internal to external or vice versa.

### 2.7.3 Longitudinal aggregation of detailed field of education (dFOE)

Once dFOE recoding was completed as described in Section 2.6, logic had to be developed to determine dFOE at different levels of aggregation and over time.

The levels of aggregation that were considered included the following:

- Overall level for the individual by course level. This is the simplest level of aggregation, as each course enrolled in will have a single classification, therefore there is no need to aggregate information at this level.
- Overall level for the individual over the course of their entire enrolment in ITE. For the majority of students, who only ever complete a single ITE program, there will be a single classification of dFOE that applies to them—however, this may not be true for students who are enrolled in multiple programs over the course of their studies.
- Overall level for the individual by program level. For most students, they will only have ever completed one program of ITE study but the dFOE classification might vary for an individual that has enrolled in both undergraduate and postgraduate programs between 2005 and 2017.

**Table 2-20** shows the complicated enrolment history for an individual who:

- was enrolled and completed one undergraduate program at one institution between 2007 and 2010
- commenced a postgraduate program in 2009 and dropped out by the end of 2010 at a second institution
- commenced a postgraduate program in 2011 at the same institution
- commenced a second undergraduate program in 2012 at a second institution.

The dFOE value for the student varied. For example, if an individual studied early childhood education but did not complete, but undertook another course in secondary education and completed, then ‘SEC’ would be assigned instead of ‘MIX’.

*Table 2-20: Classification of detailed field of education (example using dummy data)*

| 306 (institution) | 313 (student ID) | Reference year | 922 (commencement indicator) | Completion year | G_LEVEL (graduate level) | COURSE (courses taken; alphabetical labelling) | 46x (modified dFOE) | CLASS_YR (dFOE by enrolment year) | CLASS_COU (dFOE by course) | CLASS_GLVL (dFOE by graduate level) | CLASS_PERSON (dFOE by individual) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 9999 | ABC1234 | 2012 | 1 | . | UG | A | 070103 | PRI | PRI | PRI | PRI |
| 9999 | ABC1234 | 2013 | 2 | . | UG | A | 079999 | EO | PRI | PRI | PRI |
| 9999 | ABC1234 | 2014 | 2 | . | UG | A | 079999 | EO | PRI | PRI | PRI |
| 8888 | ABC5555 | 2007 | 1 | . | UG | A | 070101 | EC | MIX | MIX | MIX |
| 8888 | ABC5555 | 2008 | 2 | . | UG | A | 070101 | EC | MIX | MIX | MIX |
| 8888 | ABC5555 | 2009 | 2 | . | UG | A | 070103 | PRI | MIX | MIX | MIX |
| 8888 | ABC5555 | 2010 | 2 | 2010 | UG | A | 070103 | PRI | MIX | MIX | MIX |
| 8888 | ABC5555 | 2011 | 1 | . | PG | B | 070103 | PRI | PRI | PRI | MIX |
| 8888 | ABC5555 | 2012 | 2 | 2012 | PG | B | 070103 | PRI | PRI | PRI | MIX |
| 7777 | ABC9999 | 2009 | 1 | . | PG | A | 070101 | EC | MIX | SEC | SEC |
| 7777 | ABC9999 | 2010 | 2 | . (dropped) | PG | A | 070103 | PRI | MIX | SEC | SEC |
| 7777 | ABC9999 | 2011 | 1 | . | PG | B | 070105 | SEC | SEC | SEC | SEC |
| 7777 | ABC9999 | 2012 | 2 | 2012 | PG | B | 070105 | SEC | SEC | SEC | SEC |

**Table 2-21** shows the number of enrolments in each detailed field of education before and after recoding.

Note that the data were summarised using dFOE by course (CLASS_COU—meaning that a student's program enrolment at any time is taken into account). This means that people who have enrolled in more than one program over time are likely to be allocated to the 'mixed program' category. This category has increased substantially following the recoding process. More importantly, the unknown category ('education other') has reduced over time. Students allocated to that category were subsequently reallocated into other fields of education (EC, PRI and SEC).

