



Australian Government

**Australian Institute of
Health and Welfare**

Development of a national education and training data standards strategy and implementation plan



Australian Government

**Australian Institute of
Health and Welfare**

*Authoritative information and statistics
to promote better health and wellbeing*

Development of a national education and training data standards strategy and implementation plan

Australian Institute of Health and Welfare

Canberra

Cat. no. EDU 4

The Australian Institute of Health and Welfare is a major national agency which provides reliable, regular and relevant information and statistics on Australia's health and welfare. The Institute's mission is authoritative information and statistics to promote better health and wellbeing.

© Australian Institute of Health and Welfare 2015



This product, excluding the AIHW logo, Commonwealth Coat of Arms and any material owned by a third party or protected by a trademark, has been released under a Creative Commons BY 3.0 (CC-BY 3.0) licence. Excluded material owned by third parties may include, for example, design and layout, images obtained under licence from third parties and signatures. We have made all reasonable efforts to identify and label material owned by third parties.

You may distribute, remix and build upon this work. However, you must attribute the AIHW as the copyright holder of the work in compliance with our attribution policy available at <www.aihw.gov.au/copyright/>. The full terms and conditions of this licence are available at <<http://creativecommons.org/licenses/by/3.0/au/>>.

Enquiries relating to copyright should be addressed to the Head of the Digital and Media Communications Unit, Australian Institute of Health and Welfare, GPO Box 570, Canberra ACT 2601.

A complete list of the Institute's publications is available from the Institute's website <www.aihw.gov.au>.

ISBN 978-1-74249-731-0 (PDF)

ISBN 978-1-74249-732-7 (Print)

Suggested citation

Australian Institute of Health and Welfare 2015. Development of a national education and training data standards strategy and implementation plan. Cat. no. EDU 4. Canberra: AIHW.

Australian Institute of Health and Welfare

Board Chair
Dr Mukesh C Haikerwal AO

Acting Director
Ms Kerry Flanagan PSM

Any enquiries about or comments on this publication should be directed to:

Digital and Media Communications Unit
Australian Institute of Health and Welfare
GPO Box 570
Canberra ACT 2601
Tel: (02) 6244 1000
Email: info@aihw.gov.au

Published by the Australian Institute of Health and Welfare

This publication is printed in accordance with ISO 14001 (Environmental Management Systems) and ISO 9001 (Quality Management Systems). The paper is sourced from sustainably managed certified forests.



Please note that there is the potential for minor revisions of data in this report. Please check the online version at <www.aihw.gov.au> for any amendments.

Contents

- Acknowledgments..... iv**
- Abbreviations..... v**
- Summary viii**
- 1 Introduction.....1**
- 2 Establishment of scope (Phase 1)3**
 - 2.1 Review of preliminary work3
 - 2.2 Stakeholder communication and engagement strategy4
 - 2.3 Scope of the national data standards strategy4
- 3 Review of data collections (Phase 2).....8**
 - 3.1 Data mapping8
 - 3.2 Questionnaires.....13
 - 3.3 Results.....14
 - 3.4 Priority data items.....16
- 4 Development of the national data standards strategy (Phase 3)18**
 - 4.1 National data standards strategy and implementation plan.....18
 - 4.2 Assessment of the comparability and consistency of core data items with other sectors29
 - 4.3 Preliminary scoping of the feasibility of data linkage30
 - 4.4 Project finalisation.....32
- Appendix A: Stakeholder consultation.....33**
- Appendix B: Data mapping.....35**
- Appendix C: Description of data collections and data items mapped.....88**
- Appendix D: Questionnaires151**
- Appendix E: Stakeholder and data custodian consultation159**
- Appendix F: Summary of policy priorities162**
- Appendix G: Data collection review timeframes and processes.....166**
- Appendix H: Data flows.....179**
- References186**
- List of tables190**
- List of figures191**

Acknowledgments

The Australian Institute of Health and Welfare (AIHW) would like to acknowledge funding provided under the 2013–14 Transforming Education and Training Information in Australia (TETIA) Forward Work Program to undertake this project.

The AIHW gratefully acknowledges members of the Data Strategy Group and the former Strategic Cross-sectoral Data Committee for their contributions to the project, as well as the data custodians and stakeholders who provided valuable input.

The project was conducted by the AIHW Maternal Health, Children, Youth and Families Unit. The project team thanks Conan Liu and Fadwa Al-Yaman from the Indigenous and Children's Group, AIHW, for providing assistance and guidance.

Abbreviations

ABS	Australian Bureau of Statistics
ACARA	Australian Curriculum, Assessment and Reporting Authority
ACER	Australian Council for Educational Research
AEDC	Australian Early Development Census (formerly the Australian Early Development Index)
AIHW	Australian Institute of Health and Welfare
ANZSCO	Australian and New Zealand Standard Classification of Occupations
AQF	Australian Qualifications Framework
ARIA	Accessibility/Remoteness Index of Australia
ASCED	Australian Standard Classification of Education
ASCL	Australian Standard Classification of Languages
ASGC	Australian Standard Geographical Classification
ASGS	Australian Statistical Geography Standard
AVETMISS	Australian Vocational Education and Training Management Information Statistical Standard
CD	Collection District
COB	country of birth
COAG	Council of Australian Governments
DOB	date of birth
DSG	Data Strategy Group
ECD RDS	Early Childhood Development Researchable Data Set
ECEC	early childhood education and care
FOE	field of education
HECS	Higher Education Contribution Scheme
HEIMS	Higher Education Information Management System
HELP	Higher Education Loan Program
ICT	information and communication technology
ID	identifier

IEA	International Association for the Evaluation of Educational Achievement
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
LSAY	Longitudinal Surveys of Australian Youth
MCEECDYA	Ministerial Council for Education, Early Childhood Development and Youth Affairs (superseded by the Standing Council on School Education and Early Childhood)
METeOR	AIHW's Metadata Online Registry
NAP	National Assessment Program
NAPLAN	National Assessment Program – Literacy and Numeracy
NCVER	National Centre for Vocational Education Research
NDIS	National Disability Insurance Scheme
NEA	National Education Agreement
NECECC	National Early Childhood Education and Care Collection
NERA	National Education Reform Agreement
NSSC	National Schools Statistics Collection
NSSCWG	National Schools Statistics Collection Working Group
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
SA1	Statistical Area level 1
SACC	Standard Australian Classification of Countries
SCDC	Strategic Cross-sectoral Data Committee
SCSEEC	Standing Council on School Education and Early Childhood (superseded by the Education Council as of 1 July 2014)
SEIFA	Socio-Economic Indexes for Areas
SES	socioeconomic status
SLA	statistical local area
SLK	statistical linkage key
SLK-581	The statistical linkage key comprising letters 2, 3 and 5 of surname, letters 2 and 3 of given name, 8 digits of date of birth and sex
TAA	Test Administration Authority

TAFE	technical and further education
TETIA	Transforming Education and Training Information in Australia
TIMSS	Trends in International Mathematics and Science Study
TYIMS	Training Youth Information Management System
URL	unit record level
USI	Unique Student Identifier
VET	vocational education and training

Summary

Background

The Australian Institute of Health and Welfare (AIHW) undertook a project in 2013–14 to develop a national data standards strategy and implementation plan to enhance the comparability, quality and coherence of information across the education and training sectors, including early childhood education, school education, vocational education and training (VET) and higher education. The project was divided into 3 phases.

Phase 1: Establishment of scope

Phase 1 primarily involved defining the scope of the project, including the data collections and data items in scope of the national data standards strategy. This phase also involved a review of the preliminary work that formed the basis of the project and the development of a strategy to guide consultation and communication with stakeholders.

Phase 2: Review of data collections

Phase 2 involved an extensive data mapping exercise to review the consistency and alignment of data items with existing national data standards and to identify data gaps.

Questionnaires were developed and circulated to data custodians and other stakeholders to collect detailed information on reporting outputs, existing plans for data standards improvements, revisions to enhance comparability in data standards, and opportunities for improvements based on data collection cycles and data linkage capability.

These processes informed broad recommendations and the identification of priorities (including priority data items) for the national data standards strategy.

Phase 3: Development of the national data standards strategy

Phase 3 involved developing the national data standards strategy and implementation plan. This was informed by the priorities identified in Phase 2 of the project and the collection and synthesis of extra information. This phase also included preparing a project report to document the process and background information used to inform the strategy and implementation plan (this report).

Project finalisation

The national data standards strategy and implementation plan were provided to the Data Strategy Group (DSG) in 2014 for consideration as part of the DSG forward work program.

The strategy provided a vision and goal for education and training data, and priority areas to achieve these.

The implementation plan outlined specific projects and resources in each priority area. It identified opportunities for improvements to national standards in the short, medium and long term, taking into consideration the potential benefits to be gained and relevance to policy priorities, the level of work needed to implement changes, and the level of priority assigned by data custodians and stakeholders.

1 Introduction

The AIHW developed a national data standards strategy and implementation plan for the education and training sectors. The strategy covers the early childhood education, school education, VET and higher education sectors.

This work was undertaken under the Transforming Education and Training Information in Australia (TETIA) 2013–14 Forward Work Program, overseen by the DSG¹ and previously by the Strategic Cross-sectoral Data Committee (SCDC) (to September 2014).

The purpose of the project was to:

1. increase cooperation and engagement with stakeholders across sectors, by developing a comprehensive and agreed strategy for improvements to data quality and coherence of information across sectors
2. enhance nationally comparable data and data that facilitate state and territory comparison with national data, including those over time
3. increase the flexibility of data to be drawn together to address national needs, and answer research and policy questions in a timely manner.

The project was divided into 3 phases:

- Phase 1: Establishment of scope (October to December 2013)
- Phase 2: Review of data collections (January to March 2014)
- Phase 3: Development of national data standards strategy and final project report (March to October 2014).

These phases and their components are depicted in Figure 1.1.

This project report summarises the activities and process undertaken in each phase of the project (chapters 2–4) to develop the national data standards strategy and implementation plan for the education and training sectors, and the finalisation of the project.

¹ The Data Strategy Group provides high-level strategic data policy advice to the Australian Education, Early Childhood Development and Youth Affairs Senior Officials Committee (AEEYSOC) to support the SCSEEC Strategic Reform Framework and the development of evidence-based policy, including monitoring and reporting on school education and early childhood outcomes.

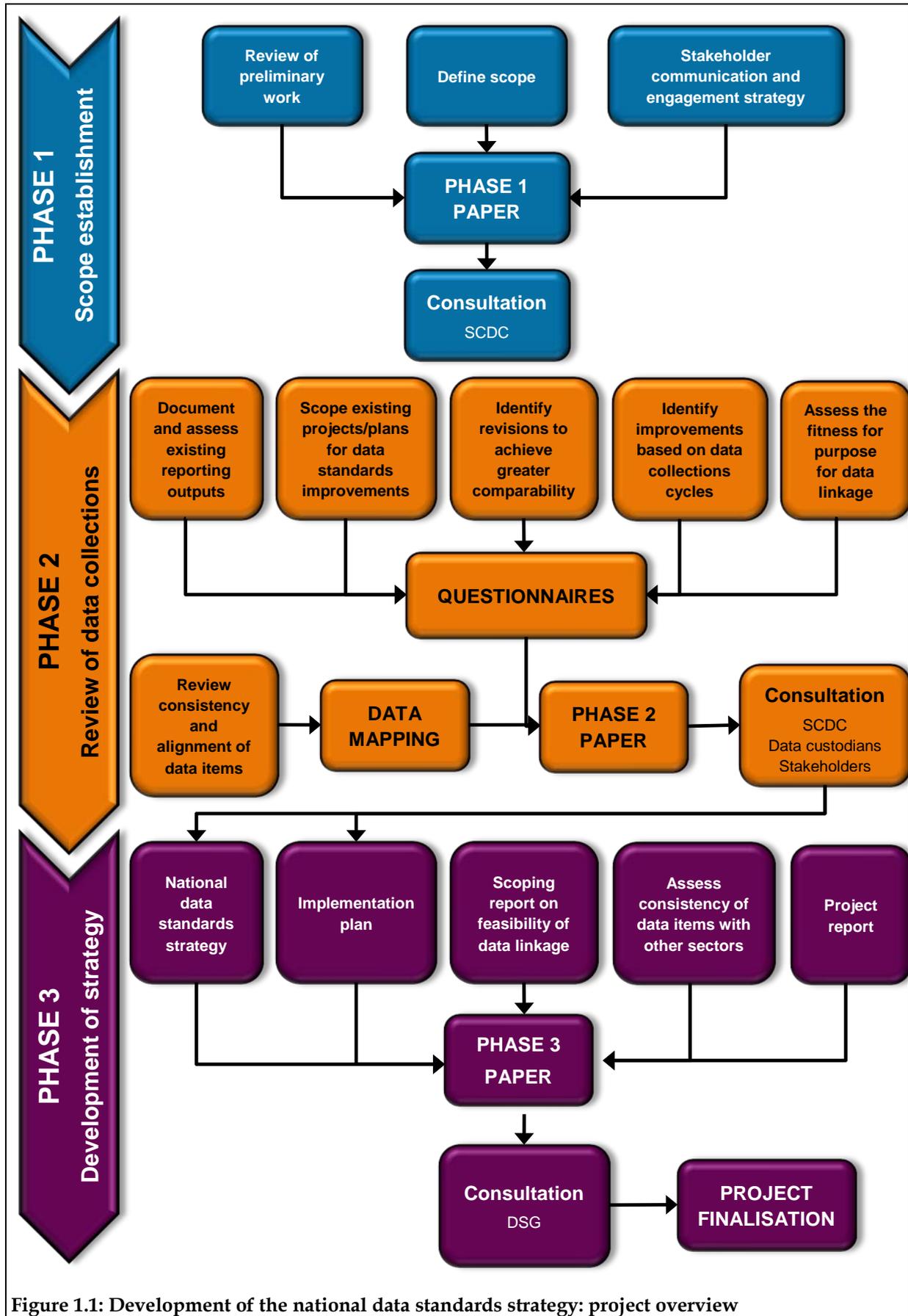


Figure 1.1: Development of the national data standards strategy: project overview

2 Establishment of scope (Phase 1)

Phase 1 primarily involved defining the scope of the project, including the data collections and data items 'in scope' of the national standards strategy. This phase also involved reviewing the preliminary work done by the SCDC Secretariat that formed the basis of the project and developing a strategy to guide consultation and communication with stakeholders.

This work culminated in a paper on Phase 1 of the project that was presented to SCDC for consultation.

2.1 Review of preliminary work

In February 2013, the SCDC Secretariat prepared a preliminary report outlining the potential for data standards development for education and training data collections as part of the TETIA 2013-14 Forward Work Program.

The preliminary report assessed 11 key data collections considered to potentially be suitable for future data linkage projects. Focus was given to the identification of core data items, namely socio-demographic items, which were assessed for their availability and linkage comparability across data collections.

In March 2013, the preliminary report was provided to data custodians responsible for the assessed data collections for feedback on progressing future data standards work and implementing changes. Data custodians raised a number of key issues, including: the scope of a national data standards strategy; change in data sets over time; leveraging existing initiatives; addressing known issues and prioritisation; investigation of all key variables required for data linkage; privacy, legislation and data accessibility; scope, reasons and policy intent of data collections; and timing and governance processes to implement changes.

Based on the feedback from data custodians, the key recommendations for the development of a national data standards strategy were to:

- develop a scoping statement in consultation with stakeholders, including the specification of metadata and response units as key aspects of the standards
- review the assessment of data collections in consultation with stakeholders to ensure the metadata and collection specification information is up to date
- consult with stakeholders to acknowledge and leverage existing initiatives to improve data standards
- consult with data custodians to address known issues, scope, reasons and policy intent of data collections thereby increasing data custodian investment and implementation of the strategy
- develop an implementation strategy for the national data standards strategy by collating collection cycles, data standards review cycles and governance processes in consultation with stakeholders to review timeframes and processes for implementing changes to data collections.

The preliminary report and these recommendations formed the basis of the current project.

2.2 Stakeholder communication and engagement strategy

A stakeholder communication and engagement strategy was developed for the project, in consultation with the SCDC. The strategy outlined the key stakeholders and the methods for engagement throughout the project. Key elements of the final strategy are summarised below.

Objectives

The objectives of the stakeholder communication and engagement strategy were to:

- ensure that an appropriate range of stakeholders were engaged and participated in the consultation
- incorporate advice and feedback from relevant experts, including those with education, training, statistical/technical and government policy/administration expertise
- foster a sense of ownership of and commitment to the outcomes of the project among stakeholders.

Stakeholders

Stakeholders for the project were identified as falling into 3 broad groups, noting that some stakeholders may be included in more than 1 group:

1. those engaged in determining policy direction, development and implementation
2. data providers, data managers, data custodians and those reporting education and training data
3. other relevant expert representation.

A complete list of the stakeholders consulted is included at Appendix A.

Consultation

The parts of the project that were identified as key points for consultation were:

- establishment of scope (Phase 1) – identified as a key consultation point with SCDC
- review of data collections (Phase 2) – identified as a key consultation point with SCDC and broader stakeholders including data custodians, education and training sector representatives (state/territory and Australian Government and non-government), and key committees, councils and advisory groups
- development of the national data standards strategy (Phase 3) – identified as a key consultation point with SCDC (and DSG following the handover of the project in September 2014) and data custodians.

These consultation points are indicated in Figure 1.1.

2.3 Scope of the national data standards strategy

Phase 1 of the project also involved developing a scoping statement in consultation with the SCDC. The scoping statement documented the purpose and benefits of the project, identified

the education and training data collections and data items in scope of the strategy, and described the project outline and deliverables.

The AIHW sought feedback from the SCDC on the scoping statement, and in particular in regard to the data collections and data items for inclusion. The SCDC was asked:

1. Have all relevant collections been included in the scope? Are there additional national data collections that should be included?
2. Have all relevant data items been included in the scope? Are there additional national data items that should be investigated?