*Table 2-21: Before and after DFOE recoding*

| | Before | | | | | After | | | | | Total N |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | EC | PRI | SEC | MIX | EO | EC | PRI | SEC | MIX | EO | |
| 2005 | 7389 | 25412 | 16919 | 288 | 12798 | 7543 | 27403 | 20290 | 4455 | 3115 | 62806 |
| 2006 | 7117 | 25687 | 16763 | 506 | 14422 | 7323 | 27877 | 20105 | 5681 | 3509 | 64495 |
| 2007 | 7811 | 26266 | 17167 | 667 | 16247 | 8095 | 28536 | 20367 | 6339 | 4821 | 68158 |
| 2008 | 8103 | 26061 | 16867 | 763 | 15839 | 8170 | 28172 | 19735 | 6637 | 4919 | 67633 |
| 2009 | 8501 | 26139 | 17889 | 953 | 15795 | 8557 | 28213 | 20792 | 6567 | 5148 | 69277 |
| 2010 | 9610 | 27419 | 19128 | 951 | 15686 | 9684 | 29566 | 22263 | 6618 | 4663 | 72794 |
| 2011 | 10200 | 27903 | 19930 | 944 | 15498 | 10313 | 29717 | 23562 | 6592 | 4291 | 74475 |
| 2012 | 11657 | 29311 | 20616 | 966 | 15650 | 11783 | 31049 | 24648 | 6694 | 4026 | 78200 |
| 2013 | 11860 | 28835 | 22137 | 1049 | 15735 | 11988 | 30568 | 26408 | 7009 | 3643 | 79616 |
| 2014 | 12502 | 28630 | 24319 | 1184 | 14751 | 12670 | 30765 | 28301 | 7167 | 2483 | 81386 |
| 2015 | 13665 | 31098 | 24855 | 1237 | 14528 | 14044 | 33470 | 28655 | 7553 | 1661 | 85383 |
| 2016 | 13286 | 29814 | 25140 | 1100 | 17791 | 14111 | 33785 | 29210 | 7727 | 2298 | 87131 |
| 2017 | 13530 | 32885 | 26303 | 852 | 18520 | 14497 | 37406 | 31077 | 6826 | 2284 | 92090 |

* EC = early childhood; PRI = primary; SEC = secondary; MIX = mixed program; EO = education other

The numbers of students allocated in each detailed field of education is likely to change over time as more accurate data become available.

### 2.7.4 Classification of enrolled student by detailed field of education

Students are counted as enrolled for each year they have an enrolment record.

In the example highlighted in **Table 2-22**, the student has been counted as an enrolled student for each year between 2007 and 2012. The student would have been counted as enrolled in undergraduate programs for each year between 2007 and 2010 and enrolled in postgraduate programs in 2011 and 2012.

*Table 2-22: Enrolments by detailed field of education (example using dummy data)*

| 306 (institution) | 313 (student ID) | Reference year | 922 (commencement indicator) | Completion year | G_LEVEL (graduate level) | CLASS_YR (dFOE by enrolment year) | CLASS_COU (dFOE by course) | 330 (type of attendance; 1 = Full-time, 2 = Part-time) | 330_COU (330 by course; 3 = Mixed type) | Age |
|---|---|---|---|---|---|---|---|---|---|---|
| 8888 | ABC5555 | 2007 | 1 | . | UG | EC | MIX | 1 | 3 | 20 |
| 8888 | ABC5555 | 2008 | 2 | . | UG | EC | MIX | 1 | 3 | 21 |
| 8888 | ABC5555 | 2009 | 2 | . | UG | PRI | MIX | 2 | 3 | 22 |
| 8888 | ABC5555 | 2010 | 2 | 2010 | UG | PRI | MIX | 1 | 3 | 23 |
| 8888 | ABC5555 | 2011 | 1 | . | PG | EO | PRI | 1 | 1 | 24 |
| 8888 | ABC5555 | 2012 | 2 | 2012 | PG | PRI | PRI | 1 | 1 | 25 |

### 2.7.5 Classification of commencing student by detailed field of education

For the examples highlighted in **Table 2-23** below, the student would have been classified as commencing an early childhood program in 2007, and education other in 2011. Note that for this purpose, students are allocated to a particular dFOE on the basis of the program they are enrolled in. If enrolled in multiple programs over time, students will be allocated to the relevant dFOE that applies at that time.

*Table 2-23: Commencements (example using dummy data)*

| 306 (institution) | 313 (Student ID) | Reference year | 922 (commencement indicator) | Completion year | G_LEVEL (graduate level) | CLASS_YR (dFOE by enrolment year) | CLASS_COU (dFOE by course) | 330 (type of attendance; 1 = Full-time, 2 = Part-time) | 330_COU (330 by course; 3 = Mixed type) | Age |
|---|---|---|---|---|---|---|---|---|---|---|
| 8888 | ABC5555 | 2007 | 1 | . | UG | EC | MIX | 1 | 3 | 20 |
| 8888 | ABC5555 | 2008 | 2 | . | UG | EC | MIX | 1 | 3 | 21 |
| 8888 | ABC5555 | 2009 | 2 | . | UG | PRI | MIX | 2 | 3 | 22 |
| 8888 | ABC5555 | 2010 | 2 | 2010 | UG | PRI | MIX | 1 | 3 | 23 |
| 8888 | ABC5555 | 2011 | 1 | . | PG | EO | PRI | 1 | 1 | 24 |
| 8888 | ABC5555 | 2012 | 2 | 2012 | PG | PRI | PRI | 1 | 1 | 25 |

### 2.7.6 Classification of completing student by detailed field of education

For the examples highlighted in **Table 2-24** below, the student would have been classified as having completed a mixed education undergraduate program in 2010 and a primary education postgraduate program in 2012. As noted in the previous section, students are allocated to a particular dFOE on the basis of the program they are enrolled in. If enrolled in multiple programs over time, students will be allocated to the relevant dFOE that applies at that time.

*Table 2-24: Completions (examples using dummy data)*

| 306 (institution) | 313 (student ID) | Reference year | 922 (commencement indicator) | Completion year | G_LEVEL (graduate level) | CLASS_YR (dFOE by enrolment year) | CLASS_COU (dFOE by course) | 330 (type of attendance; 1 = Full-time, 2 = Part-time) | 330_COU (330 by course; 3 = Mixed type) | Age |
|---|---|---|---|---|---|---|---|---|---|---|
| 8888 | ABC5555 | 2007 | 1 | . | UG | EC | MIX | 1 | 3 | 20 |
| 8888 | ABC5555 | 2008 | 2 | . | UG | EC | MIX | 1 | 3 | 21 |
| 8888 | ABC5555 | 2009 | 2 | . | UG | PRI | MIX | 2 | 3 | 22 |
| 8888 | ABC5555 | 2010 | 2 | 2010 | UG | PRI | MIX | 1 | 3 | 23 |
| 8888 | ABC5555 | 2011 | 1 | . | PG | EO | PRI | 1 | 1 | 24 |
| 8888 | ABC5555 | 2012 | 2 | 2012 | PG | PRI | PRI | 1 | 1 | 25 |

## 2.8 First-year retention calculation

First- to second-year retention applies to undergraduate programs of more than two years duration. For this reason, only students enrolled in Bachelor Pass programs were selected for this analysis. The Bachelor Pass (BPASS) programs are the most common, comprising 90% of all undergraduate students enrolled in ITE during 2017.

For a given commencement year, the retention rate was calculated as the number of students who were still enrolled (i.e. retained) in the following year. A further check was made on the enrolment type (i.e. whether they were flagged as a continuing student during the second year of the program).

The retention rate was calculated by dividing the total number of students continuing in the second year by the number of students who commenced the prior year. Individuals who completed on the same year they commenced the program were excluded.

## 2.9 Student residence vs study location

The analysis in the Pipeline Report compares the state or territory of the home residence of each student with the state or territory location of the ITE program in which the student is enrolled.

The state of the provider was obtained from the HESDC data element *Appendix A : Higher Education, VET & PIR provider codes and names*.

For students studying at a provider with multiple campus locations, it is not possible to determine the campus the students are attending from the information available in the HESDC course file. In the absence of other data, the students are attributed to the group studying in their home state of residence.

## 2.10 Determining commencement and completion year

The data does not distinguish between start-of-year commencement, therefore the time to completion maybe overstated by 6–12 months in some cases.

Student cohorts are grouped by year of commencement to allow for direct comparison of completion rates. Students are identified as commencing students via the field 'commencing student indicator' which is designated as '1' for a commencing student and '2' for a student who is enrolled as a continuing student in a program of study.

Students are identified as completing students when they have the field 'completion year' populated with a valid year. Data interrogation has shown that some students 'appear' in the longitudinal data between 2005 and 2017 without a commencing student indicator of '1'. Similarly, some students are enrolled in the requisite number of subjects to complete an ITE program, but then 'disappear' without a completion year populated.

Analysis of the QILT and HESDC data showed that there was a small number of survey respondents (n=47) who did not have a completion year populated, but who had completed the Graduate Outcomes Survey.

It is therefore likely that the number of commencing students and completing students is understated in the following analysis. Further investigation will be required to determine the extent to which this occurs once access to teacher workforce data is available—it will be possible to identify the students who did not complete as denoted by the absence of completion year data in the ITE data, but who are registered teachers.

## 2.11 Statistics

Once all the above data transformations were complete, all derived statistics—e.g. total number of students, mean age, median ATAR—were derived using standard descriptive statistical functions available in SAS Enterprise Guide software for each level of aggregated data.

# Appendix 1: Reference tables used

1. Australian Standard Classification of Education (ASCED) and Coder, 2001 (https://www.abs.gov.au/ausstats/abs@.nsf/0/E7779A9FD5C8D846CA256AAF001FCA5C)
2. Higher education, VET & PIR provider codes and names (https://heimshelp.education.gov.au/resources/appendices#AppendixA)
3. Standard Australian Classification of Countries (SACC), 2016 (https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1269.02016?OpenDocument)
4. Australian and New Zealand Standard Industrial Classification (ANZSIC), 2006 (Revision 2.0) (https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1292.02006%20(Revision%202.0)?OpenDocument)