As a result of this consultation process, the number of data collections and data items within scope was expanded. The additional data collections and data items are indicated in tables 2.1 and 2.2 respectively.

Final scope

The scope of the strategy is to support the development and implementation of national data standards for a core set of 53 data items across 15 national education and training data collections. The data collections and data items to be covered by the strategy were identified by the SCDC and do not include all data collections relevant to education and training. The focus is on administrative collections and censuses with an education and training focus (with the exception of PISA, TIMSS and PIRLS, which are studies of a sample of the student population) that are potentially suitable for future data linkage projects. Table 2.1 lists the data collections in scope of the strategy and the data custodians with primary responsibility for the use, disclosure and protection of the data.

The data items initially identified for inclusion were core data items required to enable data exchange and linkage, primarily socio-demographic data items such as name, sex, date of birth and address. This scope was expanded following consultation with the SCDC in Phase 1 of the project to include extra items of policy and research interest. The final data items in scope of the strategy have been grouped into 5 categories:

- core **socio-demographic** data items
- core **education** data items
- data items that are integral for **data linkage**
- **disability status** and other health data items
- **family composition** data items.

A detailed list of the data items is in Table 2.2.

Table 2.1: National education and training data collections (15) in scope of the national data standards strategy

Data collection	Data custodian(s)
Australian Early Development Census (AEDC)	Australian Government Department of Education and Training
Census of Population and Housing	ABS
Higher Education Statistics Collection	Australian Government Department of Education and Training
National Apprenticeship and Traineeship Collection	NCVER, on behalf of the state and territory and Australian governments
National Assessment Program (NAP) ^{(a)(b)}	ACARA, state and territory test administration authorities
National Early Childhood Education and Care Collection	ABS, state and territory education departments, Australian Government Department of Education and Training
National Early Education and Care Workforce Census ^(b)	Australian Government Department of Education and Training
National Schools Attendance Collection ^(b)	ACARA, state and territory education departments, Catholic and independent schools, and Australian Government Department of Education and Training
National Schools Statistics Collection (NSSC)	ABS, state and territory education departments, and Catholic and independent schools
Programme for International Student Assessment (PISA)	ACER, Australian Government Department of Education and Training
Progress in International Reading Literacy Study (PIRLS) ^(b)	ACER, Australian Government Department of Education and Training
Trends in International Mathematics and Science Study (TIMSS)	ACER, Australian Government Department of Education and Training
National VET Provider Collection	NCVER on behalf of the state and territory and Australian governments
National VET in Schools	NCVER on behalf of the state and territory and Australian governments

(a) The NAP includes the National Assessment Program—Literacy and Numeracy (NAPLAN), NAP Sample Assessments—science literacy, civics and citizenship and information and communication technology (ICT) literacy and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) Data collections added to scope following SCDC consultation in Phase 1. For NAP, this refers to NAP Sample Assessments only.

Table 2.2: Detailed list of data items

Data items by category	
Socio-demographic data items	Education data items
Age	Educational institution—identifier
Country of birth	Educational institution—address line
Date of birth	Educational institution—address suburb
Geography—address line	Educational institution—address postcode
Geography—address suburb	Educational institution—address state/territory
Geography—address postcode	Educational institution—concurrent
Geography—address state/territory	Educational institution—previous
Geography—Mesh Block	Educational institution—type
Geography—Collection District (ASGC)	Attendance/enrolment status ^(a)
Geography—SLA (ASGC)	Field of education ^(a)
Geography—SA1 (ASGS)	Educational achievement ^(a)
Geography—SA2 (ASGS)	Highest school level completed
Geography—remoteness	Highest non-school qualification completed
Income	
Indigenous status	Statistical linkage data items
Language spoken at home	Date of birth
Occupation	Identifier code—non-SLK
Parent’s country of birth	Identifier code—SLK
Parent’s education—highest school level completed	Geography—address postcode
Parent’s education—highest level of qualification	Geography—Mesh Block
Parent’s education—highest level of non-school qualification	Geography—address suburb
Parent occupation	Geography—SA1 (ASGS)
Proficiency in English	Name—family name
Sex	Name—given name
SES for person/family ^(b)	Sex
SES for area ^(b)	
Year of arrival	Health and disability data items
	Disability or special needs status/flag ^(a)
	Disability type ^(a)
Family data items	
Family composition ^(a)	

(a) Data items added to scope following SCDC consultation in Phase 1.

(b) These data items are not directly collected but, where they exist, are derived from other data items (for example, geography data items for SES for area).

3 Review of data collections (Phase 2)

Phase 2 of the project involved collecting detailed information on the data collections through a data mapping exercise and consultation with stakeholders and data custodians on in-scope data collections and additional information needs. More specifically, the detailed tasks were to:

- review the consistency and alignment of data items with existing national standards, including the identification of data gaps
- document and assess existing reporting outputs and identify areas for improvement in relation to data standards in data collections in scope
- scope existing projects/plans for data improvements in relation to data standards in data collections in scope
- identify revisions required to data items, collections and reporting requirements to achieve greater comparability across programs, to improve the ability of data to inform policy issues; and to improve data collection and reporting methods
- identify opportunistic improvements (to data systems) based on existing data collection cycles/data systems upgrades
- assess the fitness for purpose of data collections for data linkage.

The results of this phase were provided to the SCDC and other stakeholders for consultation. These results are summarised in this chapter.

3.1 Data mapping

A comprehensive review of the consistency and alignment of data items between education and training data collections and with existing national data standards was conducted to inform priorities for the strategy. A summary of the review is included in this section; details of methods and results of the assessment are set out in Appendix B.

The core data items across the 5 different categories (socio-demographic, education, disability and health, family and data linkage) were assessed for consistency with national data standards in METeOR, the AIHW metadata online registry. Metadata for the in-scope education and training data collections was compared to nationally endorsed data standards within the health, housing, community services and early childhood sectors, to determine the consistency of definitions and coding options between the data collections and with national standards. These national standards in METeOR were used to provide a single comparator for each data item, so that all education and training data standards were compared with the same standard. It is recognised that data items from education and training data collections that are not collected consistently with the national standards used in this comparison are still collected to agreed standards within the specific data collection or education and training sector (for example, within the school or VET sector).

Fifty-three data items were assessed for each of the 11 data collections included in the mapping. This resulted in a total of 583 data items that were mapped and assessed for consistency against national standards, where available. An effort indicator was then applied to classify the level of effort needed to align data items with national standards.

Table 3.1 and Figure 3.1 compare the data item standards of individual collections with the national data standards in METeOR.

Overall:

- 29% (139) of data items were fully consistent with national data standards. No effort is required for these data items to be consistent with existing standards.
- Some effort would be required to improve the consistency of 25% (119) of data items that are derivable or mappable to standards.
- 5% (24) of data items were not consistent with national data standards and considerable effort would be required to align these data items with existing national standards.
- The largest proportion of data items (41% or 197) includes items that are considered to be relevant but do not exist in the collection, are unable to be determined due to a lack of information, or are collected but a national standard does not exist. This group of data items would also require considerable effort to ensure comparability.
- 104 data items were graded as not applicable. These data items are excluded from the per cent calculations in this section.

Socio-demographic data items:

- 30% of data items were entirely consistent with national standards. Language spoken at home, country of birth, date of birth, and geography address items (suburb, state, postcode) were the most consistent with standards across the data collections.
- 29% were derivable or mappable to existing standards. Age was the item most commonly derived.
- Only 7% of data items were not consistent. Remoteness, parent's country of birth and parent occupation were inconsistent across 3 data collections.
- 34% of data items were either relevant but not collected, unable to be determined or were collected but a national standard does not exist. Data items on geography, proficiency in English, year of arrival and parental information were the most common items not collected but deemed relevant to collections.
- 68 data items were determined to be not applicable to a data collection.

Education data items:

- 25% of items were fully consistent with national standards.
- 19% were derivable or mappable.
- 28% of items were assessed as potentially relevant to the data collection but not currently collected, most commonly concurrent and previous educational institution data items.
- 19% of data items do not currently have an existing national standard in METeOR; for example, attendance or enrolment status and educational institution type.

Disability data items:

- 2 items for disability were reviewed. The first was a flag for having disability and the other described disability type.
- 5 collections collected the disability flag consistently with the national standard.
- No collection collected the disability type item according to, or mappable to, the national standard.

Family composition data item:

- A single data item for family composition was reviewed, with only 2 collections collecting this item. Both were consistent or mappable to the national data standard.
- The data item was graded as relevant but not collected for the other data collections.

Statistical linkage data items:

- More than half of the statistical linkage data items were either fully consistent, derivable or mappable to national data standards (62%).
- Statistical linkage key, given name and family name are consistent, derivable or mappable to the national data standards in the Higher Education Statistics Collection, NECECC and NSSC. While all of the collections other than the Census collect name items, this information may not be available for linkage.

Table 3.2 provides an overview of the level of effort required within each data collection for all data items to be consistent with existing national data standards in METeOR. It is based only on data items that are determined to be applicable to the individual data collection, and is ranked according to the total effort required to meet existing national data standards.

- The data collections contained between 49% and 95% of core data items in scope of the strategy that would require some or considerable effort to meet existing national data standards.
- Between 11% and 44% of data items were identified as requiring some effort across the collections and 20% to 76% as requiring considerable effort.

Table 3.1: Summary of data item consistency with national data standards, by 11 data collections^(a)

Effort indicator	Group	AEDC	Apprentices	Census	Higher Education	NAP ^(b)	NECECC	NSSC ^(c)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools	Total	
													No.	%
No effort required	Consistent	11	12	17	11	6	22	17	2	2	21	18	139	29
Some effort required	Derivable or mappable	6	12	10	19	10	5	14	11	8	12	12	119	25
Considerable effort required	Not consistent	2	0	1	4	3	0	1	5	6	1	1	24	5
	Unable to determine or unavailable	26	19	13	9	25	18	14	24	26	7	16	197	41
Not applicable		8	10	12	10	9	8	7	11	11	12	6	104	..
<i>Total</i>		53	53	53	53	53	53	53	53	53	53	53	583	..
Total (excluding not applicable)		45	43	41	43	44	45	46	42	42	41	47	479	100

.. Not applicable (category does not apply).

(a) Refer to Appendix B for the rationale for the inclusion of 11 rather than all 15 data collections in the review of data item consistency and additional notes on several of the data collections.

(b) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(c) The assessment of the NSSC is based on unit record level (URL) components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Note: Refer to Appendix B for an explanation of the categories used in this table.

Source: Appendix Table B2.

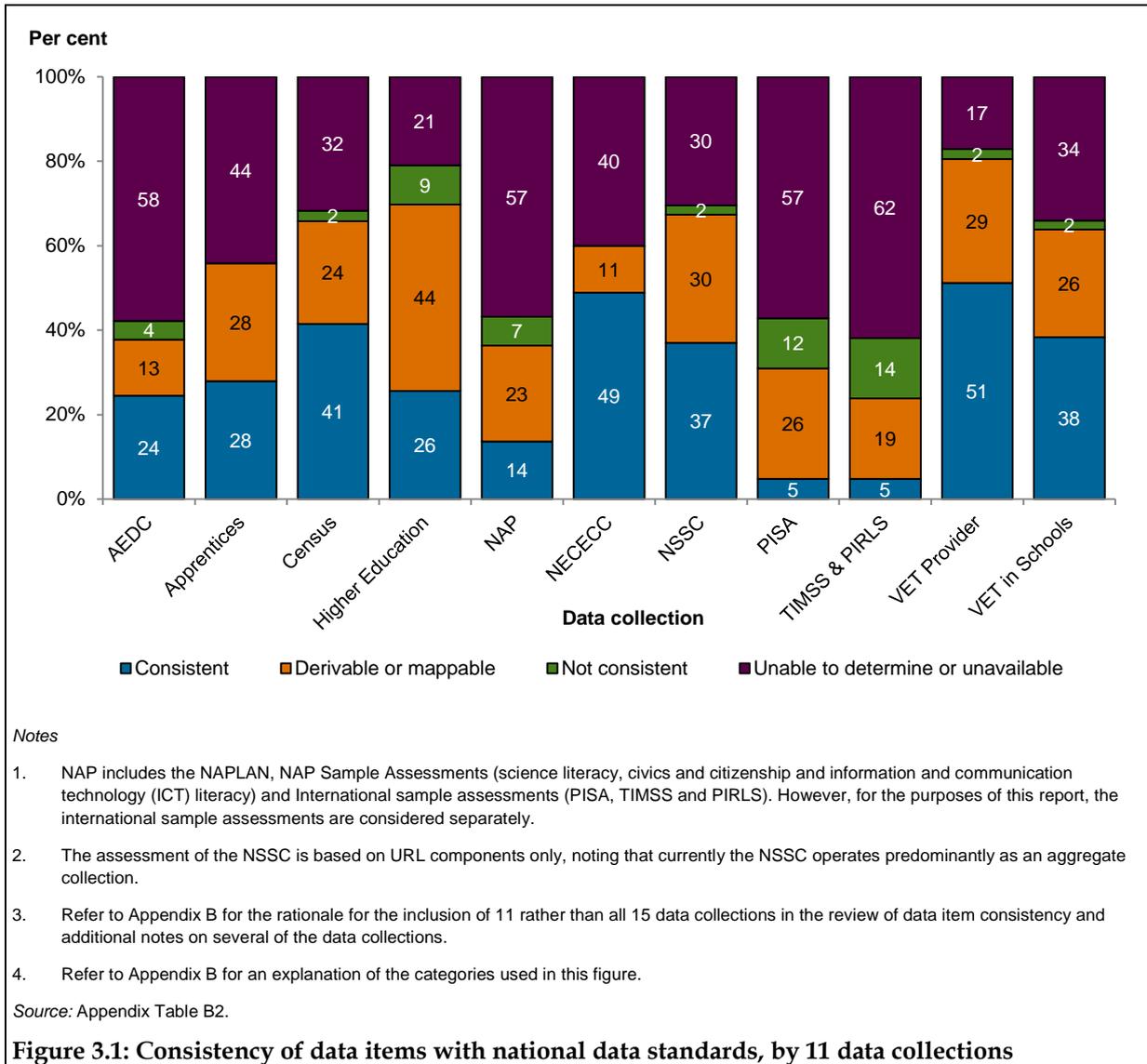


Table 3.2: Data items requiring some or considerable effort to meet national data standards, by data collection^(a)

Data collection	Data items					
	Some effort required		Considerable effort required		Total with effort required ^(d)	
	Number	Per cent	Number	Per cent	Number	Per cent
PISA	11	26	29	69	40	95
TIMSS & PIRLS	8	19	32	76	40	95
NAP ^(b)	10	23	28	64	38	86
AEDC	6	13	28	62	34	76
Higher Education	19	44	13	30	32	74
Apprentices	12	28	19	44	31	72
NSSC ^(c)	14	30	15	33	29	63
VET in Schools	12	26	17	36	29	62
Census	10	24	14	34	24	59
NECECC	5	11	18	40	23	51
VET Provider	12	29	8	20	20	49

(a) Refer to Appendix B for the rationale for the inclusion of 11 rather than all 15 data collections in the review of data item consistency and additional notes on several of the data collections.

(b) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(c) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

(d) 'Some effort required' and 'Considerable effort required' categories combined. The denominator is the total number of data items minus the number of not applicable data items (refer to the last row of Table 3.1 for the denominators used).

Notes

1. Refer to Appendix B for an explanation of the categories used in this table.
2. Components may not sum to totals due to rounding.

Source: Appendix Table B3.

3.2 Questionnaires

Questionnaires were developed for stakeholders and data custodians to gather information on in-scope data collections in relation to:

- reporting outputs
- existing projects/plans for improvements to data collections and recently completed projects
- identification of areas for improvement in data collections and their priority in relation to policy, to increase comparability and to improve collection and reporting methods
- data collection and review cycles in order to identify opportunities for improvements
- assessing the current fitness for purpose of data collections for data linkage.

A copy of each of the questionnaires circulated is at Appendix D.

The questionnaire consultation period was from January to April 2014. One hundred and fifty questionnaires were sent to stakeholders and data custodians across 35 different

organisations (Appendix A). Of the 35 organisations, 17 (49%) provided at least 1 response, with many of the larger organisations submitting responses from multiple areas within the organisation. Non-responders were followed up.

A summary of the feedback received from data custodians and stakeholders via the questionnaires is available in Appendix E.

3.3 Results

The results of the data mapping and the data custodian and stakeholder questionnaires were analysed. The key findings from both components were then collated. The following issues emerged as the highest priority from a data standards perspective.

Indigenous status, remoteness and socioeconomic status data standards

Stakeholders and data custodians indicated strong support for the collection of consistent information for these data items. These are core groups used for performance indicator reporting and the information is also important for formulating and targeting strategies.

The data mapping indicated that the Indigenous status data item was collected by 4 of 11 data collections in a manner that was fully consistent with the national standard.

Remoteness was collected or able to be derived for all data collections. Eight of the collections were consistent with the standard or able to be derived or mapped, while 3 were not consistent with the national standard. Work on remoteness data standards is currently occurring.

Socioeconomic status information, while available or derivable for all data collections, is based on different geographies and/or different indexes and therefore is not collected in a consistent manner between collections. Previous work done for the former SCDC in relation to assessing the coherence of socioeconomic status measures in education and training data collections has suggested that this information need may be best addressed within sectors in the short term.

For a number of data collections, remoteness and socioeconomic status data items are able to be derived, but are not specifically collected. Adding these as derived data items to the collections would increase the availability of information on these groups.

Availability of information on country of birth, proficiency in English and year of arrival in Australia

Strong support for increasing the availability of consistent information for these items in the data collections emerged from the consultation. This was then checked against the data mapping, which showed that there was already a moderately high level of consistent information collected for country of birth information, although some scope for further consistency remains.

The data mapping also indicated that proficiency in English and year of arrival in Australia were collected by less than half of the national education and training data collections in scope of this project (4 and 2 collections respectively). Since most collections do not collect this information, it may be feasible to assess whether data linkage would be a viable option where this information is needed.

Disability

The data mapping and consultation identified that disability information was lacking in availability and consistency. Although concerns were raised about the appropriateness and the ability of education and training data collections to accurately collect this information, a need for consistent standards was expressed through the consultation.

Further work would need to be done to determine the appropriateness and feasibility of this, and whether a single standard would be possible across the sectors, or whether linkage with other collections would be viable to obtain this information. Considerable work has occurred on the collection of information on students with disability in the school sector (SCSEEC 2014).

Education

A number of education data items were identified in the data mapping as lacking in availability and/or consistency: educational institution concurrent, educational institution previous, educational institution type, attendance/enrolment status.

The data mapping indicated that information on educational institution (concurrent, previous, type) was lacking in availability and consistency across potentially relevant collections. These are therefore areas where the addition of data items and the development of standards could be considered, and this could also increase the information available for data linkage.

Parental information

The availability and consistency of data items relating to parents was identified in the data mapping as an area of weakness, given the importance of this information in relation to children's educational performance and other outcomes.

Data items relating to parents' education and occupation were available from 5 and 4 data collections respectively, while parents' country of birth was collected for 3. For parents' country of birth and occupation, there was little or no consistency with existing national standards.

Data linkage

This phase of the project involved a preliminary assessment of the feasibility of data linkage for the education and training data collections in scope of the project. Each collection was assessed individually according to basic criteria that could be applied consistently across the data collections. Six of 11 data collections were assessed to have the core and supplementary linkage items requisite for data linkage. These results did not take into account that some data collections have legislative provisions which restrict access to the core and supplementary data linkage items.

It was noted that the availability of small area geographic data items would assist the accuracy of data linkage across the data collections, particularly in the absence of name, address and SLK information. The addition of small area geographic data items (Mesh Block or SA1) to national education and training data collections that do not currently collect them could therefore be considered, noting that SA1 may be more feasible than Mesh Block.

Further information on the feasibility of data linkage is available in Chapter 4 of this report; however, for a more detailed assessment, further work would need to be done to determine specific linkages between collections that would yield the greatest benefits in terms of addressing policy and research priorities. This would allow a more detailed linkage feasibility assessment to be undertaken between the relevant collections.

Other data items

Some other data items were identified by stakeholders for collection and standardisation, including family composition, sibling order, refugee status, prior education, child protection status and health care card status. However, these did not emerge as priority data items.

It was also noted that there is an increasing demand for reporting on vulnerable children.

3.4 Priority data items

Based on the data mapping and consultation with data custodians and stakeholders, a number of data elements and activities were put forward as priorities for further development, including:

- increasing the availability of consistent information on country of birth; proficiency in English; year of arrival in Australia; visa and/or migrant status; education items, specifically educational institution concurrent, educational institution previous, educational institution type, attendance/enrolment status; and disability
- improving consistency with existing national data standards for Indigenous status and remoteness
- improving the availability of parental information given the importance of this information in relation to children's educational performance and other outcomes
- improving the availability of small area geographic data items to assist the accuracy of data linkage across the data collections.

Subsequent consultation with the SCDC and other stakeholders provided broad support for each of these areas. The need to align the national data standards strategy, and any specific strategies to improve the availability and consistency of data items, with policy priorities and clear benefits was strongly noted by stakeholders. Disability items were noted to be highly relevant to current policy, while the collection of proficiency in English and parental information was identified as important to inform service provision and resource allocation. The availability and consistency of socioeconomic status information was strongly prioritised and considered necessary to enhance nationally comparable data to inform policy and research priorities. Increasing the availability of smaller level geography measures, such as Mesh Block and SA1, was identified as a means to enhance the quality of socioeconomic status and remoteness, as well as to improve the quality of potential data linkage. Data linkage was identified as a cost-effective method to obtain information on parental characteristics, socio-demographics and disability.

The priority data items were refined following consultation and consideration of policy priorities (see Chapter 4 for further discussion of policy priorities). The items are:

- Indigenous status
- remoteness
- socioeconomic status
- disability
- proficiency in English
- parental education and occupation
- geography – SA1

- education data items: educational institution concurrent, educational institution previous, educational institution type, attendance/enrolment status.

4 Development of the national data standards strategy (Phase 3)

Phase 3 of the project built on the priority data items identified as a result of Phase 2 and the synthesis of additional information collected from the Phase 2 questionnaires to inform the national data standards strategy and implementation plan. Phase 3 also included a number of extra tasks:

- an assessment of the consistency and comparability of core data items with other sectors (health, community services and housing and homelessness)
- preliminary scoping of the feasibility of data linkage
- a project report summarising the process and background information for the development of the strategy and implementation plan (this report).

This chapter gives an overview of these tasks.

4.1 National data standards strategy and implementation plan

The national data standards strategy and implementation plan were developed based on the outcomes of the Phase 2 information gathering, analysis and consultation processes. These processes informed the vision, goal and priorities of the strategy.

The implementation plan outlines the projects to be undertaken to achieve the priorities and goal of the strategy. Based on information gathered in Phase 2, the implementation plan identifies opportunities for improvements to national standards in the short, medium and long term, taking into consideration the potential benefits to be gained and relevance to policy priorities, the level of work needed to implement changes, and the level of priority assigned by data custodians and stakeholders. It also takes into account existing and planned projects that aim to improve collections and standards, as well as data collection cycles and data flows, to identify opportunities for changes to be implemented.

This section describes information developed in Phase 3 of the project that, in addition to the Phase 2 output, informed the strategy and implementation plan.

Policy priorities

There is substantial investment at all levels of government to provide quality education and training and to ensure successful transitions to school, further education, training and work. The factors and outcomes associated with these transitions overlap with other policy areas of health, disability, equity and workforce development.

Through the Phase 2 consultation process, stakeholders emphasised the need to align the national data standards strategy, and any specific strategies to improve the consistency of data items, with policy priorities. A review of current government policy priorities was undertaken and used to inform the strategy. A summary of relevant priorities in each area as at October 2014 is presented in this section, with further detail provided in Appendix F.

A common theme within each area is the enhancement of knowledge management and innovation to enable the monitoring and reporting of outcomes to build an evidence base. A

national data standards strategy establishes an approach to quality, consistent national education data items that will assist in following student education pathways and outcomes from early childhood to adulthood, building the evidence base to assist research and policy and enabling a longitudinal view of education, training and outcomes for Australians. This will enhance the understanding of educational outcomes, and the social and economic returns from investing in education.

The addition or modification of data items in education and training collections to improve consistency requires considerable investment of resources to review, standardise and endorse changes, and subsequently implement changes across multiple systems. Aligning the strategy with policy priorities, and particularly where there are critical gaps in reference to policy, provides a sound basis and demonstrated value for the implementation of changes, and is consistent with the vision of the strategy.

Early childhood

It is the vision of the National Early Childhood Development Strategy for all children to have the best start in life to create a better future for themselves and the nation. Continued evidence identifies the profound impact of the early years on future health, development and wellbeing. As such, it is a priority for all children to have access to affordable, quality education. Consequently, several National Partnership Agreements are in place to ensure quality early childhood education that is accountable and provides clarity around expectations via National Quality Standards and a rating system.

Aboriginal and Torres Strait Islander people

Indigenous Australians have poorer results than non-Indigenous Australians in school attendance, Year 12 attainment, literacy and numeracy and employment. To help in closing the gap between the outcomes of Indigenous and other Australians, specific reforms and action plans exist to attain targets for Indigenous Australians with a focus on transitions to school, further education and work.

Disability and health

The objective of the National Disability Agreement is that ‘people with disability and their carers have an enhanced quality of life and participate as valued members of the community’ (COAG 2012c). The National Partnership for More Support for Students with Disabilities assists the opportunities of students with disabilities by providing support through an inclusive environment in schools to help students with disabilities in their educational experience, outcomes and transitions to further study or work. The Intergovernmental Agreement for the National Disability Insurance Scheme (NDIS) Launch provides the foundation for governments to work together to develop and implement the first stage of the NDIS to ensure that people with disability and their carers receive access to the required support.

The importance of consistent data on disability is emphasised in the National Disability Strategy, particularly in relation to the collection of data on school students with disability and the reporting of educational outcomes for students with disability.

Education and training

All Australian Education Ministers agreed to the *Melbourne declaration on educational goals for young Australians* in 2008. The goals under the Melbourne Declaration are that Australian

schooling promotes equity and excellence, and that all young Australians become successful learners, confident and creative individuals, and active and informed citizens.

The National Education Agreement (NEA) and the National Education Reform Agreement (NERA) set out the objective that 'all Australian school students acquire the knowledge and skills to participate effectively in society and employment in a globalised economy'. The agreements contribute to the achievement of the outcomes that: all children are engaged in and benefiting from schooling; students meet basic literacy and numeracy standards and that levels of achievement are improving; schooling promotes the social inclusion and reduces the education disadvantage of children, especially Indigenous children; Australian students excel by international standards; and young people make a successful transition from school to work and further study.

A number of performance indicators in the NEA and the NERA draw on data from collections in scope of the strategy.

The Higher Education Participation Programme aims to remove disadvantage as a barrier to educational attainment. This includes eliminating the effects of disadvantage such as low socioeconomic status, disability, refugee status, homelessness and remoteness to ensure all Australians have the ability to attain an education.

Workforce

Several government funds and initiatives exist to upskill and enhance the workforce of Australia. These policies exist across different education sectors, including schools, training and higher education. For example, the School Pathways Program undertaken in selected participating schools aims to increase the pool of students with prerequisites in science, technology, engineering and mathematics to continue further education to support the defence sector. The objective of the National Agreement for Skills and Workforce Development is to achieve a VET system that delivers a more productive and highly skilled workforce, enabling all working-age Australians to participate effectively in the labour market and contribute to Australia's economic future. The Industry Skills Fund offers loans to enhance the completion rate of apprenticeships, while the Early Childhood Workforce Initiatives aid university completion of those studying early childhood to build workforce capacity to deliver quality early childhood education and care programs.

Recent and future developments in data standards and linkage

A national data standards strategy that builds on previous and existing work to improve the consistency and quality of data standards will help to ensure that the strategy will achieve the intended outcomes more efficiently and in line with existing initiatives. The investment and engagement of data custodians in the strategy is more likely if the strategy aligns with established and agreed priorities, which may have activities, networks, processes and infrastructure in place already.

Data standards

Table 4.1 outlines the recent and current developments in relation to data standards for each of the education and training data collections in scope of the strategy, as at June 2014, based on information provided by data custodians.

Table 4.1: Data standards developments for the 15 national education and training data collections, as at June 2014

Data collection	Developments	Status
Australian Early Development Census (AEDC)	Background information items aligned with Standing Council on School Education and Early Childhood standards. Collection geography aligned with Australian Statistical Geography Standard	Complete
Census of Population and Housing	Improvements are planned for the 2016 collection to update its relevance to modern Australia	In progress
Higher Education Statistics Collection	Modification of OS-HELP submission in 2013 to include additional information about Australian students studying in Asia to support program implementation and policy development	Complete
National Apprenticeship and Traineeship Collection	Updates as a result of changes to the Standard (AVETMISS), which is currently under review	In progress
	Consideration of including the Unique Student Identifier (USI)	In progress
	Outcomes from the national harmonisation project currently underway, in particular principle 6 which relates to the agreement between Australian and state and territory governments about data to be collected including common understanding of data definitions, consistent views on quality and transparency of data and agreed protocol for ownership and review	In progress
National Assessment Program—Literacy and Numeracy (NAPLAN)	Changes to the Writing scale and subsequent key performance indicator reporting in 2012 to improve educational outcomes	Complete
	The delivery of NAPLAN online will be implemented from 2017 on an opt-in basis over two to three years	In progress
National Assessment Program (NAP) Sample Assessments	NAP Sample Assessment delivery moved online for educational and administrative improvements	Complete
National Early Childhood Education and Care Collection	Inclusion of new service provider data element to enable the reporting of the new performance indicators to support the National Partnership Agreement on Universal Access to Early Childhood Education	Complete
National Early Childhood Education and Care Workforce Census	Data items and questions under the National ECEC Workforce Census have remained largely consistent from 2010 to 2013 with minor changes to the wording of questions reflecting changes to relevant data standards	Complete
National Schools Attendance Collection	Standards formed in 2012 with implementation for different sectors staggered between 2013 and 2015	In progress
National Schools Statistics Collection (NSSC)	Refinement and introduction of standards for Aboriginal and Torres Strait Islander status; remoteness indicator; average age of students; school participation rates for students aged 6 to 19 years	Complete
	The NSSC is primarily based on aggregate data with some states and territories supplying government school data in unit record form. The ABS engages bilaterally with data custodians to increase the proportion of data supplied in unit record form	In progress
Programme for International Student Assessment (PISA)	Nil	—

(continued)

Table 4.1 (continued): Data standards developments for the 15 national education and training data collections, as at June 2014

Data collection	Developments	Status
Progress in International Reading Literacy Study (PIRLS)	Nil	—
Trends in International Mathematics and Science Study (TIMSS)	Nil	—
VET Provider Collection	<p>Release 7 of the Standard (AVETMISS) for VET Provider was adopted from 2014. This captures new data items (more detailed address information for students to permit reporting to SA1 geographical area, the USI when it is introduced, and identifying skill sets) to support reforms in the VET sector</p> <p>The USI is a number that a student retains throughout their lifetime and will act as a key to their training information, giving students a single point of access to their VET records. The adoption of a USI is likely to improve the quality of VET data and linkage of VET data across and within sections, and therefore could inform best practice for the potential implementation of a USI in other educational sectors</p>	Complete
	Minor update to classification values during 2014 to accommodate the introduction of total VET activity, more frequent reporting, and the incorporation of the VET in Schools Collection	Complete
VET in Schools	Up to 2013, this collection has been an administrative arrangement under AVETMISS but will be integrated into the VET Provider Collection from 2014 using Release 7 of the Standard (AVETMISS)	Complete

Data linkage

The National Education Data Linkage Roadmap is a roadmap for coordinated action by all states and territories and the Australian Government to enhance and build upon previous and future investment in national data linkage infrastructure through the Population Health Research Network. It seeks to do this by broadening the base of data custodians engaged and opening up these research capabilities into the early childhood and education sectors, in order to better support the evidence-based policy and practice needed to improve outcomes for children and young people.

The national data standards strategy is complementary to the National Education Data Linkage Roadmap. The goal of the strategy to improve consistency in national data standards supports and increases the capacity for linkage of data collections under the Roadmap.

There are a number of specific data linkage projects, either underway or planned, involving one or more of the national education and training data collections that are in scope of this project. This information has been obtained from the environmental scan undertaken for the TETIA strategic plan and 3-year work program, as well as stakeholder and data custodian consultation. Projects are grouped by the predominant stage of life that they relate to.

Early childhood and school-age

The Productivity Commission's report (2015) on the inquiry into child care and early childhood learning recommended the creation of a linked data set:

The Australian Government should establish a program to link information for each child from the National ECEC Collection to information from the Child Care Management System, the Australian Early Development Census, and NAPLAN testing results to establish a longitudinal database. Where possible, this should also be linked to other key administration data sets and Censuses.

There are a number of relevant data linkage projects that have been recently completed or are currently underway.

- The Early Childhood Development Researchable Data Set (ECD RDS) project involves the creation of a linked data set that will include data about children from birth to the early years of schooling and contain information about their health, development and educational outcomes. The proposed data collections include birth, perinatal, early childhood education, AEDC and NAPLAN. Once established, the ECD RDS would foster research in early childhood development covering health, human services, and early childhood education and care across the transition from birth to the early years of school education (AIHW 2014a).
- The Building National Data Linkage Capability for the AEDC project aims to improve understanding of the impact of differing jurisdictional services and policy settings on child development and trajectories. The core objective is to establish national linkage capability for the AEDC that can be used by all jurisdictions. The project is being undertaken by the SA Department for Education and Childhood Development and the Telethon Kids Institute.
- The Measuring Educational Outcomes over the Life-course data integration project aims to facilitate improved understanding of learning pathways from early childhood education and schooling through to further education and employment. This study has involved the integration of a range of datasets including the early childhood education, AEDC, school enrolments, NAPLAN, and Census of Population and Housing data. This study has produced analysis that assesses the impact of personal, family, social, and economic characteristics on child development and student achievement, and the longitudinal outcomes of students after school (ABS 2014a, 2014b).
- Two ABS research papers have been released as part of the ABS Census Data Enhancement Education Quality Study, which assess the feasibility and quality of linkage methods for government school enrolment records from the National Schools Statistics Collection (NSSC) and the 2011 Census of Population and Housing to expand the evidence base for education and training policy (ABS 2013c, 2013d).

Youth

- The Building capability to maximise use of VET and Census of Population and Housing data project assessed the feasibility of linkage between the VET in Schools and Census of Population and Housing data, and produced analysis of the socioeconomic characteristics and initial destinations of students participating (or not participating) in VET in schools (ABS 2014c, 2014d).
- Linking the Longitudinal Surveys of Australian Youth (LSAY) with NAPLAN is a pilot project involving linking NAPLAN scores to LSAY to examine the correlations and differences between NAPLAN testing (focused on Year 9) and PISA scores (collected

when most students are in Year 10). The aim is to demonstrate whether it is feasible to link LSAY with another data set, and is the first step in ultimately extending the value of LSAY as a source of information on youth transitions. The work is being done by the NCVET and is funded by the Department of Education and Training.

- The *Data linkage and statistical matching: options for the Longitudinal Surveys of Australian Youth* discussion paper explored 2 issues. The first was the potential for linking data from existing administrative collections to LSAY. The specific administrative data sources considered for linkage with LSAY included NAPLAN, Medicare Australia, Centrelink, the Higher Education Statistics Collection, the National VET Provider Collection and the Census of Population and Housing. The second was the feasibility of combining data from LSAY and the Longitudinal Study of Australian Children (LSAC) (Gemici & Nguyen 2013).