# Appendix 2: Modified subject grouping table

| ASCED / 464 | Detailed fields | Grouping in Pipeline Report (Subjects) |
|---|---|---|
| 010000 | NATURAL AND PHYSICAL SCIENCES | Mathematical Sciences |
| 010100 | Mathematical Sciences | Mathematics |
| 010101 | Mathematics | Mathematics |
| 010103 | Statistics | Mathematics |
| 010199 | Mathematical Sciences, n.e.c. | Mathematics |
| 010300 | Physics and Astronomy | Physics |
| 010301 | Physics | Physics |
| 010303 | Astronomy | Physics |
| 010500 | Chemical Sciences | Chemistry |
| 010501 | Organic Chemistry | Chemistry |
| 010503 | Inorganic Chemistry | Chemistry |
| 010599 | Chemical Sciences, n.e.c. | Chemistry |
| 010700 | Earth Sciences | Earth and Environmental Science |
| 010701 | Atmospheric Sciences | Earth and Environmental Science |
| 010703 | Geology | Earth and Environmental Science |
| 010705 | Geophysics | Earth and Environmental Science |
| 010707 | Geochemistry | Earth and Environmental Science |
| 010709 | Soil Science | Earth and Environmental Science |
| 010711 | Hydrology | Earth and Environmental Science |
| 010713 | Oceanography | Earth and Environmental Science |
| 010799 | Earth Sciences, n.e.c. | Earth and Environmental Science |
| 010900 | Biological Sciences | Biology |
| 010901 | Biochemistry and Cell Biology | Biology |
| 010903 | Botany | Biology |
| 010905 | Ecology and Evolution | Biology |
| 010907 | Marine Science | Biology |
| 010909 | Genetics | Biology |
| 010911 | Microbiology | Biology |
| 010913 | Human Biology | Biology |
| 010915 | Zoology | Biology |
| 010999 | Biological Sciences, n.e.c. | Biology |
| 019900 | Other Natural and Physical Sciences | Other Science |
| 019901 | Medical Science | Other Science |
| 019903 | Forensic Science | Other Science |
| 019905 | Food Science and Biotechnology | Food Technology |
| 019907 | Pharmacology | Other Science |

| | | |
|---|---|---|
| 019909 | Laboratory Technology | Other Science |
| 019999 | Natural and Physical Sciences, n.e.c. | Other Science |
| 020000 | INFORMATION TECHNOLOGY | ICT |
| 020100 | Computer Science | ICT |
| 020101 | Formal Language Theory | ICT |
| 020103 | Programming | ICT |
| 020105 | Computational Theory | ICT |
| 020107 | Compiler Construction | ICT |
| 020109 | Algorithms | ICT |
| 020111 | Data Structures | ICT |
| 020113 | Networks and Communications | ICT |
| 020115 | Computer Graphics | ICT |
| 020117 | Operating Systems | ICT |
| 020119 | Artificial Intelligence | ICT |
| 020199 | Computer Science, n.e.c. | ICT |
| 020300 | Information Systems | ICT |
| 020301 | Conceptual Modelling | ICT |
| 020303 | Database Management | ICT |
| 020305 | Systems Analysis and Design | ICT |
| 020307 | Decision Support Systems | ICT |
| 020399 | Information Systems, n.e.c. | ICT |
| 029900 | Other Information Technology | ICT |
| 029901 | Security Science | ICT |
| 029999 | Information Technology, n.e.c. | ICT |
| 030000 | ENGINEERING AND RELATED TECHNOLOGIES | Engineering |
| 030100 | Manufacturing Engineering and Technology | Engineering |
| 030101 | Manufacturing Engineering | Engineering |
| 030103 | Printing | Engineering |
| 030105 | Textile Making | Engineering |
| 030107 | Garment Making | Engineering |
| 030109 | Footwear Making | Engineering |
| 030111 | Wood Machining and Turning | Engineering |
| 030113 | Cabinet Making | Engineering |
| 030115 | Furniture Upholstery and Renovation | Engineering |
| 030117 | Furniture Polishing | Engineering |
| 030199 | Manufacturing Engineering and Technology, n.e.c. | Engineering |
| 030300 | Process and Resources Engineering | Engineering |
| 030301 | Chemical Engineering | Engineering |
| 030303 | Mining Engineering | Engineering |

| | | |
|---|---|---|
| 030305 | Materials Engineering | Engineering |
| 030307 | Food Processing Technology | Engineering |
| 030399 | Process and Resources Engineering, n.e.c. | Engineering |
| 030500 | Automotive Engineering and Technology | Engineering |
| 030501 | Automotive Engineering | Engineering |
| 030503 | Vehicle Mechanics | Engineering |
| 030505 | Automotive Electrics and Electronics | Engineering |
| 030507 | Automotive Vehicle Refinishing | Engineering |
| 030509 | Automotive Body Construction | Engineering |
| 030511 | Panel Beating | Engineering |
| 030513 | Upholstery and Vehicle Trimming | Engineering |
| 030515 | Automotive Vehicle Operations | Engineering |
| 030599 | Automotive Engineering and Technology, n.e.c. | Engineering |
| 030700 | Mechanical and Industrial Engineering and Technology | Engineering |
| 030701 | Mechanical Engineering | Engineering |
| 030703 | Industrial Engineering | Engineering |
| 030705 | Toolmaking | Engineering |
| 030707 | Metal Fitting, Turning and Machining | Engineering |
| 030709 | Sheetmetal Working | Engineering |
| 030711 | Boilermaking and Welding | Engineering |
| 030713 | Metal Casting and Patternmaking | Engineering |
| 030715 | Precision Metalworking | Engineering |
| 030717 | Plant and Machine Operations | Engineering |
| 030799 | Mechanical and Industrial Engineering and Technology, n.e.c. | Engineering |
| 030900 | Civil Engineering | Engineering |
| 030901 | Construction Engineering | Engineering |
| 030903 | Structural Engineering | Engineering |
| 030905 | Building Services Engineering | Engineering |
| 030907 | Water and Sanitary Engineering | Engineering |
| 030909 | Transport Engineering | Engineering |
| 030911 | Geotechnical Engineering | Engineering |
| 030913 | Ocean Engineering | Engineering |
| 030999 | Civil Engineering, n.e.c. | Engineering |
| 031100 | Geomatic Engineering | Engineering |
| 031101 | Surveying | Engineering |
| 031103 | Mapping Science | Engineering |
| 031199 | Geomatic Engineering, n.e.c. | Engineering |
| 031300 | Electrical and Electronic Engineering and Technology | Engineering |
| 031301 | Electrical Engineering | Engineering |