Whole population

- The Australian Census Longitudinal Dataset currently uses data from the 2006 (wave 1) and 2011 (wave 2) Censuses of Population and Housing to build a longitudinal picture of Australian society. With almost 1 million records in the sample in the first release, it is the largest longitudinal data set in Australia (ABS 2013a).

Data flow and collection cycles

Information was obtained from data custodians regarding data collection cycles, including the frequency with which data items and standards are reviewed, the process for implementing changes to data standards and the flow of data from collection to reporting. Flowcharts depicting this information for each data collection are in appendixes G and H. The timeframes for the review of data standards are also summarised in Table 4.2 and presented visually in Table 4.3. The information in these appendixes and tables is based on the most recent cycle of the data collection, and is subject to change for future cycles. The information available for each data collection varies, and is dependent on the input from data custodians.

The information on timeframes for data standards review can be used to identify opportunities for changes to be made in relation to data standards for these collections based on existing data collection cycles/data systems upgrades. Given the substantial differences between the data collections in terms of timeframes and processes for the review of data standards, data collection-specific approaches would be needed to implement any changes.

Table 4.2: Frequency of data collection and review of data standards

Data collection	Frequency of collection and review	Next review	Data standards review period
AEDC	3-yearly, reviewed every cycle	2018 collection year	January to June 2017
Census of Population and Housing	5-yearly, reviewed every cycle	Submissions for 2016 Census closed May 2013 Next opportunity will be 2021 Census	November 2017 to October 2019
Higher Education Statistics Collection	Student data collected 4 times per year, staff data collected annually, reviewed annually	2016 collection year	First half of 2015
National Apprentices and Trainees Collection	Annual, major review approximately every 5 years, minor updates on an as-needs basis	Major review estimated to occur for the 2019 collection year Minor updates for the 2016 collection year	Major review: late 2016 to early 2018 Minor updates: late 2014 to mid 2015
NAP ^(a)	Annual (NAPLAN) and 3-yearly rolling basis (NAP Sample assessments, reviewed as required)	ACARA's data standards manual follows ABS and Education Council requirements and can be updated at any point in time	No fixed review schedule
National Early Childhood Education and Care Collection	Annual, reviewed every cycle	2015 collection year	July 2014 to March 2015
National Early Childhood Education and Care Workforce Census	3-yearly, reviewed every cycle	2016 collection year	2015
National Schools Attendance Collection	Twice yearly from 2015, no fixed review schedule	No fixed review schedule	First half of 2015
National Schools Statistics Collection	Annual, no fixed review schedule	No fixed review schedule	No fixed review schedule
PISA	3-yearly, reviewed every cycle	2018 collection year	2016
PIRLS	5-yearly, reviewed every cycle	2016 collection year	2019
TIMSS	4-yearly, reviewed every cycle	2019 collection year	2017
VET Provider Collection	Annual, major review approximately every 5 years, minor updates on an as-needs basis	Major review estimated to occur for the 2019 collection year Minor updates for the 2016 collection year	Major review: late 2016 to early 2018 Minor updates: late 2014 to mid 2015
VET in Schools	Annual, major review approximately every 5 years, minor updates on an as-needs basis	Major review estimated to occur for the 2019 collection year Minor updates for the 2016 collection year	Major review: late 2016 to early 2018 Minor updates: late 2014 to early 2015

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

Data flow information shows the data supply chain from the collection stage through to reporting outputs, and highlights the data custodians and other stakeholders involved in this process (Appendix D). A common flow of data includes: data collection, submission, processing, verification, analysis, output/reporting development, approvals and release.

Some collections may involve multiple stakeholders and data custodians throughout this process, such as service providers and schools, state and territory departments and research agencies or Australian Government departments (for example, AEDC). Other collections have fewer stakeholders/custodians in the data supply chain, such as state and territory departments and a research agency or Australian Government department (for example, VET Provider), and the Census has only the ABS from collection through to output and reporting.

This demonstrates the multiple levels of stakeholders and data systems that are impacted by changes to data standards and need to be considered in the implementation of any changes, and how this differs depending on the data collection.

Table 4.3: Estimated data standards review periods (shaded cells), by data collection, 2014 to 2021

Data collection	Year and month																																		
	2014						2015						2016						2017																
	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N
Australian Early Development Census																																			
Census of Population and Housing																																			
Higher Education Statistics Collection																																			
National Apprentices and Trainees Collection ^(a)																																			
National Apprentices and Trainees Collection ^(b)																																			
NAP ^{(c)(d)}																																			
National Early Childhood Education and Care Collection																																			
National Early Childhood Education and Care Workforce Census																																			
National Schools Attendance Collection ^(d)																																			
National Schools Statistics Collection ^(d)																																			
PISA																																			
PIRLS																																			
TIMSS																																			
VET Provider Collection ^(a)																																			
VET Provider Collection ^(b)																																			
VET in Schools ^(a)																																			
VET in Schools ^(b)																																			

(continued)

Table 4.3 (continued): Estimated data standards review periods (shaded cells), by data collection, 2014 to 2021

Data collection	Year and month																																															
	2018												2019												2020												2021											
	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D
Australian Early Development Census																																																
Census of Population and Housing																																																
Higher Education Statistics Collection																																																
National Apprentices and Trainees Collection ^(a)																																																
National Apprentices and Trainees Collection ^(b)																																																
NAP ^(c)																																																
National Early Childhood Education and Care Collection																																																
National Early Childhood Education and Care Workforce Census																																																
National Schools Attendance Collection ^(d)																																																
National Schools Statistics Collection ^(d)																																																
PISA																																																
PIRLS																																																
TIMSS																																																
VET Provider Collection ^(a)																																																
VET Provider Collection ^(b)																																																
VET in Schools ^(a)																																																
VET in Schools ^(b)																																																

(a) Minor update.

(b) Major review.

(c) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(d) No fixed schedule.

4.2 Assessment of the comparability and consistency of core data items with other sectors

In Phase 2, an extensive data mapping exercise was undertaken on the data standards used in different education and training collections. The assessment of the comparability and consistency of core data items with other sectors is an extension of this data mapping. The assessment was based on national data standards from the health, housing and homelessness, and community services sectors, where available, and identifies opportunities for linkage of education and training data to these sectors to inform policy and research. This section summarises the results of this assessment.

Methods

The data standards registered in the national health, community services and housing and homelessness data dictionaries for core data items were compared with those for the in-scope education and training data collections to explore the availability and consistency of data standards between the sectors.

The assessment algorithm and coding principles developed for the data mapping exercise undertaken in Phase 2 of this project were used to assess the available metadata and grade the level of consistency between education and training data items and national data standards in other sectors (refer to the methods described in Appendix B for further details).

Results

Overall, the education and training data collections have the highest level of consistency with data standards in the community services national data dictionary at 26%. Consistency with standards in the health and housing and homelessness dictionaries was slightly lower at 24% and 22% respectively.

A further 26% of data items are derivable or mappable to standards within the community services dictionary. Smaller proportions of data items were classified as derivable or mappable to data standards in the health (21%) and housing and homelessness data dictionaries (13%).

Only a small proportion of data items were classified as not consistent with standards in the community services, health and housing and homelessness data dictionaries (between 4% and 6%).

The largest proportion of data items were classified as unable to be determined or unavailable, due to either insufficient information to determine consistency, a relevant data item not being collected in the education and training data collections or because of the lack of a standard in the national data dictionaries. The proportion of data items in this category was highest when compared with the housing and homelessness data dictionary (60%), followed by health (51%) and community services (43%).

Discussion

Comparability of data standards is important to enable linkage between education and training data collections and data collections in other sectors to inform policy and research.

Consistency or comparability of data standards, particularly for statistical linkage data items and key socio-demographic data items, increases the ability to successfully link data sets and the quality of the resulting linked data. The highest level of comparability between education and training data items and national health, housing and homelessness, and community services data dictionaries was for the statistical linkage data item category, followed by socio-demographic, education, family composition and disability.

While the assessment of the consistency and comparability of data items across sectors is important to inform data linkage, it is also useful to identify gaps in the consistency of data standards across sectors, particularly in areas of policy priority.

Therefore, a dual focus on the consistency and comparability of data items relevant to both data linkage and to policy priorities across sectors would provide a platform to prioritise any future data development.

4.3 Preliminary scoping of the feasibility of data linkage

Data linkage is a possible method to address data gaps in education and training data collections. Where feasible and appropriate, data linkage can reduce data provider and respondent burden and maximise efficiency of data collection.

In Phase 2, a preliminary assessment was made of the availability of data items suitable for the linkage of person-level data sets that are available for the different education and training sectors. In Phase 3, the purpose of the preliminary scoping of the feasibility of data linkage was to examine in more detail the feasibility of using data linkage to integrate the various data collections. The focus was on the technical feasibility of data linkage based on the data items available from the in-scope data collections and the type of linkage that could potentially be undertaken using these.

Prior to undertaking data linkage, a number of issues need to be considered including privacy and legislative issues such as ethics approval requirements, data custodian approvals, receipt and storage of data and agreement on publications of analysis resulting from the linked data. Access to the linked data for external researchers is another consideration for some data linkage projects. These issues need to be considered when obtaining data from multiple sources at the national and jurisdictional levels for the purpose of linkage, and can take considerable time to navigate and resolve (for further details, see AIHW 2014a). These other issues that affect linkage projects were considered briefly, excluding legislative frameworks that may affect linkage protocols and data dissemination. Further information on the linkage of Commonwealth data for statistical and research purposes is available from the National Statistical Service website (National Statistical Service 2015).

This section presents a summary of the preliminary scoping of the feasibility of data linkage.

Data linkage methods

Data linkage is a powerful tool for identifying multiple appearances of individuals within a data set and for linking the information across data sets. Because the information recorded for an individual may vary from data set to data set – due to either differences in reporting (for example, in first name) or errors – a robust linkage process should allow for some discrepancy in characteristics.

There are 2 main types of data linkage that are relevant for the data sets considered here:

- key-based record linkage (also known as deterministic record linkage), in which the linkage of records is based on exact agreement of the linkage variables. Variation in reporting can be allowed for by using a number of different keys
- probabilistic record linkage, in which the linkage of records in 2 files is based on the probabilities of agreement and disagreement between a range of linkage variables. Probabilistic linkage allows for variation in reporting by allowing probabilities of agreement to be less than 1 and probabilities of disagreement to be greater than 0.

Key-based linkage is commonly used when neither a unit identifier nor full name data are available, but there are other data items that in combination can be used to link records. Probabilistic linkage is generally used when full name information, along with other demographic data, is available. Because not all of the education and training data sets contain full name information, the type of linkage to be used would depend on the 2 data sets being linked.

Privacy protection

The protection of the privacy of individuals is very important when linking data sets. There are several aspects to consider throughout a data linkage project to preserve privacy and confidentiality, including legislative frameworks, governance arrangements, ethics committee approval (if required), use of identifying and analysis variables, managing risks around dissemination of data and archiving of data.

The High Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes are also a key consideration.

Two main ways of protecting privacy during data linkage include:

1. excluding full name and possibly other identifying items from the data sets being linked (de-identifying)
2. using processes that ensure that only identifying data, and other items which assist accurate linkage, are included in the data sets used for record linkage, and that identifying data are excluded from analysis data sets (the separation principle).

Linkage feasibility

Many of the collections being considered for linkage have either given and family name information or the data necessary to construct SLK-581. Consequently, it is expected that many of the data collections could be linked to provide data sets for statistical analysis. However, the staff and provider National ECEC Workforce Census collections have no data suitable for linking and so are unable to be linked either to each other or to other collections. Also, the NSSC is primarily an aggregate data collection; however, continued engagement with stakeholders may lead to an increased quantity of data being provided at the unit record level for government schools in the future. Among the student-level collections, those carried out for international comparisons (PISA, TIMSS and PIRLS) have no name data, which limits linkage possibilities for these collections.

Apart from name data, a number of collections have person-level identifiers that may be useful when linking within-sector data collections. In particular, student identifier (ID) could be such a possible candidate for the collections covering school students (although the consistency of a student's ID across schools, sectors and over time needs to be determined).

Within the higher education and VET sectors, respectively, CHESSN (Commonwealth Higher Education Student Support Number, note that this is not available for all students) and USI (unique student identifier, introduced for all students undertaking nationally accredited training from 1 January 2015) may be useful for linking across years.

Data consistency

The consistency with which linkage variables are reported across the data sets affects how accurately linkage processes can match records for the same person across collections. Inconsistencies can result from simple reporting or processing errors (for example, in date of birth and sex). However, some variables can be expected to change over time. Three areas that have been identified as likely to impact on data linkage between education and training collections include:

- change of residential address
- consistency of reported name
- stability of person identifiers.

4.4 Project finalisation

A paper on Phase 3 of the project was provided to the DSG in October 2014 for consideration as part of the DSG forward work program. This paper included the following items:

- national data standards strategy
- implementation plan
- assessment of the consistency and comparability of core data items with other sectors (health, community services and housing and homelessness)
- preliminary scoping report on the feasibility of data linkage
- project report (this report).

Appendix A: Stakeholder consultation

Table A1 lists the organisations consulted during the project. Representatives from these organisations also represent a range of committees and groups, including the Data Strategy Group, Data and Performance Measurement Principal Committee, Early Childhood Data Sub Group, Learner Identity Management Subgroup, National Senior Officials Committee and the former Strategic Cross-sectoral Data Committee.

Table A1: Organisations consulted during the development of the national data standards strategy

Organisation
Australian Bureau of Statistics
Australian Capital Territory Education and Training Directorate
Australian Children's Education and Care Quality Authority
Australian Council for Educational Research
Australian Curriculum, Assessment and Reporting Authority
Australian Government Department of Education and Training
Australian Government Department of Industry
Australian Institute for Teaching and School Leadership
Australian Workforce and Productivity Agency
Board of Studies, Teaching and Educational Standards NSW
Catholic Education Commission (National, NSW, Qld and SA)
COAG Reform Council
Education Services Australia
Independent Schools Council of Australia
McGaw Group
National Advisory for Tertiary Education, Skills and Employment (Secretariat to SCOTese)
National Centre for Vocational Education Research
National Schools Interoperability Program (NSIP)
New South Wales Department of Education and Communities
Northern Territory Department of Education and Children's Services
Northern Territory Department of Education and Training
Productivity Commission
Queensland Department of Education, Training and Employment
School Curriculum and Standards Authority (WA)
Skills Tasmania
South Australian Department for Education and Child Development
South Australian Department of Further Education, Employment, Science and Technology
Standing Council on School Education and Early Childhood
Tasmanian Department of Education

(continued)

Table A1 (continued): Organisations consulted during the development of the national data standards strategy

Organisation
Telethon Institute for Child Health Research
Universities Australia
University of Canberra
Victorian Department of Education and Early Childhood Development
Western Australian Department of Education
Western Australian Department of Training and Workforce Development

Appendix B: Data mapping

This appendix presents the results of the detailed mapping of data items across the in-scope national education and training data collections. The purposes of this exercise were to:

- review the consistency and alignment of data items between collections and with existing national standards
- analyse commonalities, gaps and areas for improvement
- assess the fitness for purpose for data linkage and identify areas for data development.

Data items were assessed for consistency with national standards where they exist and are relevant, and similarities and differences between the data collections were also identified. Variations in data definitions and response categories can affect the comparability of data available for cross-sectoral analysis. Variations may be necessary to accurately reflect the nature of a specific data collection and the purpose for which it was developed. This may not be a problem where data reported according to different response categories can be mapped to a common higher level of coding that is meaningful across the data collections of interest. However, if variations between response categories or underlying definitions result in data that cannot be mapped or compared then the capacity of data collections to support policy, planning and performance monitoring across sectors may be compromised.

This appendix discusses the methods and results of the data mapping overall for all of the core data items identified, and also with respect to the following categories:

- core socio-demographic data items
- core education data items
- data items that are integral for data linkage
- disability status and other health data items
- family composition data items.

Data collections included in mapping

Of the 15 education and training data collections within scope of this project, 13 were included in the data mapping exercise.

The National Schools Attendance Collection was excluded as it is an aggregate rather than a unit record level (URL) collection. Similarly, only the URL components of the National Early Childhood Education and Care Collection (NECECC) and the NSSC have been included. The NSSC currently operates predominantly as an aggregate data collection, rather than URL collection.

The National ECEC Workforce Census was initially included in the mapping, but the vast majority of data items selected for mapping were not comparable with the other collections, due to the National ECEC Workforce Census being at the service, staff and worker levels rather than the child/student level. It was therefore excluded from the final data mapping.

The NAPLAN and the NAP Sample Assessments, and the TIMSS and PIRLS, collect the same core data items and have therefore been grouped and counted as single collections for data mapping purposes (that is, NAPLAN and NAP Sample Assessments have been counted as 1 and TIMSS and PIRLS have been counted as 1). The data mapping discussion therefore refers to 11 data collections.

Methods

A list of 53 key data items was developed in consultation with the SCDC (Table 2.2) during Phase 1 of the project. The next step involved obtaining metadata about the data collections, the majority of which was sourced online. National standards were sourced from METeOR – an online metadata registry for health, community services, early childhood education and housing and homelessness metadata – via the AIHW website: <www.aihw.gov.au>.

These national standards were used to provide a single comparator for each data item, so that all education and training data standards were compared with the same standard. This highlights the differences and similarities in how the data items are collected between the education and training data collections and allows the information to be summarised against a common base. Given that the data items assessed are predominantly core data items that are relevant to multiple sectors, this was considered to be a reasonable approach. However, it is recognised that data items from education and training data collections that are not collected consistently with the national standards used in this comparison are collected to agreed standards within the specific data collection or education and training sector (for example, within the school or VET sector).