| | | |
|---|---|---|
| 031303 | Electronic Engineering | Engineering |
| 031305 | Computer Engineering | Engineering |
| 031307 | Communications Technologies | Engineering |
| 031309 | Communications Equipment Installation and Maintenance | Engineering |
| 031311 | Powerline Installation and Maintenance | Engineering |
| 031313 | Electrical Fitting, Electrical Mechanics | Engineering |
| 031315 | Refrigeration and Air Conditioning Mechanics | Engineering |
| 031317 | Electronic Equipment Servicing | Engineering |
| 031399 | Electrical and Electronic Engineering and Technology, n.e.c. | Engineering |
| 031500 | Aerospace Engineering and Technology | Engineering |
| 031501 | Aerospace Engineering | Engineering |
| 031503 | Aircraft Maintenance Engineering | Engineering |
| 031505 | Aircraft Operation | Engineering |
| 031507 | Air Traffic Control | Engineering |
| 031599 | Aerospace Engineering and Technology, n.e.c. | Engineering |
| 031700 | Maritime Engineering and Technology | Engineering |
| 031701 | Maritime Engineering | Engineering |
| 031703 | Marine Construction | Engineering |
| 031705 | Marine Craft Operation | Engineering |
| 031799 | Maritime Engineering and Technology, n.e.c. | Engineering |
| 039900 | Other Engineering and Related Technologies | Engineering |
| 039901 | Environmental Engineering | Engineering |
| 039903 | Biomedical Engineering | Engineering |
| 039905 | Fire Technology | Engineering |
| 039907 | Rail Operations | Engineering |
| 039909 | Cleaning | Engineering |
| 039999 | Engineering and Related Technologies, n.e.c. | Engineering |
| 040000 | ARCHITECTURE AND BUILDING | Architecture |
| 040100 | Architecture and Urban Environment | Architecture |
| 040101 | Architecture | Architecture |
| 040103 | Urban Design and Regional Planning | Architecture |
| 040105 | Landscape Architecture | Architecture |
| 040107 | Interior and Environmental Design | Architecture |
| 040199 | Architecture and Urban Environment, n.e.c. | Architecture |
| 040300 | Building | Building |
| 040301 | Building Science and Technology | Building |
| 040303 | Building Construction Management | Building |
| 040305 | Building Surveying | Building |
| 040307 | Building Construction Economics | Building |

| | | |
|---|---|---|
| 040309 | Bricklaying and Stonemasonry | Building |
| 040311 | Carpentry and Joinery | Building |
| 040313 | Ceiling, Wall and Floor Fixing | Building |
| 040315 | Roof Fixing | Building |
| 040317 | Plastering | Building |
| 040319 | Furnishing Installation | Building |
| 040321 | Floor Coverings | Building |
| 040323 | Glazing | Building |
| 040325 | Painting, Decorating and Sign Writing | Building |
| 040327 | Plumbing | Building |
| 040329 | Scaffolding and Rigging | Building |
| 040399 | Building, n.e.c. | Building |
| 050000 | AGRICULTURE, ENVIRONMENTAL AND RELATED STUDIES | Agriculture and Environmental Studies |
| 050100 | Agriculture | Agriculture and Environmental Studies |
| 050101 | Agricultural Science | Agriculture and Environmental Studies |
| 050103 | Wool Science | Agriculture and Environmental Studies |
| 050105 | Animal Husbandry | Agriculture and Environmental Studies |
| 050199 | Agriculture, n.e.c. | Agriculture and Environmental Studies |
| 050300 | Horticulture and Viticulture | Agriculture and Environmental Studies |
| 050301 | Horticulture | Agriculture and Environmental Studies |
| 050303 | Viticulture | Agriculture and Environmental Studies |
| 050500 | Forestry Studies | Agriculture and Environmental Studies |
| 050501 | Forestry Studies | Agriculture and Environmental Studies |
| 050700 | Fisheries Studies | Agriculture and Environmental Studies |
| 050701 | Aquaculture | Agriculture and Environmental Studies |
| 050799 | Fisheries Studies, n.e.c. | Agriculture and Environmental Studies |
| 050900 | Environmental Studies | Agriculture and Environmental Studies |
| 050901 | Land, Parks and Wildlife Management | Agriculture and Environmental Studies |
| 050999 | Environmental Studies, n.e.c. | Agriculture and Environmental Studies |
| 059900 | Other Agriculture, Environmental and Related Studies | Agriculture and Environmental Studies |
| 059901 | Pest and Weed Control | Agriculture and Environmental Studies |
| 059999 | Agriculture, Environmental and Related Studies, n.e.c. | Agriculture and Environmental Studies |
| 060000 | HEALTH | Medicine |
| 060100 | Medical Studies | Medicine |
| 060101 | General Medicine | Medicine |
| 060103 | Surgery | Medicine |
| 060105 | Psychiatry | Medicine |
| 060107 | Obstetrics and Gynaecology | Medicine |
| 060109 | Paediatrics | Medicine |