An assessment algorithm (Figure B1) describes the process used to assess consistency with national standards and between data collections. A data item must be clear, concise, unambiguous, comprehensive and provide sufficient information to ensure that all those who collect, provide, analyse and use the data, clearly understand its meaning. If a data item does not meet these criteria, it is not possible to assess for consistency with national standards. In these cases, data items in the mapping exercise have been graded as ‘unable to determine’.

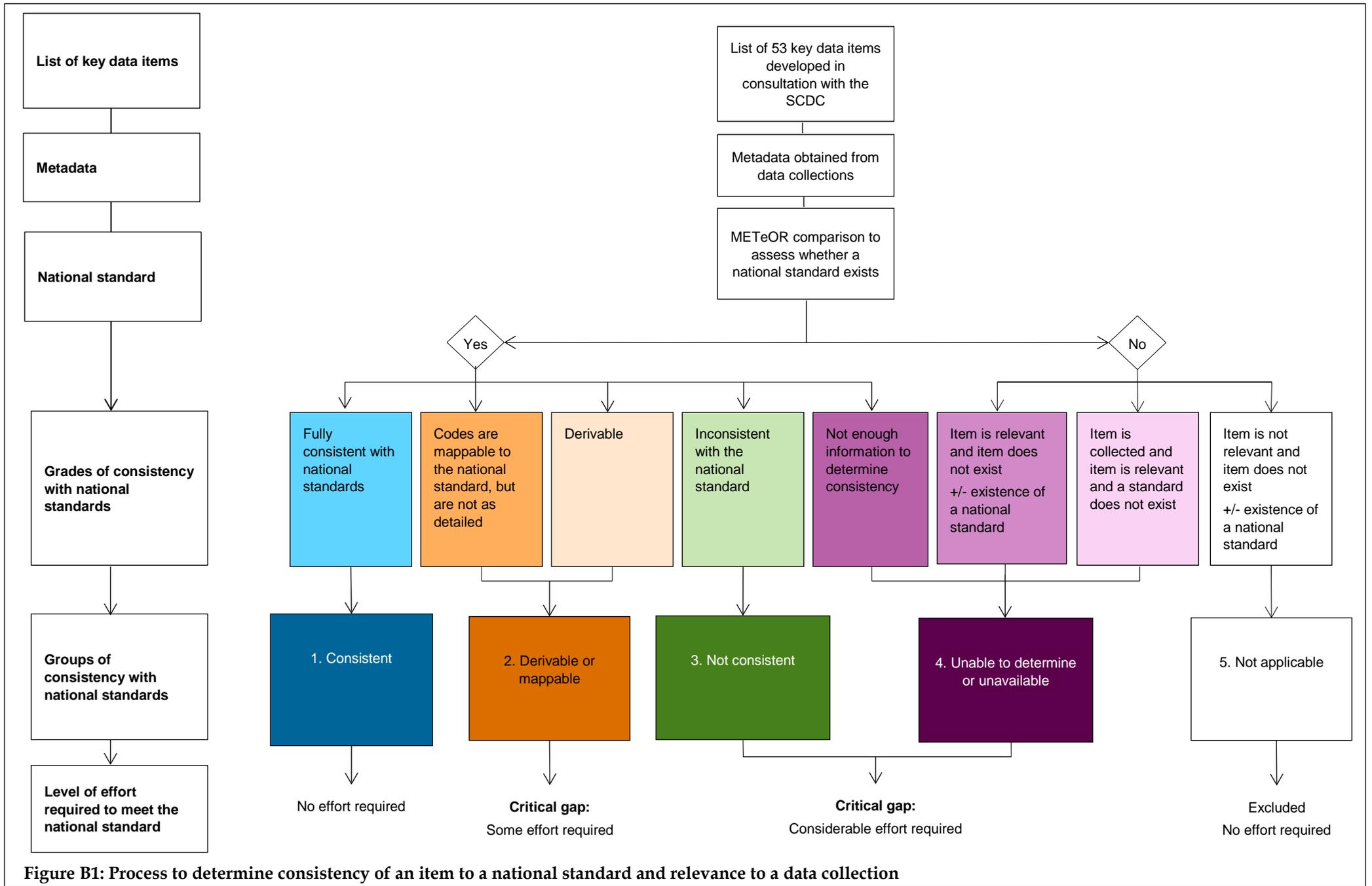
The 2 key attributes or fields of a data item used as the basis for assessing consistency were:

- definition (what is it you want to know?)
- code set/response categories (what is the range of possible responses?)

The assessment was undertaken by the AIHW based primarily on publicly available data collection documentation. However, not all of the data collections examined as part of this project had associated documentation (for example, data dictionaries or data guides). Rather than excluding these collections from the analysis, they were assessed on the basis of the question and response categories where this information was available.

The data item information that was used for the data mapping is included in Appendix C, along with other key information for the in-scope data collections.

Data custodians reviewed and provided feedback on the data mapping, which has been incorporated into this report.



Grades of consistency with national standards

Eight grades of consistency with national standards were used to assess the education and training data items with those in the national data dictionaries. These 8 grades were used to create the detailed tables in this appendix:

- fully consistent with national standard
- mappable to national standard
- derivable
- inconsistent
- not enough information to determine consistency
- item is relevant but does not exist in collection
- item is collected but no standard exists
- not applicable.

These grades were then aggregated to 5 groups of consistency to provide summary information (Table B1 and described below).

An effort indicator was then applied to classify the level of effort required to align data items with national standards. Items that require effort ('some effort required' and 'considerable effort required') to align data items with national standards identify gaps in consistency and comparability between data items.

Data collections that had items that were derivable or mappable would need some effort to attain full consistency with a national standard, whereas inconsistent or non-existent but relevant data items would need considerably more effort.

Items that were graded as 'Not applicable' were excluded from further assessment.

Table B1: Groups (5) of consistency

	Group	Meaning	Effort indicator (to meet the national standard)
	1. Consistent	Fully consistent with national standards	No effort required
Critical gaps	2. Derivable or mappable	Codes are mappable to the national standard, but are not as detailed as the national standard, or data items are not collected but may be derivable from another item	Some effort required
	3. Not consistent	Inconsistent with the national standard	Considerable effort required
	4. Unable to determine or unavailable	Unable to determine consistency, or item does not exist but may be relevant to the collection, or national standard does not exist	Considerable effort required
	5. Not applicable	Data items that are not collected and not relevant to a particular collection	Not applicable

1. Consistent group

This group contains items that are fully consistent and comparable with the national data standard. As such, all attributes of the data item are entirely consistent with the most recent version of the relevant national standard, and the codes are identical to or can be aggregated to the highest level of the national standard. There would be no effort required to meet the national standard as it has already been attained and is ongoing.

2. Derivable or mappable group

This group means that the national standard defines codes that are more detailed than, or has some codes/response categories that are not relevant to, the data collection. Ideally, the coding categories should be mappable to at least the highest level of the Data domain in the national standard. This should not prevent data collections from using only those coding categories that are relevant to the collection, provided they are codes that can be ‘mapped’ or ‘aggregated’ to at least the highest level of the codes in the relevant national standard.

In addition, some data items may not be collected, but can be derived from another data item that is collected and mapped to the national standard. Some effort would be required to map or derive these data items.

3. Not consistent group

These data items have definitions and/or codes that are not consistent in meaning with relevant national or international standards, and will not yield data that is comparable to other data sets. Data items have been coded as not consistent only in situations where it is clear a relevant national standard exists that captures similar information to that defined in the data item.

4. Unable to determine or unavailable

Includes:

- **Not enough information to determine consistency**

A data item should be clear, concise, unambiguous, comprehensive and provide sufficient information to ensure that all those who collect, provide, analyse and use the

data clearly understand its meaning. If a data item does not meet these criteria then it has been graded as 'unable to determine'.

- **Item is relevant, but does not exist in collection**

This category means that the data item is potentially relevant to the collection, but is not part of the collection. These are areas where it could be considered whether it would be useful and worthwhile to add these data items to collections or obtain via data linkage; or it may indicate that it is not possible or appropriate to collect or derive this data item (for example, for legislative or privacy reasons).

- **Item collected but standard does not exist**

Data items that have no existing relevant national standard in METeOR may represent areas where a national standard is needed. However, the absence of a national standard does not always imply a need for one, as some data elements are specific to a particular data collection and may not be applicable to others.

5. Not applicable

This group comprises data items that have existing national data standards, but are not applicable to a particular education and training data collection (and are not collected in that collection). Items that were graded as 'Not applicable' were excluded from further assessment.

Detailed data mapping

Details of the metadata used to inform the mapping can be found in Appendix C, along with further descriptions of each data collection and details of the data items that have been mapped.

Table B2 summarises the number of data items mapped to each of the grades of consistency categories. This information has been used to construct the figures in the results section of this appendix.

Table B3 provides the detailed results of the data mapping for each data item across the 11 data collections.

The key on the following page shows the colour coding used in Tables B2 and B3, and is followed by a description of the categories.

Key	Meaning
Data item name	National data standard and METeOR ID
Data item name	Fully consistent with national standards
Data item name	Codes are mappable to the national standard, but are not as detailed as the national standard
Data item name	Derivable
Data item name	Inconsistent with the national standard
Data item name	Not enough information to determine consistency
Data item name	Item is relevant but does not exist in collection (a national standard may or may not exist)
Data item name	Item collected and relevant but national standard does not exist
Not applicable	Not applicable

National data standard and METeOR ID

The name and METeOR ID of the national data standard, where one exists for the data item. This is the standard that the rest of the categories are mapped to.

Fully consistent with national standards

All attributes of the data item are entirely consistent with the most recent version of the relevant national standard, and the codes are identical to, or can be aggregated to, the highest level of the national standard.

Codes are mappable to the national standard, but are not as detailed

This category means that the national standard defines codes that are more detailed than, or has some codes/response categories that are not relevant to, the data collection. Ideally, the coding categories should be mappable to at least the highest level of the Data domain in the national standard. This should not prevent data collections from using only those coding categories that are relevant to the collection, provided they are codes that can be 'mapped' or 'aggregated' to at least the highest level of the codes in the relevant national standard.

Derivable or mappable

Data items that are not collected, but can be derived from another data item that is collected, and mapped to the national standard.

Inconsistent with national standards

These data items have definitions and/or codes that are not consistent in meaning with relevant national or international standards, and will not yield data that is comparable to other data sets. Data items have been coded as not consistent only in situations where it is clear a relevant national standard exists that captures similar information to that defined in the data item.

Not enough information to determine consistency

A data item should be clear, concise, unambiguous, comprehensive and provide sufficient information to ensure that all those who collect, provide, analyse and use the data clearly understand its meaning. If a data item does not meet these criteria then it has been graded as 'unable to determine'.

Item is relevant, but does not exist in collection

This category means that the data item is potentially relevant to the collection, but is not part of the collection. These are areas where it could be considered whether it would be useful and worthwhile to add these data items to collections; or it may indicate that it is not possible to collect or derive this data item (for example, for legislative or privacy reasons). A national standard may or may not exist for data items in this category.

Item collected but standard does not exist

Data items that have no existing relevant national standard may represent areas where a national standard is needed. However, the absence of a national standard does not always imply a need for one, as some data elements are specific to a particular data collection and would not be applicable to others.

Not applicable

Data items that have existing national standards, but are not applicable to a particular data collection (and not collected in that collection).

Table B2: Summary of mapping of education and training data collections with national standards (number and per cent of data items selected for mapping)

Data item status	AEDC		Apprentices		Census		Higher Education		NAP ^(a)		NECECC		NSSC ^(b)		PISA		TIMSS & PIRLS		VET Provider		VET in Schools		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Socio-demographic data items																								
Fully consistent	8	30	7	26	10	37	4	15	3	11	10	37	6	22	0	0	0	0	11	41	9	33	68	23
Mappable	2	7	2	7	4	15	6	22	3	11	0	0	2	7	6	22	4	15	2	7	2	7	33	11
Derivable	0	0	4	15	3	11	6	22	2	7	4	15	8	30	1	4	1	4	3	11	3	11	35	12
Inconsistent	1	4	0	0	1	4	1	4	2	7	0	0	1	4	4	15	5	19	0	0	0	0	15	5
Unable to determine	6	22	0	0	2	7	0	0	9	33	0	0	0	0	0	0	1	4	0	0	0	0	18	6
Item is relevant	6	22	5	19	4	15	1	4	3	11	9	33	5	19	9	33	9	33	1	4	8	30	60	20
Collected but no standard	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Not applicable	4	15	9	33	3	11	9	33	5	19	4	15	5	19	7	26	7	26	10	37	5	19	68	23
<i>Total items mapped</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>27</i>	<i>100</i>	<i>297</i>	<i>100</i>
Educational data items																								
Fully consistent	1	8	1	8	2	15	3	23	1	8	4	31	4	31	1	8	1	8	5	38	5	38	28	20
Mappable	2	15	2	15	1	8	1	8	4	31	0	0	2	15	2	15	1	8	3	23	3	23	21	15
Derivable	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Inconsistent	0	0	0	0	0	0	1	8	1	8	0	0	0	0	0	0	1	8	0	0	0	0	3	2
Unable to determine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	23	3	23	0	0	0	0	6	4

(continued)

Table B2 (continued): Summary of consistency of education and training data collections with national standards (number and per cent of data items selected for mapping)

Data item status	AEDC		Apprentices		Census		Higher Education		NAP ^(a)		NECECC		NSSC ^(b)		PISA		TIMSS & PIRLS		VET Provider		VET in Schools		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Educational data items (continued)																								
Item is relevant	2	15	8	62	0	0	5	38	2	15	2	15	3	23	2	15	2	15	2	15	3	23	31	22
Collected but no standard	4	31	1	8	2	15	2	15	1	8	3	23	2	15	1	8	1	8	2	15	2	15	21	15
Not applicable	4	31	1	8	8	62	1	8	4	31	4	31	2	15	4	31	4	31	1	8	0	0	33	23
<i>Total items mapped</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>13</i>	<i>100</i>	<i>143</i>	<i>100</i>
Disability data items																								
Fully consistent	0	0	0	0	1	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5
Mappable	1	50	1	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	50	1	50	4	18
Derivable	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Inconsistent	1	50	0	0	0	0	2	100	0	0	0	0	0	0	0	0	0	0	1	50	1	50	5	23
Unable to determine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Item is relevant	0	0	1	50	1	50	0	0	2	100	2	100	2	100	2	100	2	100	0	0	0	0	12	55
Collected but no standard	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Not applicable	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Total items mapped</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>2</i>	<i>100</i>	<i>22</i>	<i>100</i>

(continued)

Table B2 (continued): Summary of consistency of education and training data collections with national standards (number and per cent of data items selected for mapping)

Data item status	AEDC		Apprentices		Census		Higher Education		NAP ^(a)		NECECC		NSSC ^(b)		PISA		TIMSS & PIRLS		VET Provider		VET in Schools		Total			
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%		
Family composition data items																										
Fully consistent	0	0	0	0	1	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	
Mappable	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	100	0	0	0	0	0	0	0	1	9
Derivable	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Inconsistent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Unable to determine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Item is relevant	1	100	1	100	0	0	1	100	1	100	1	100	1	100	0	0	1	100	1	100	1	100	1	100	9	82
Collected but no standard	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Not applicable	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Total items mapped</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>1</i>	<i>100</i>	<i>11</i>	<i>100</i>		
Statistical linkage data items																										
Fully consistent	2	20	4	40	3	30	4	40	2	20	8	80	7	70	1	10	1	10	5	50	4	40	41	37		
Mappable	1	10	3	30	1	10	3	30	1	10	0	0	0	0	1	10	2	20	3	30	3	30	18	16		
Derivable	0	0	0	0	1	10	3	30	0	0	1	10	2	20	0	0	0	0	0	0	0	0	7	6		
Inconsistent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	10	0	0	0	0	0	0	1	1		
Unable to determine	7	70	0	0	1	10	0	0	7	70	0	0	0	0	2	20	3	30	0	0	0	0	20	18		

(continued)

Table B2 (continued): Summary of consistency of education and training data collections with national standards (number and per cent of data items selected for mapping)

Data item status	AEDC		Apprentices		Census		Higher Education		NAP ^(a)		NECECC		NSSC ^(b)		PISA		TIMSS & PIRLS		VET Provider		VET in Schools		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Item is relevant	0	0	3	30	3	30	0	0	0	0	1	10	1	10	5	50	4	40	1	10	2	20	20	18
Collected but no standard	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Not applicable	0	0	0	0	1	10	0	0	0	0	0	0	0	0	0	0	0	0	1	10	1	10	3	3
<i>Total items mapped</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>10</i>	<i>100</i>	<i>110</i>	<i>100</i>
All data items																								
Fully consistent	11	21	12	23	17	32	11	21	6	11	22	42	17	32	2	4	2	4	21	40	18	34	139	24
Mappable	6	11	8	15	6	11	10	19	8	15	0	0	4	8	10	19	7	13	9	17	9	17	77	13
Derivable	0	0	4	8	4	8	9	17	2	4	5	9	10	19	1	2	1	2	3	6	3	6	42	7
Inconsistent	2	4	0	0	1	2	4	8	3	6	0	0	1	2	5	9	6	11	1	2	1	2	24	4
Unable to determine	13	25	0	0	3	6	0	0	16	30	0	0	0	0	5	9	7	13	0	0	0	0	44	8
Item is relevant	9	17	18	34	8	15	7	13	8	15	15	28	12	23	18	34	18	34	5	9	14	26	132	23
Collected but no standard	4	8	1	2	2	4	2	4	1	2	3	6	2	4	1	2	1	2	2	4	2	4	21	4
Not applicable	8	15	10	19	12	23	10	19	9	17	8	15	7	13	11	21	11	21	12	23	6	11	104	18
Total items mapped	53	100	53	100	53	100	53	100	53	100	53	100	53	100	53	100	53	100	53	100	53	100	583	100

- (a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.
- (b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Note: Components may not sum to totals due to rounding.