| | | |
|---|---|---|
| 060111 | Anaesthesiology | Medicine |
| 060113 | Pathology | Medicine |
| 060115 | Radiology | Medicine |
| 060117 | Internal Medicine | Medicine |
| 060119 | General Practice | Medicine |
| 060199 | Medical Studies, n.e.c. | Medicine |
| 060300 | Nursing | Nursing |
| 060301 | General Nursing | Nursing |
| 060303 | Midwifery | Nursing |
| 060305 | Mental Health Nursing | Nursing |
| 060307 | Community Nursing | Nursing |
| 060309 | Critical Care Nursing | Nursing |
| 060311 | Aged Care Nursing | Nursing |
| 060313 | Palliative Care Nursing | Nursing |
| 060315 | Mothercraft Nursing and Family and Child Health Nursing | Nursing |
| 060399 | Nursing, n.e.c. | Nursing |
| 060500 | Pharmacy | Pharmacy |
| 060501 | Pharmacy | Pharmacy |
| 060700 | Dental Studies | Health |
| 060701 | Dentistry | Dentistry |
| 060703 | Dental Assisting | Health |
| 060705 | Dental Technology | Health |
| 060799 | Dental Studies, n.e.c. | Health |
| 060900 | Optical Science | Science |
| 060901 | Optometry | Science |
| 060903 | Optical Technology | Health |
| 060999 | Optical Science, n.e.c. | Health |
| 061100 | Veterinary Studies | Science |
| 061101 | Veterinary Science | Science |
| 061103 | Veterinary Assisting | Health |
| 061199 | Veterinary Studies, n.e.c. | Health |
| 061300 | Public Health | Health |
| 061301 | Occupational Health and Safety | Health |
| 061303 | Environmental Health | Health |
| 061305 | Indigenous Health | Health |
| 061307 | Health Promotion | Health |
| 061309 | Community Health | Health |
| 061311 | Epidemiology | Health |
| 061399 | Public Health, n.e.c. | Health |

| | | |
|---|---|---|
| 061500 | Radiography | Health |
| 061501 | Radiography | Health |
| 061700 | Rehabilitation Therapies | Health |
| 061701 | Physiotherapy | Health |
| 061703 | Occupational Therapy | Health |
| 061705 | Chiropractic and Osteopathy | Health |
| 061707 | Speech Pathology | Health |
| 061709 | Audiology | Health |
| 061711 | Massage Therapy | Health |
| 061713 | Podiatry | Health |
| 061799 | Rehabilitation Therapies, n.e.c. | Health |
| 061900 | Complementary Therapies | Health |
| 061901 | Naturopathy | Health |
| 061903 | Acupuncture | Health |
| 061905 | Traditional Chinese Medicine | Health |
| 061999 | Complementary Therapies, n.e.c. | Health |
| 069900 | Other Health | Nutrition |
| 069901 | Nutrition and Dietetics | Nutrition |
| 069903 | Human Movement | Health |
| 069905 | Paramedical Studies | Health |
| 069907 | First Aid | Health |
| 069999 | Health, n.e.c. | Health |
| 070000 | EDUCATION | Education Other |
| 070100 | Teacher Education | Education Other |
| 070101 | Teacher Education: Early Childhood | Early Childhood |
| 070103 | Teacher Education: Primary | Primary |
| 070105 | Teacher Education: Secondary | Secondary |
| 070107 | Teacher-Librarianship | Librarianship |
| 070109 | Teacher Education: Vocational Education and Training | VET |
| 070111 | Teacher Education: Higher Education | Higher Education |
| 070113 | Teacher Education: Special Education | Special Education |
| 070115 | English as a Second Language Teaching | English ALD |
| 070117 | Nursing Education Teacher Training | Nursing |
| 070199 | Teacher Education, n.e.c. | Education Other |
| 070300 | Curriculum and Education Studies | Education Other |
| 070301 | Curriculum Studies | Education Other |
| 070303 | Education Studies | Education Other |
| 079900 | Other Education | Education Other |
| 079999 | Education, n.e.c. | Education Other |