Source: Appendix Table B3.

Table B3: Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Socio-demographic data items												
Age	Person—age 303794	Age	Derivable from date of birth	Derived from date of birth, where provided, otherwise stated age is used.	Derivable from date of birth	Derivable from date of birth	Age	Derivable from date of birth	Derivable from birth month and birth year	Derivable from date of birth items; birth day, birth month, birth year	Derivable from date of birth	Derivable from date of birth
Country of birth	Person—country of birth (Standard Australian Classification of Countries 2011) 459973	Country of birth (SACC)	Country identifier (SACC)	Country of birth of person (SACC)	Country of birth code (SACC)	Held at jurisdiction's test administration authority (TAA) (no further information)		Student's country of birth (SACC)	Country of birth national categories—self	Born in Australia	Country identifier (SACC)	Country identifier (SACC)
Date of birth	Person—date of birth 287007	Date of birth	Date of birth	Not available on the output dataset. Only Age in years is available.	Date of birth	Date of birth	Date of birth	Date of birth	Birth month Birth year	Date of student birth day Date of student birth month Date of student birth year	Date of birth	Date of birth

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Socio-demographic data items (continued)												
Geography address— address line	Person (address)— address line 286620	Street address		Apartment/Flat/Unit number Street number Street name ^(g)	Address of permanent home residence— Part 1 Address of permanent home residence— Part 2 (not available for all students) ^(g)	Held at jurisdiction's TAA (no further information)	Address line				Address building/ property name Address flat/unit details Address street number Address street name ^(g)	
Geography address— suburb	Address— suburb/town/locality name 429889	Local community	Address location— suburb, locality or town	Suburb/ locality	Name of Suburb/ Locality (not available for all students) ^(g)	Held at jurisdiction's TAA (no further information)	Suburb/ town/locality name				Address location— suburb, locality or town	Address location— suburb, locality or town
Geography address— state	Person—Australian state/territory identifier (ASGS) 286919	State ID	State identifier	State/ territory	Residential address— state	Held at jurisdiction's TAA (no further information)	Australian state/ territory identifier	Derivable from area of usual residence Mesh Block			State identifier	State identifier
Geography address— postcode	Address— Australian postcode 429894	Postcode	Postcode	Postal area (allocated based on SA1)	Address of permanent home residence postcode	Held at jurisdiction's TAA (no further information)	Australian postcode	Derivable from area of usual residence Mesh Block			Postcode	Postcode

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Socio-demographic data items (continued)												
Geography ASGS— Mesh Block ^(c)	Address— Mesh Block, code (ASGS 2011)	Derivable from street address, local community and state ID		Mesh Block	Derivable	Potentially derivable from information held at jurisdiction's TAA (no further information)	Derivable from address line, suburb and state	Area of usual residence Mesh Block			Not applicable ^(h)	Not applicable ^(h)
Geography ASGS— SA1	Address— statistical area SA1 (ASGS, 2011) 457287	Derivable from street address, local community and state ID		Derivable from Mesh Block	Derivable	Potentially derivable from information held at jurisdiction's TAA (no further information)	Statistical area level 1	Derivable from area of usual residence Mesh Block			Statistical area level 1 identifier ^(h)	^(h)
Geography ASGS— SA2	Address— statistical area SA2 (ASGS, 2011) 457289	Derivable from street address, local community and state ID	Derivable from postcode and suburb	Derivable from Mesh Block	Derivable	Potentially derivable from information held at jurisdiction's TAA (no further information)	Derivable from SA1	Derivable from SA1 and area of usual residence Mesh Block			Statistical area level 2 identifier	Derivable from postcode and suburb

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Socio-demographic data items (continued)												
Geography ASGC – Collection District	Person— area of usual residence, geographic location CD code (ASGC) 455512	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Area of usual residence (until 2013)	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Geography ASGC—SLA	Person— geographic location SLA (ASGC 2011) 455542	SLA code	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Geography— remote-ness	Person— geographic remoteness, classification (ASGS-RA) 569240	Local Community ARIA category	Derivable from postcode and suburb	Remoteness	Derivable from postcode	Geolocation (MCEECDYA Geographical Location Classification)	Derivable from SA1	Derivable from area of usual residence Mesh Block	Geographic location (MCEECDYA school's location classification)	Geographic location (MCEECDYA school's location classification)	Derivable from postcode and suburb	Remoteness (ARIA+) region (derived from SA2 which is derived from postcode and suburb)
Income	Person— gross income 302481	Not applicable	Not applicable	Total personal income (weekly)	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Indigenous status	Person— Indigenous status 291036	ATSI type	Indigenous status identifier	Indigenous status	Aboriginal and Torres Strait Islander code	Indigenous status	Indigenous status	ATSI status	Indigenous background	Indigenous background	Indigenous status identifier	Indigenous status identifier

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Socio-demographic data items (continued)												
Language spoken at home	Person—main language spoken at home (ASCL) 460125	Language ID (ASCL)	Language identifier (ASCL)	Language spoken at home (ASCL)	Language spoken at home code (ASCL)	Main language other than English spoken at home (ASCL)		Student's main language other than English spoken at home (ASCL)	Language at home	Often speak English at home	Language identifier (ASCL)	Language identifier (ASCL)
Occupation	Person—occupation (main) (ANZSCO 2006) 350899 Person—occupation (main) (ANZSCO 2013) 566159	Not applicable	ANZSCO identifier	Occupation	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	ANZSCO identifier	ANZSCO identifier
Parents' country of birth	Person—country of birth (Standard Australian Classification of Countries 2011) 459973		Not applicable	Birthplace of male parent Birthplace of female parent	Not applicable				Country of birth International–Mother Country of birth International–Father	Mother born in Australia Father born in Australia	Not applicable	

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Socio-demographic data items (continued)												
Parent's highest level of qualification	Person— level of highest educational attainment (ASCED, 2001) 321069		Not applicable		Highest educational attainment by parent/guardian 1 Highest educational attainment by parent/guardian 2 (not available for all students)	Derivable from Parent/Guardian school education items and Parent/Guardian non-school qualification items		Derivable from Parent/Guardian school education items and Parent/Guardian non-school qualification items	Educational level of Mother (ISCED) Educational level of Father (ISCED)	Highest level of education Mother (ISCED) Highest level of education Father (ISCED)	Not applicable	
Parent's highest school level completed	Person— highest year of school completed 375998		Not applicable		Not applicable	Parent/Guardian 1 school education Parent/Guardian 2 school education		Parent/Guardian 1 school education Parent/Guardian 2 school education	Not applicable	Not applicable	Not applicable	
Parent's highest non-school qualification	Person— level of highest non-school qualification 398769		Not applicable		Not applicable	Parent/Guardian 1 non-school education Parent/Guardian 2 non-school education		Parent/Guardian 1 highest non-school qualification Parent/Guardian 2 highest non-school qualification	Not applicable	Not applicable	Not applicable	

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Socio-demographic data items (continued)												
Parent occupation	Person—occupation (main) (ANZSCO 2006) 350899		Not applicable		Not applicable	Parent/ guardian 1 occupation group Parent/ guardian 2 occupation group		Parent/ guardian 1 occupation group Parent/ guardian 2 occupation group	Mother occupation (ISCO) Father occupation (ISCO)	Mother occupation Father Occupation	Not applicable	
Proficiency in English	Person—proficiency in spoken English 270203	Proficiency in English		Proficiency in spoken English							Proficiency in spoken English identifier	Proficiency in spoken English identifier (optional)
Sex	Person—sex 287316	Gender	Sex	Sex	Gender code	Sex	Sex	Sex	Sex	Sex of student	Sex	Sex
SES for area	Socio-Economic Indexes for Areas (SEIFA) cluster 2011 477111 ^(d)	Local Community SEIFA score Category	Derivable from SA2 (derived from postcode and suburb)	Derivable from postcode (based on SA1)	Derivable from Address of permanent home residence postcode	Potentially derivable from information held at jurisdiction's TAA (no further information)	Derivable from SLA (ASGC) (2013 and earlier)/SA1 (ASGS) (2014 onwards)	Derivable from area of usual residence Mesh Block	Index of economic, social and cultural status	Economic disadvantage	Derivable from SA2	Derivable from SA2 (derived from postcode and suburb)
SES for person/ family		Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Socio-demographic data items (continued)												
Year of arrival	Person—year of first arrival in Australia 269929			Year of arrival in Australia	Year of arrival in Australia							
Education data items												
Educational institution ID/name	Service provider organisation—organisation identifier 414987	School ID	Training organisation identifier	Not applicable	Higher education provider code	School ID School name	Service provider organisation identifier/ name	System school ID	School ID	School ID	Training organisation identifier	Training organisation identifier
Educational institution address—address line	Service provider organisation (address)—address line 290315			Not applicable		Street name	Service provider organisation—Address line	School Address— Apartment/ Flat number School Address— Street number School Address— Street name	School address ^(g)	School address ^(g)	Address first line Address second line	Address first line Address second line
Educational institution—suburb	Address—suburb/town/locality name 429889	School suburb		Not applicable	Campus location (not available for all students)	School suburb	Suburb/town/locality name	School address—suburb	School address ^(g)	School address ^(g)	Address location – suburb, locality or town	Address location – suburb, locality or town
Educational institution address—state	Service provider organisation—state/territory 289083	School state ^(e)		Not applicable		State ^(e)	Service provider organisation—Australian state/territory identifier	State/territory	State	State	State identifier	State identifier

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Education data items (continued)												
Educational institution address—postcode	Address—Australian postcode 429894			Not applicable	Postcode or overseas country code location of Higher Education/ VET Provider campus/ delivery location	School postcode		School address—postcode	School address ⁽⁹⁾	School address ⁽⁹⁾	Postcode	Postcode
Educational institution – concurrent		Dual placement		Not applicable							VET in Schools flag	
Educational institution—previous		Non-parental care/other educational program before entering school		Not applicable			Preschool program repeater indicator					
Educational institution type		School type	Not applicable	Type of educational institution attending	Not applicable	Sector	Service activity type	School affiliation	School sector	School sector	Not applicable	School type (parent school)
Attendance or enrolment status		Attendance various items	Full-time identifier	Full-time/part-time student status	Type of attendance code	Not applicable	Preschool program enrolment/ attendance indicator	Full-time or Part-time Status of Enrolment	Not applicable	Not applicable		

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Education data items (continued)												
Field of education	Person—field of education (ASCED, 2001) 270172	Not applicable		Non-school qualification: Field of study (ASCED)	Field of education code (Field of Education Classification)	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Subject field of education identifier (6-digit ASCED) Program field of education identifier (4-digit ASCED)	Subject field of education identifier (6-digit ASCED)
Educational achievement (results)		Not applicable		Not applicable	Unit of study completion status	Not applicable	Not applicable		Not applicable	Not applicable	Outcome identifier—national	Outcome identifier – national
Highest school level completed	Person—highest year of school completed 375998	Not applicable	Highest school level completed	Highest year of school completed		Student grade level	Not applicable	Grade of student's enrolment	Grade	Grade ID	Highest school level completed	Highest school level completed
Highest non-school qualification	Person—level of highest non-school qualification 398769	Not applicable	Prior educational achievement identifier	Non-school qualification: Level of study	Highest educational participation prior to commencement	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Prior educational achievement identifier	Prior educational achievement identifier

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Disability data items												
Special needs/ Disability flag	Person— disability status 312934 Person— activity limitation indicator 533815	Special needs status	Disability flag	Core activity need for assistance	Disability						Disability flag	Disability flag
Special needs/ Disability type	Person— disability group 337532	Behaviour problem Chronic illness Emotional problem Hearing impairment Learning disability Physical disability Speech impairment Visual impairment Other enduring problems			Disability						Disability type identifier	Disability type identifier

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Family composition data items												
Family composition	Household—family composition 343206			Family composition					Family structure			
Statistical linkage data items^(f)												
Date of birth	Person—date of birth 287007	Date of birth	Date of birth	Not available on the output dataset. Only Age in years is available.	Date of birth	DOB	Date of birth	Date of birth	Birth month Birth year	Date of students birth day Date of students birth month Date of students birth year	Date of birth	Date of birth
Identifier code—non-SLK	Person—person identifier 290046	Student ID	Client identifier – apprenticeship Client identifier – TYIMS	Not applicable	Commonwealth Higher Education Student Support Number (not available for all students)	Student ID ⁽ⁱ⁾		System student ID	Student ID	Student ID	Client identifier – apprenticeships	Client identifier

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Statistical linkage data items^(f) (continued)												
Identifier code— SLK	Record linkage key, code 581 349895	Derivable from date of birth, family name, given name and gender			Derivable from date of birth, student surname, student given name and gender code	Derivable from information held at jurisdiction's TAA (no further information)	Statistical linkage key 581	Statistical linkage key		Derivable from date of birth items, family name, given name and sex of student		
Name – Family name	Person (name)— family name 453750 Person—letters of family name 349481	Collected and held securely (no further information available)	Name for encryption ^(g)	⁽ⁱ⁾	Student surname (not available for all students) ^(g)	Held at jurisdiction's TAA (no further information)	Letters of Family Name	Letters of student's enrolled family name	Family name	Family name	Name for encryption ^(g)	Name for encryption ^(g)
Name – Given name	Person (name)— given name 453734 Person—letters of given name 349483	Collected and held securely (no further information available)	Name for encryption ^(g)	⁽ⁱ⁾	Student given name— first (not available for all students) ^(g)	Held at jurisdiction's TAA (no further information)	Letters of Given Name	Letters of student's enrolled given name	Given name	Given name	Name for encryption ^(g)	Name for encryption ^(g)
Geography address – suburb	Address— suburb/town/locality name 429889	Local community	Address location- suburb, locality or town	Suburb/ locality	Name of Suburb/ Locality (not available for all students) ^(g)	Held at jurisdiction's TAA (no further information)	Suburb/town/locality name				Address location – suburb, locality or town	Address location – suburb, locality or town

(continued)

Table B3 (continued): Detailed mapping of education and training data collections with national standards

Data item	Standard name/ METeOR ID	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools
Statistical linkage data items^(f) (continued)												
Geography address—postcode	Address—Australian postcode 429894	Postcode	Postcode	Postal area (allocated based on SA1)	Address of permanent home residence postcode	Held at jurisdiction's TAA (no further information)	Australian postcode	Derivable from area of usual residence Mesh Block			Postcode	Postcode
Geography ASGS – Mesh Block ^(c)	Address—Mesh Block, code (ASGS 2011)	Derivable from address line, suburb and state		Mesh Block	Derivable	Potentially derivable from information held at jurisdiction's TAA (no further information)	Derivable from address line, suburb and state	Area of usual residence of Mesh Block code			Not applicable ^(h)	Not applicable ^(h)
Geography ASGS—SA1	Address—statistical area SA1 (ASGS, 2011) 457287	Derivable from address line, suburb and state		Derivable from Mesh Block	Derivable	Potentially derivable from information held at jurisdiction's TAA (no further information)	Statistical area level 1	Derivable from area of usual residence of Mesh Block			Statistical area level 1 identifier ^(h)	^(h)
Sex	Person—sex 287316	Gender	Sex	Sex	Gender code	Sex	Sex	Sex	Sex	Sex of student	Sex	Sex

(a) The NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

(c) Standard has been developed but has not been endorsed. It is expected that this will be adopted as a national data standard in the future.

- (d) 2006 Socio-Economic Indexes for Areas (SEIFA), soon to be superseded by 2011 SEIFA.
- (e) Address—Australian state/territory identifier 430134 is an additional standard which could be used whereby the AEDC would be fully consistent and NAP presumably mappable.
- (f) Includes some items that also appear in other categories.
- (g) These items are not available for analysis or linkage purposes.
- (h) The VET sector has previously determined that it does not require geographical reporting to below SA1 level.
- (i) Surname and first name are collected in the Census but this information is destroyed post-statistical processing unless respondents have explicitly agreed to their name-identified information being retained. This information is not available for any purpose (including analysis or data linkage), within a 99 year closed access period.
- (j) The student ID held by ACARA is different to the student ID held by schools and therefore can not be used for linkage purposes. ACARA student identifiers are not consistent over time.

Sources: ABS 2011, 2013a; ACARA 2013; ACER 2012, 2014; AIHW 2014b; Australian Government Department of Education 2013; IEA 2013a; IEA 2013b; NCVET 2013a, 2013b, 2013c; Social Research Centre 2013.

Results

Overall consistency with national data standards

Fifty-three data items covering socio-demographic, education, statistical linkage, disability and family composition categories from each of the 11 data collections were assessed. This resulted in a total of 583 data items that were assessed for consistency and mapped against national standards in METeOR, where available (Table B4).

Overall (Table B4, Figure B2):

- across all data collections, 24% (139) of data items were consistent with national standards; within a collection, consistency ranged between 4% and 42% (1–21 data items).
- across all collections, 20% (119) of data items were derivable or mappable to national standards; within a collection, between 9 and 36% (5–19) of data items were derivable or mappable. With some effort, these data items could potentially be made consistent with national standards.
- 4% (24) of data items across all collections were not consistent with national standards. Within a collection, inconsistency ranged between 0–11% (0–6) and considerable effort would be required to align with national standards.
- 34% (197) of data items across all collections were either relevant but did not exist in the data collection, did not have sufficient information to determine the level of consistency or were collected but a national standard does not exist. Within a collection, the range was 15–51% (8–27) of data items. Considerable effort would be required to further review the data items, revise to align with national standards and/or develop national standards where they do not currently exist.

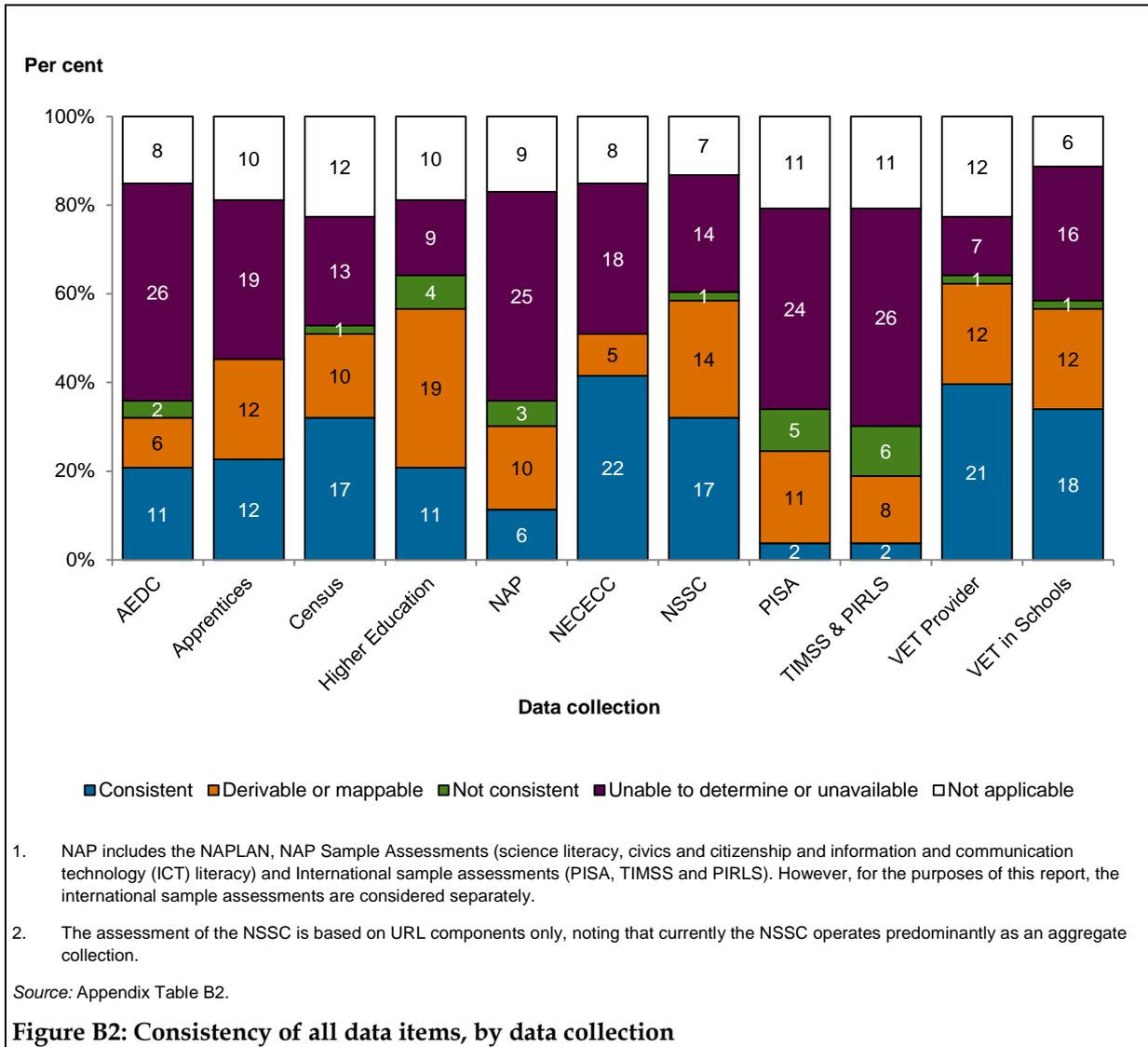


Table B4: Summary of data standard consistency across collections

Effort indicator	Group	AEDC	Apprentices	Census	Higher Education	NAP ^(a)	NECECC	NSSC ^(b)	PISA	TIMSS & PIRLS	VET Provider	VET in Schools	Total	
													No.	%
No effort required	Consistent	11	12	17	11	6	22	17	2	2	21	18	139	24
Some effort required	Derivable or mappable	6	12	10	19	10	5	14	11	8	12	12	119	20
Considerable effort required	Not consistent	2	0	1	4	3	0	1	5	6	1	1	24	4
	Unable to determine	26	19	13	9	25	18	14	24	26	7	16	197	34
Not applicable		8	10	12	10	9	8	7	11	11	12	6	104	18
Total		53	53	53	53	53	53	53	53	53	53	53	583	100

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B2.

Consistency by data item category

Data items were also examined for consistency with national standards in METeOR according to 5 specific categories:

- core socio-demographic data items
- core education data items
- disability status and other health data items
- family composition data items
- data items that are integral for data linkage.

Socio-demographic data items

Twenty-seven data items relating to socio-demographic information were assessed across the 11 data collections (297 data items in total).

Twenty-two per cent (68 items) of data items were graded as entirely consistent with national standards (Figure B3). The data items with the greatest level of consistency across the collections and with the national standards were language spoken at home, country of birth, date of birth, suburb, state and postcode, although there were still some collections that were not consistent with existing standards or did not collect these items.

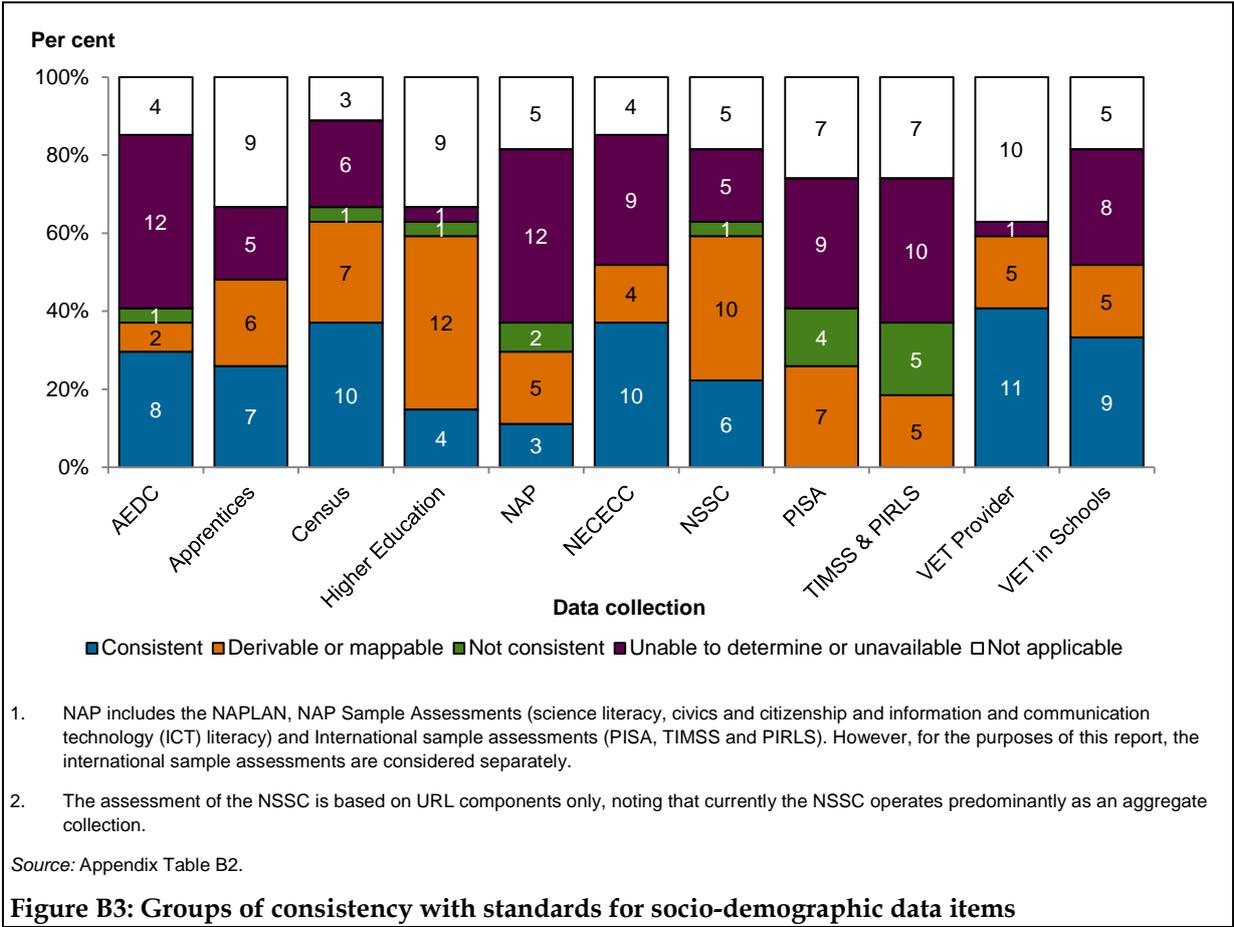


Figure B3: Groups of consistency with standards for socio-demographic data items

An additional 23% or 68 data items were assessed as able to be mapped or derived from other items to existing standards (Table B5), which would require some additional effort to fully align with national standards. Sex was consistent or mappable across all collections and

age was consistent or derivable for all collections, with 8 collections able to apply the standard to this item via date of birth items and the remaining 3 collections collecting this item separately. Date of birth appears to be consistent or mappable for all collections except PISA, where information is available. Indigenous status, while fully consistent or mappable for all mapped collections, had 7 collections that were mappable rather than fully consistent, meaning that the codes are not as detailed as the national standard.

Geography items relating to the usual residence of the child/student are derivable or mappable to national standards on some level for all collections except PISA, TIMSS and PIRLS. Where Mesh Block, postcode or suburb were collected or could be derived, state and statistical area were also coded as derivable where not separately available.

Only a small proportion of data items were assessed as not consistent with the national standards (5%) (Table B7). The data items with the greatest inconsistency with the standards were geography – remoteness, parents' country of birth and parent occupation (3 data collections were inconsistent with the standard). Twenty-six per cent (78 data items) were either non-existent, difficult to determine or were collected but a national standard did not exist. The majority of these (60 of 78) were coded as relevant to the collection, but did not exist in the collection (Table B6). For example, year of arrival in Australia, proficiency in English, parental education, parental occupation and parental country of birth data items are potentially relevant to a number of data collections, but are not collected by most. Data custodians could consider the feasibility of adding such data items to their collections or using data linkage to obtain this information where required; noting that considerable effort may be necessary to achieve this and the benefits of the information should be balanced against the burden of collecting it.

For 6% (18) of data items, there was not enough information available from the metadata for the data collection to determine consistency with the national standard.

Almost one-quarter (23%) of items were graded as not applicable to the relevant collection; including personal income and occupation information for children, parental information for older youths/adults, SES for person/family (no agreed measure exists), and geographic data items based on the Australian Standard Geographical Classification (ASGC) where the collection has transitioned to the Australian Statistical Geography Standard (ASGS) (Table B8).

In terms of individual collections, the Census of Population and Housing had the highest proportion of data items that were consistent, mappable, or able to be derived to the national data standards (63%), followed by the Higher Education Statistics Collection, the National Schools Statistics Collection and the National VET Provider Collection (59%).

Table B5: Socio-demographic data items that are derivable or mappable, by data collection

AEDC (2)	Apprentices (6)	Census (7)	Higher Education (12)	NAP^(a) (5)	NECECC (4)	NSSC^(b) (10)	PISA (7)	TIMSS & PIRLS (5)	VET Provider (5)	VET in Schools (5)
Remoteness	Age	ASGS—SA1	Age	Age	Mesh Block	Age	Age	Age	Age	Age
Sex	ASGS—SA2	ASGS—SA2	Date of birth	Parent's highest qualification	ASGS—SA2	Address—state	Country of birth	Date of birth	Remoteness	ASGS—SA2
	Remoteness	Income	Address line	Parent's highest school level	Remoteness	Address—postcode	Indigenous status	Indigenous status	Indigenous status	Indigenous status
	Indigenous status	Indigenous status	Address—suburb	Parent's highest non-school qualification	SES for area	ASGS—SA1	Language spoken at home	Parent's highest qualification	Sex	Sex
	Sex	Proficiency in English	Address—state	Sex		ASGS—SA2	Parent's highest qualification	Sex	SES for area	SES for area
	SES for area	Sex	ASGS—SA1			Remoteness	Parent occupation			
		SES for area	ASGS—SA2			Parent's highest qualification	Sex			
			Remoteness			Parent's highest school level				
			Income			Parent's highest non-school qualification				
			Indigenous status			SES for area				
			Sex							
			SES for area							

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B6: Socio-demographic data items that are relevant but do not exist in the collection

AEDC (6)	Apprentices (5)	Census (4)	Higher Education (1)	NAP^(a) (3)	NECECC (9)	NSSC^(b) (5)	PISA (9)	TIMSS & PIRLS (9)	VET Provider (1)	VET in Schools (8)
Parents' country of birth	Address line	Parent's highest qualification	Proficiency in English	Parents' country of birth	Country of birth	Address line	Address line	Address line	Year of arrival	Address line
Parent's highest qualification	Mesh Block	Parent's highest school level		Proficiency in English	Language spoken at home	Address—suburb	Address—suburb	Address—suburb		ASGS—SA1
Parent's highest school level	ASGS—SA1	Parent's highest non-school qualification		Year of arrival	Parents' country of birth	Parents' country of birth	Address—state	Address—state		Parents' country of birth
Parent's highest non-school qualification	Proficiency in English	Parent occupation			Parent's highest qualification	Proficiency in English	Address—postcode	Address—postcode		Parent's highest qualification
Parent occupation	Year of arrival				Parent's highest school level	Year of arrival	Mesh Block	Mesh Block		Parent's highest school level
Year of arrival					Parent's highest non-school qualification		ASGS—SA1	ASGS—SA1		Parent's highest non-school qualification
					Parent occupation		ASGS—SA2	ASGS—SA2		Parent occupation
					Proficiency in English		Proficiency in English	Proficiency in English		Year of arrival
					Year of arrival		Year of arrival	Year of arrival		

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B7: Socio-demographic data items inconsistent with standards

AEDC (1)	Apprentices (0)	Census (1)	Higher Education (1)	NAP^(a) (2)	NECECC (0)	NSSC^(b) (1)	PISA (4)	TIMSS & PIRLS (5)	VET Provider (0)	VET in Schools (0)
Proficiency in English		Parents' country of birth	Parent's highest qualification	Remoteness		Parent occupation	Date of birth	Country of birth		
				Parent occupation			Remoteness	Remoteness		
							Parents' country of birth	Language spoken at home		
							SES for area	Parents' country of birth		
								Parent occupation		

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B8: Socio-demographic data items not applicable by collection

AEDC (4)	Apprentices (9)	Census (3)	Higher Education (9)	NAP^(a) (5)	NECECC (4)	NSSC^(b) (5)	PISA (7)	TIMSS & PIRLS (7)	VET Provider (10)	VET in Schools (5)
ASGC—Collection District	ASGC—Collection District	ASGC—Collection District	ASGC—Collection District	ASGC—Collection District	ASGC—SLA	ASGC—Collection District	ASGC—Collection District	ASGC—Collection District	Mesh Block	Mesh Block
Income	ASGC—SLA	ASGC—SLA	ASGC—SLA	ASGC—SLA	Income	ASGC—SLA	ASGC—SLA	ASGC—SLA	ASGC—Collection District	ASGC—Collection District
Occupation	Income	SES for person/family	Income	Income	Occupation	Income	Income	Income	ASGC—SLA	ASGC—SLA
SES for person/family	Parent's country of birth		Occupation	Occupation	SES for person/family	Occupation	Occupation	Occupation	Income	Income
	Parent's highest qualification		Parent's country of birth	SES for person/family		SES for person/family	Parent's highest school level	Parent's highest school level	Parent's country of birth	SES for person/family
	Parent's highest school level		Parent's highest school level				Parent's highest non-school qualification	Parent's highest non-school qualification	Parent's highest qualification	
	Parent's highest non-school qualification		Parent's highest non-school qualification				SES for person/family	SES for person/family	Parent's highest school level	
	Parent occupation		Parent occupation						Parent's highest non-school qualification	
	SES for person/family		SES for person/family						Parent occupation	
								SES for person/family		

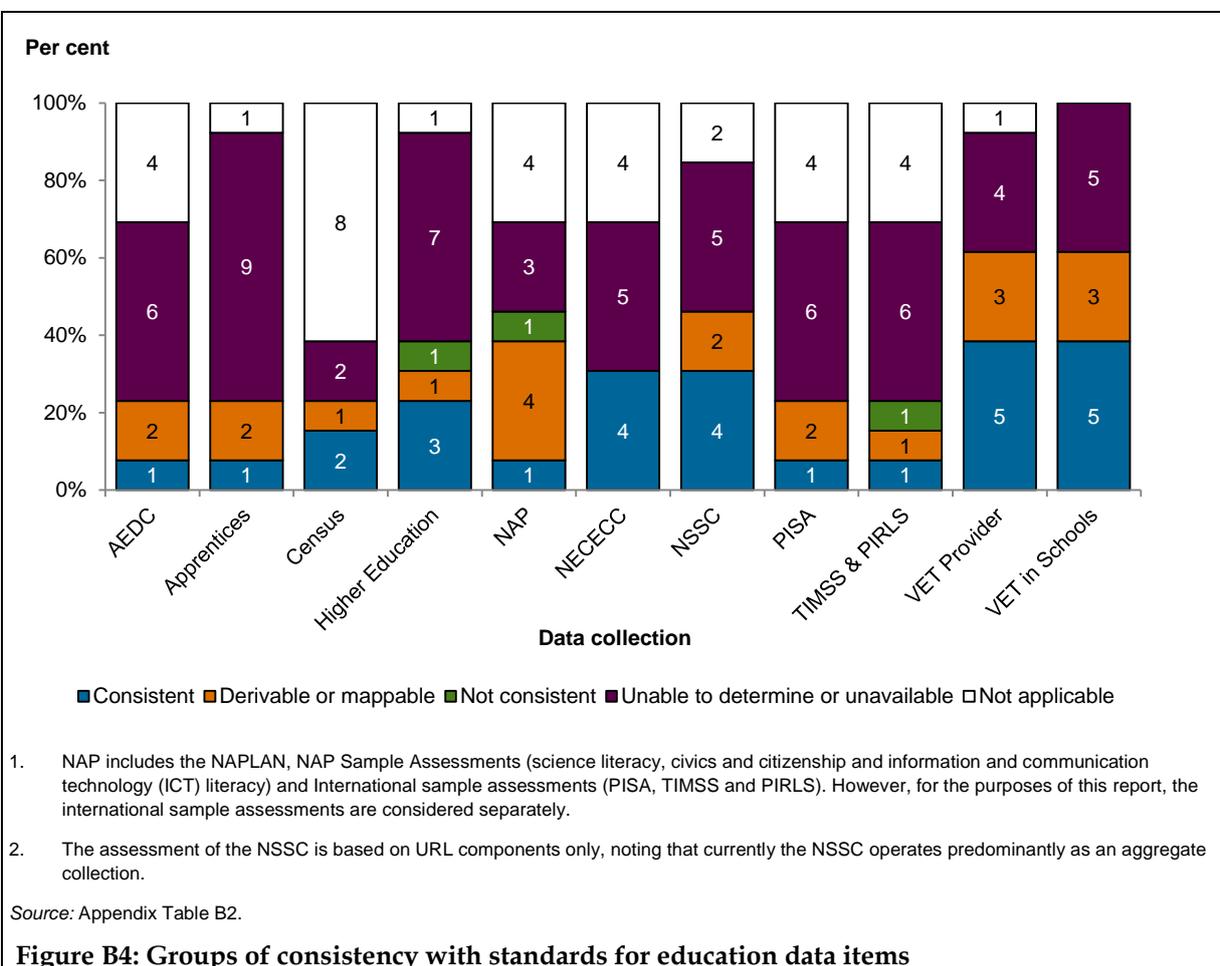
(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Education data items

Thirteen core education data items were assessed (a total of 143 data items across the 11 data collections). Twenty-eight (20%) of these data items were assessed as fully consistent with national standards. Twenty-two data items (15%) were mappable and none were derivable from other items (Figure B4 and Table B9).