| | | |
|---|---|---|
| 080000 | MANAGEMENT AND COMMERCE | Accounting |
| 080100 | Accounting | Accounting |
| 080101 | Accounting | Accounting |
| 080300 | Business and Management | Business and Management |
| 080301 | Business Management | Business and Management |
| 080303 | Human Resource Management | Business and Management |
| 080305 | Personal Management Training | Business and Management |
| 080307 | Organisation Management | Business and Management |
| 080309 | Industrial Relations | Business and Management |
| 080311 | International Business | Business and Management |
| 080313 | Public and Health Care Administration | Business and Management |
| 080315 | Project Management | Business and Management |
| 080317 | Quality Management | Business and Management |
| 080319 | Hospitality Management | Business and Management |
| 080321 | Farm Management and Agribusiness | Business and Management |
| 080323 | Tourism Management | Business and Management |
| 080399 | Business and Management, n.e.c. | Business and Management |
| 080500 | Sales and Marketing | Sales and Marketing |
| 080501 | Sales | Sales and Marketing |
| 080503 | Real Estate | Sales and Marketing |
| 080505 | Marketing | Sales and Marketing |
| 080507 | Advertising | Sales and Marketing |
| 080509 | Public Relations | Sales and Marketing |
| 080599 | Sales and Marketing, n.e.c. | Sales and Marketing |
| 080700 | Tourism | Tourism |
| 080701 | Tourism | Tourism |
| 080900 | Office Studies | Office Studies |
| 080901 | Secretarial and Clerical Studies | Office Studies |
| 080903 | Keyboard Skills | Office Studies |
| 080905 | Practical Computing Skills | Office Studies |
| 080999 | Office Studies, n.e.c. | Office Studies |
| 081100 | Banking, Finance and Related Fields | Banking, Finance and Related Fields |
| 081101 | Banking and Finance | Banking, Finance and Related Fields |
| 081103 | Insurance and Actuarial Studies | Banking, Finance and Related Fields |
| 081105 | Investment and Securities | Banking, Finance and Related Fields |
| 081199 | Banking, Finance and Related Fields, n.e.c. | Banking, Finance and Related Fields |
| 089900 | Other Management and Commerce | Other Management and Commerce |
| 089901 | Purchasing, Warehousing and Distribution | Other Management and Commerce |
| 089903 | Valuation | Other Management and Commerce |

| | | |
|---|---|---|
| 089999 | Management and Commerce, n.e.c. | Other Management and Commerce |
| 090000 | SOCIETY AND CULTURE | Humanities |
| 090100 | Political Science and Policy Studies | Humanities |
| 090101 | Political Science | Politics |
| 090103 | Policy Studies | Politics |
| 090300 | Studies in Human Society | Humanities |
| 090301 | Sociology | Humanities |
| 090303 | Anthropology | Humanities |
| 090305 | History | History |
| 090307 | Archaeology | Humanities |
| 090309 | Human Geography | Geography |
| 090311 | Indigenous Studies | Indigenous Studies |
| 090313 | Gender Specific Studies | Humanities |
| 090399 | Studies in Human Society, n.e.c. | Humanities |
| 090500 | Human Welfare Studies and Services | Humanities |
| 090501 | Social Work | Humanities |
| 090503 | Children's Services | Humanities |
| 090505 | Youth Work | Humanities |
| 090507 | Care for the Aged | Humanities |
| 090509 | Care for the Disabled | Humanities |
| 090511 | Residential Client Care | Humanities |
| 090513 | Counselling | Humanities |
| 090515 | Welfare Studies | Humanities |
| 090599 | Human Welfare Studies and Services, n.e.c. | Humanities |
| 090700 | Behavioural Science | Psychology |
| 090701 | Psychology | Psychology |
| 090799 | Behavioural Science, n.e.c. | Psychology |
| 090900 | Law | Law |
| 090901 | Business and Commercial Law | Law |
| 090903 | Constitutional Law | Law |
| 090905 | Criminal Law | Law |
| 090907 | Family Law | Law |
| 090909 | International Law | Law |
| 090911 | Taxation Law | Law |
| 090913 | Legal Practice | Law |
| 090999 | Law, n.e.c. | Law |
| 091100 | Justice and Law Enforcement | Law |
| 091101 | Justice Administration | Law |
| 091103 | Legal Studies | Law |

| | | |
|---|---|---|
| 091105 | Police Studies | Law |
| 091199 | Justice and Law Enforcement, n.e.c. | Law |
| 091300 | Librarianship, Information Management and Curatorial Studies | Librarianship |
| 091301 | Librarianship and Information Management | Librarianship |
| 091303 | Curatorial Studies | Librarianship |
| 091500 | Language and Literature | Language and Literature |
| 091501 | English Language | English |
| 091503 | Northern European Languages | LOTE |
| 091505 | Southern European Languages | LOTE |
| 091507 | Eastern European Languages | LOTE |
| 091509 | Southwest Asian and North African Languages | LOTE |
| 091511 | Southern Asian Languages | LOTE |
| 091513 | Southeast Asian Languages | LOTE |
| 091515 | Eastern Asian Languages | LOTE |
| 091517 | Australian Indigenous Languages | Australian Indigenous Languages |
| 091519 | Translating and Interpreting | LOTE |
| 091521 | Linguistics | English |
| 091523 | Literature | English |
| 091599 | Language and Literature, n.e.c. | English |
| 091700 | Philosophy and Religious Studies | Philosophy and Religious Studies |
| 091701 | Philosophy | Philosophy |
| 091703 | Religious Studies | Religious Studies |
| 091900 | Economics and Econometrics | Economics |
| 091901 | Economics | Economics |
| 091903 | Econometrics | Economics |
| 092100 | Sport and Recreation | PE |
| 092101 | Sport and Recreation Activities | PE |
| 092103 | Sports Coaching, Officiating and Instruction | PE |
| 092199 | Sport and Recreation, n.e.c. | PE |
| 099900 | Other Society and Culture | Humanities |
| 099901 | Family and Consumer Studies | Humanities |
| 099903 | Criminology | Humanities |
| 099905 | Security Services | Humanities |
| 099999 | Society and Culture, n.e.c. | Humanities |
| 100000 | CREATIVE ARTS | Creative Arts |
| 100100 | Performing Arts | Creative Arts |
| 100101 | Music | Music |
| 100103 | Drama and Theatre Studies | Drama |
| 100105 | Dance | Dance |

| | | |
|---|---|---|
| 100199 | Performing Arts, n.e.c. | Creative Arts |
| 100300 | Visual Arts and Crafts | Creative Arts |
| 100301 | Fine Arts | Fine Arts |
| 100303 | Photography | Creative Arts |
| 100305 | Crafts | Creative Arts |
| 100307 | Jewellery Making | Creative Arts |
| 100309 | Floristry | Creative Arts |
| 100399 | Visual Arts and Crafts, n.e.c. | Creative Arts |
| 100500 | Graphic and Design Studies | Creative Arts |
| 100501 | Graphic Arts and Design Studies | Creative Arts |
| 100503 | Textile Design | Creative Arts |
| 100505 | Fashion Design | Creative Arts |
| 100599 | Graphic and Design Studies, n.e.c. | Creative Arts |
| 100700 | Communication and Media Studies | Creative Arts |
| 100701 | Audio Visual Studies | Creative Arts |
| 100703 | Journalism | Journalism |
| 100705 | Written Communication | Creative Arts |
| 100707 | Verbal Communication | Creative Arts |
| 100799 | Communication and Media Studies, n.e.c. | Creative Arts |
| 109900 | Other Creative Arts | Creative Arts |
| 109999 | Creative Arts, n.e.c. | Creative Arts |
| 110000 | FOOD, HOSPITALITY AND PERSONAL SERVICES | Food and Hospitality |
| 110100 | Food and Hospitality | Food and Hospitality |
| 110101 | Hospitality | Food and Hospitality |
| 110103 | Food and Beverage Service | Food and Hospitality |
| 110105 | Butchery | Food and Hospitality |
| 110107 | Baking and Pastrymaking | Food and Hospitality |
| 110109 | Cookery | Food and Hospitality |
| 110111 | Food Hygiene | Food and Hospitality |
| 110199 | Food and Hospitality, n.e.c. | Food and Hospitality |
| 110300 | Personal Services | Personal Services |
| 110301 | Beauty Therapy | Personal Services |
| 110303 | Hairdressing | Personal Services |
| 110399 | Personal Services, n.e.c. | Personal Services |
| 120000 | MIXED FIELD PROGRAMMES | General Education Programmes |
| 120100 | General Education Programmes | General Education Programmes |
| 120101 | General Primary and Secondary Education Programmes | General Education Programmes |
| 120103 | Literacy and Numeracy Programmes | Literacy and Numeracy |
| 120105 | Learning Skills Programmes | General Education Programmes |

| | | |
|---|---|---|
| 120199 | General Education Programmes, n.e.c. | General Education Programmes |
| 120300 | Social Skills Programmes | Social Skills Programmes |
| 120301 | Social and Interpersonal Skills Programmes | Social Skills Programmes |
| 120303 | Survival Skills Programmes | Social Skills Programmes |
| 120305 | Parental Education Programmes | Social Skills Programmes |
| 120399 | Social Skills Programmes, n.e.c. | Social Skills Programmes |
| 120500 | Employment Skills Programmes | Employment Skills Programmes |
| 120501 | Career Development Programmes | Employment Skills Programmes |
| 120503 | Job Search Skills Programmes | Employment Skills Programmes |
| 120505 | Work Practices Programmes | Employment Skills Programmes |
| 120599 | Employment Skills Programmes, n.e.c. | Employment Skills Programmes |
| 129900 | Other Mixed Field Programmes | Other Mixed Field Programmes |
| 129999 | Mixed Field Programmes, n.e.c. | Other Mixed Field Programmes |

# Appendix 3: Acronyms used

| | |
|---|---|
| ABS | Australian Bureau of Statistics |
| ASCED | Australian Standard Classification of Education |
| ATAR | Australian Tertiary Admission Rank |
| ATSI | Aboriginal and Torres Strait Islander |
| ATWD | Australian Teacher Workforce Data |
| CHESSN | Commonwealth Higher Education Student Support Number |
| dFOE | Detailed Field of Education |
| ESS | Employer Satisfaction Survey |
| GOS | Graduate Outcomes Survey |
| HECS | Higher Education Contribution Scheme |
| HEIMS | Higher Education Information Management System |
| HEPs | Higher Education Providers |
| HESDC | Higher Education Student Data Collection |
| ITE | Initial Teacher Education |
| PG | Postgraduate |
| QILT | Quality Indicators for Learning and Teaching |
| SAS | Statistical Analysis Software |
| SES | Student Experience Survey |
| UG | Undergraduate |

PO Box 299,
Collins Street West, VIC 8007

Ph: 03 9944 1267

**atwd@aitsl.edu.au**
**www.aitsl.edu.au/atwd**

ABN 17 117 362 740