The data items with the greatest level of consistency were:

- 10 collections: Educational institution ID/name (note that despite being fully consistent with the standard, the identifiers are not necessarily consistent between collections)
- 8 collections: Educational institution address – state
- 7 collections: Educational institution address – suburb
- 6 collections: Highest school level completed
- 5 collections: Education institution address – address line, postcode
- 4 collections: Highest non-school qualification, Field of education.

Thirty-one data items (22%) were assessed as potentially relevant to the data collection but not collected. Of these, 20 did not have a national standard. Concurrent and previous educational institution data items were those most commonly deemed to be relevant but not collected (Table B10).

Close to one-quarter (23%) of data items were graded as not applicable; for example, highest qualification and field of education in early childhood or school-based collections (Table B12).

There were a number of education data items (21%) that are collected within the identified data collections but for which no national standard currently exists:

- Educational institution concurrent
- Educational institution previous
- Educational institution type
- Attendance or enrolment status
- Educational achievement (standard not considered appropriate for national data standard due to variation between collections).

Only a small proportion of data items were identified as not consistent with the national standards (2% or a total of 3 data items) or with not enough information available to determine (4% or 6 data items). The inconsistent data items were related to highest level of school completed for 2 data collections and highest qualification for 1 collection (Table B11). All of the data items for which consistency with the standards could not be determined were related to the address of the educational institution.

At the data collection level, the National VET Provider Collection and National VET Schools Collection had the highest proportion of data items that were consistent, derivable or mappable (54%), followed by the National Schools Statistics Collection (38%). The Census of Population and Housing collected the least number of applicable data items.

Table B9: Education data items that are derivable or mappable by collection

AEDC (2)	Apprentices (2)	Census (1)	Higher Education (1)	NAP^(a) (4)	NECECC (0)	NSSC^(b) (2)	PISA (2)	TIMSS & PIRLS (1)	VET Provider (3)	VET in Schools (3)
Educational institution—suburb	Highest school level completed	Highest non-school qualification	Educational institution—postcode	Educational institution—address line		Educational institution—state	Educational institution—state	Educational institution—state	Educational institution—state	Educational institution—state
Educational institution—state	Highest qualification			Educational institution—suburb		Highest school level completed	Highest school level completed		Highest school level completed	Highest school level completed
				Educational institution—state					Highest non-school qualification	Highest non-school qualification
				Educational institution—postcode						

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B10: Education data items that are relevant but do not exist in the collection

AEDC (2)	Apprentices (8)	Census (0)	Higher Education (5)	NAP^(a) (2)	NECECC (2)	NSSC^(b) (3)	PISA (2)	TIMSS & PIRLS (2)	VET Provider (2)	VET in Schools (3)
Educational institution—address line	Educational institution—address line		Educational institution—address line	Educational institution—concurrent	Educational institution—postcode	Educational institution—concurrent	Educational institution—concurrent	Educational institution—concurrent	Educational institution—previous	Educational institution—concurrent
Educational institution—postcode	Educational institution—suburb		Educational institution—state	Educational institution—previous	Educational institution—concurrent	Educational institution—previous	Educational institution—previous	Educational institution—previous	Attendance or enrolment status	Educational institution—previous
	Educational institution—state		Educational institution—concurrent			Educational achievement (results)				Attendance or enrolment status
	Educational institution—postcode		Educational institution—previous							
	Educational institution—concurrent		Highest school level							
	Educational institution—previous									
	Field of education									
	Educational achievement (results)									

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Note: National data standards do not exist for items that are shaded.

Source: Appendix Table B3.

Table B11: Education data items inconsistent with standards, by collection

AEDC (0)	Apprentices (0)	Census (0)	Higher Education (1)	NAP^(a) (1)	NECECC (0)	NSSC^(b) (0)	PISA (0)	TIMSS & PIRLS (1)	VET Provider (0)	VET in Schools (0)
			Highest non-school qualification	Highest school level				Highest school level		

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B12: Education data items not applicable, by collection

AEDC (4)	Apprentices (1)	Census (8)	Higher Education (1)	NAP^(a) (4)	NECECC (4)	NSSC^(b) (2)	PISA (4)	TIMSS & PIRLS (4)	VET Provider (1)	VET in Schools (0)
Field of education	Educational institution type	Educational institution ID/name	Educational institution type	Attendance or enrolment status	Field of education	Field of education	Attendance or enrolment status	Attendance or enrolment status	Educational institution type	
Educational achievement (results)		Educational institution—address line		Field of education	Educational achievement (results)	Highest non-school qualification	Field of education	Field of education		
Highest school level		Educational institution—suburb		Educational achievement (results)	Highest school level		Educational achievement (results)	Educational achievement (results)		
Highest non-school qualification		Educational institution—state		Highest non-school qualification	Highest non-school qualification		Highest non-school qualification	Highest non-school qualification		
		Educational institution—postcode								
		Educational institution—concurrent								
		Educational institution—previous								
		Educational achievement (results)								

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

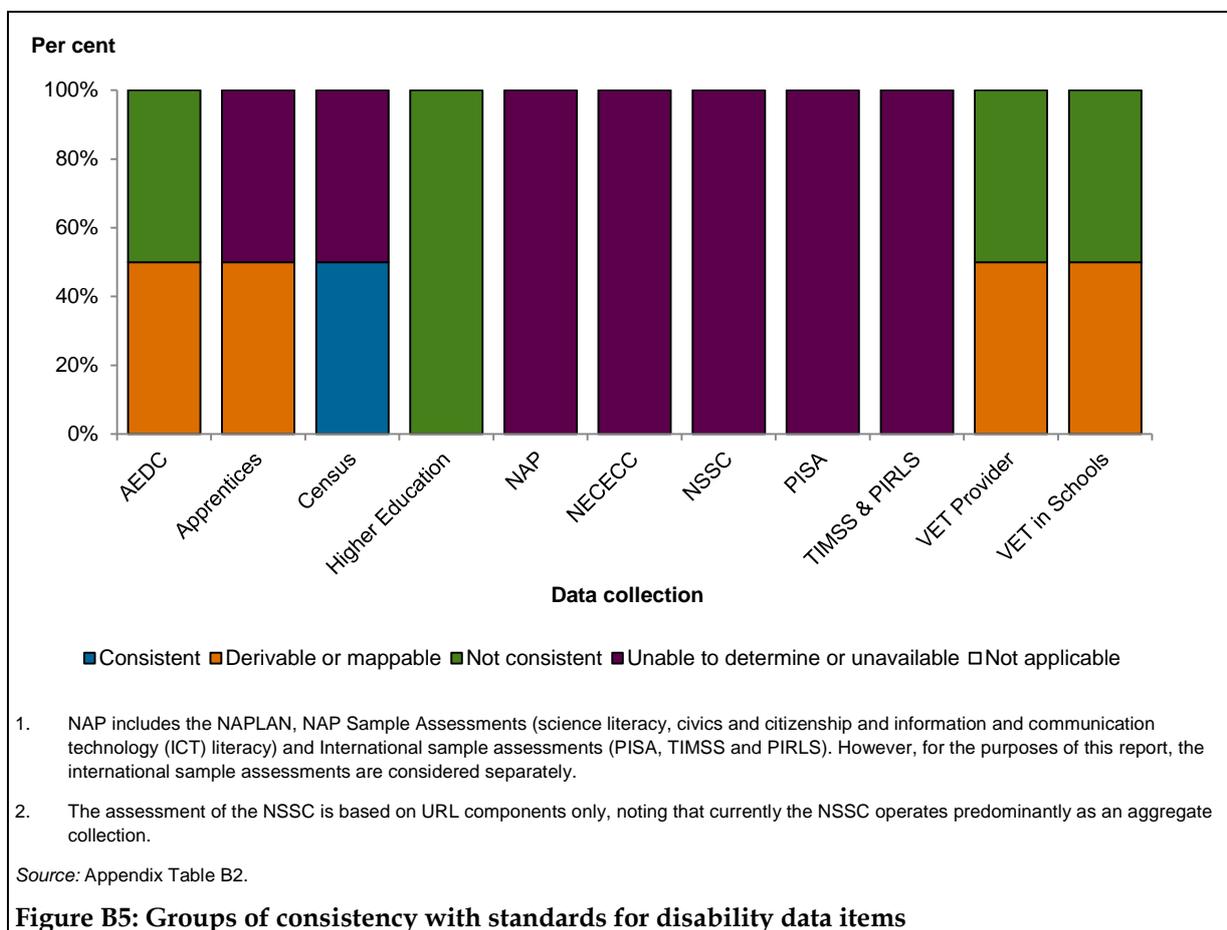
Disability data items

Three national data standards for disability data items were identified that were considered relevant to the education and training data collections. Two of the standards act as a flag to identify whether or not the person has disability (disability status and activity limitation indicator) and the other provides information on the disability type. Two types of disability data items were identified in the collections: disability flag and disability type.

There are a number of considerations related to the collection of disability information for children and students in education and training settings. These are discussed further in Appendix E under ‘Disability status’.

Six of the data collections collected the disability flag data item, and 5 of these collected the item in such a way that was consistent with or mappable to the national standards, while 1 was inconsistent (Figure B5 and Table B13).

Four data collections collected the disability type data item (all of which also collected the disability flag data item), but none of these items were collected in a manner consistent with or mappable to the national standard (Figure B5 and Table B15).



The Higher Education Statistics Collection combines the 2 disability data items and codes each disability type dichotomously. The AEDC also employs dichotomous coding to identify disability type, allowing multiple disability types to be identified. In contrast, the code set for the data standard allows only the disability group that most clearly expresses the experience of disability by a person to be recorded. The data collections also include a more comprehensive list of disabilities compared to the standard.

All of the collections that did not collect 1 or both disability data items were graded as relevant but not collected (Table B14).

Table B13: Disability data items that are derivable or mappable, by collection

AEDC (1)	Apprentices (1)	Census (0)	Higher Education (0)	NAP ^(a) (0)	NECECC (0)	NSSC ^(b) (0)	PISA (0)	TIMSS & PIRLS (0)	VET Provider (1)	VET in Schools (1)
Special needs/ Disability flag	Special needs/ Disability flag								Special needs/ Disability flag	Special needs/ Disability flag

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B14: Disability data items that are relevant but do not exist in the collection

AEDC (0)	Apprentices (1)	Census (1)	Higher Education (0)	NAP ^(a) (2)	NECECC (2)	NSSC ^(b) (2)	PISA (2)	TIMSS & PIRLS (2)	VET Provider (0)	VET in Schools (0)
	Special needs/ Disability type	Special needs/ Disability type		Special needs/ Disability flag						
				Special needs/ Disability type						

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B15: Disability data items inconsistent with standards, by collection

AEDC (1)	Apprentices (0)	Census (0)	Higher Education (2)	NAP^(a) (0)	NECECC (0)	NSSC^(b) (0)	PISA (0)	TIMSS & PIRLS (0)	VET Provider (1)	VET in Schools (1)
Special needs/ Disability type			Special needs/ Disability flag						Special needs/ Disability type	Special needs/ Disability type
			Special needs/ Disability type							

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Family composition data items

Only a single data item exists for family composition, and it is collected in 2 collections only: the Census of Population and Housing and PISA. A national data standard exists for this data item and both of the data collections are consistent with (Census) or mappable to the standard (PISA) (Figure B6 and Table B16).

The data item was graded as relevant but not collected for the other data collections (Table B17).

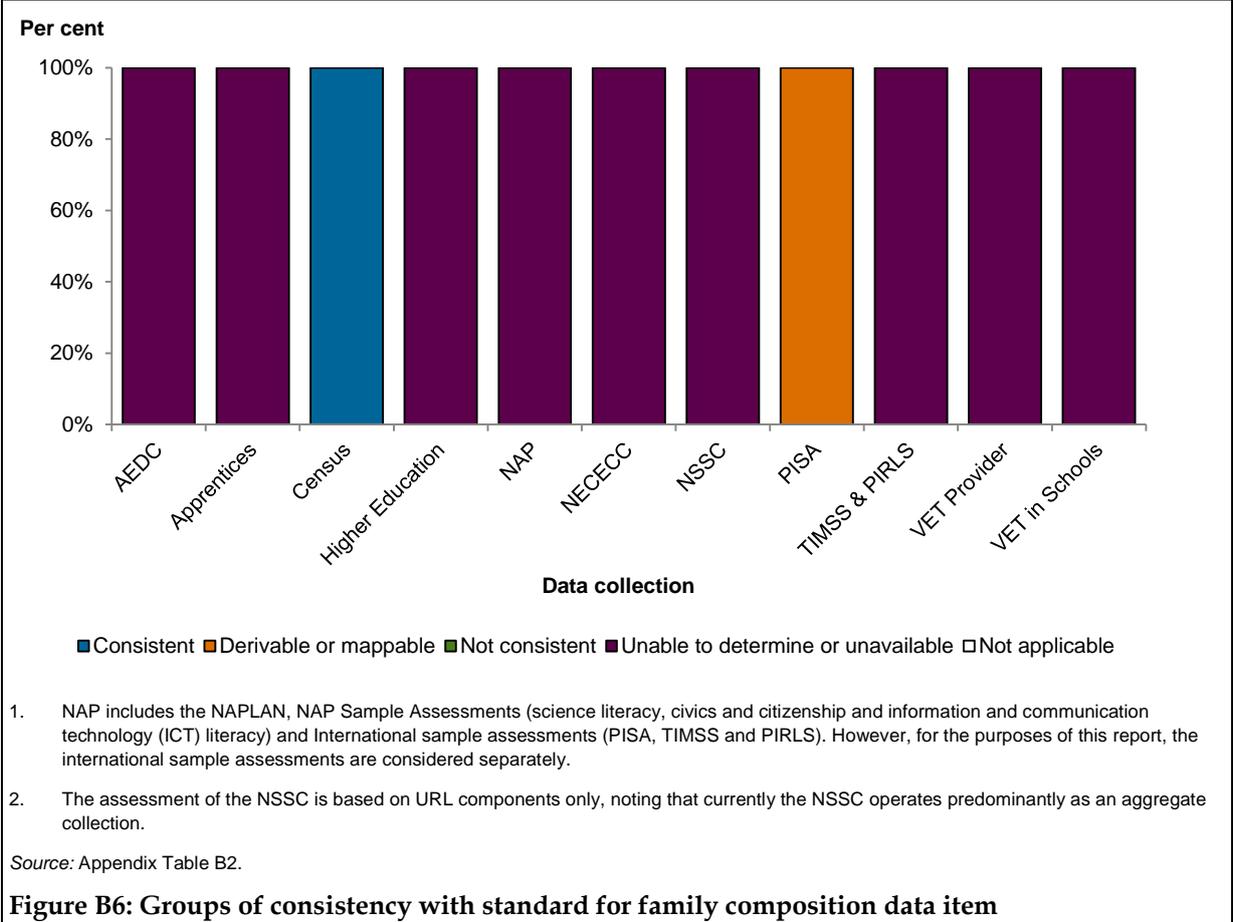


Table B16: Family composition data items that are derivable or mappable

AEDC (0)	Apprentices (0)	Census (0)	Higher Education (0)	NAP ^(a) (0)	NECECC (0)	NSSC ^(b) (0)	PISA (1)	TIMSS & PIRLS (0)	VET Provider (0)	VET in Schools (0)
							Family composition			

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B17: Family composition data items that are relevant but do not exist in the collection

AEDC (1)	Apprentices (1)	Census (0)	Higher Education (1)	NAP ^(a) (1)	NECECC (1)	NSSC ^(b) (1)	PISA (0)	TIMSS & PIRLS (1)	VET Provider (1)	VET in Schools (1)
Family composition	Family composition		Family composition	Family composition	Family composition	Family composition		Family composition	Family composition	Family composition

(a) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

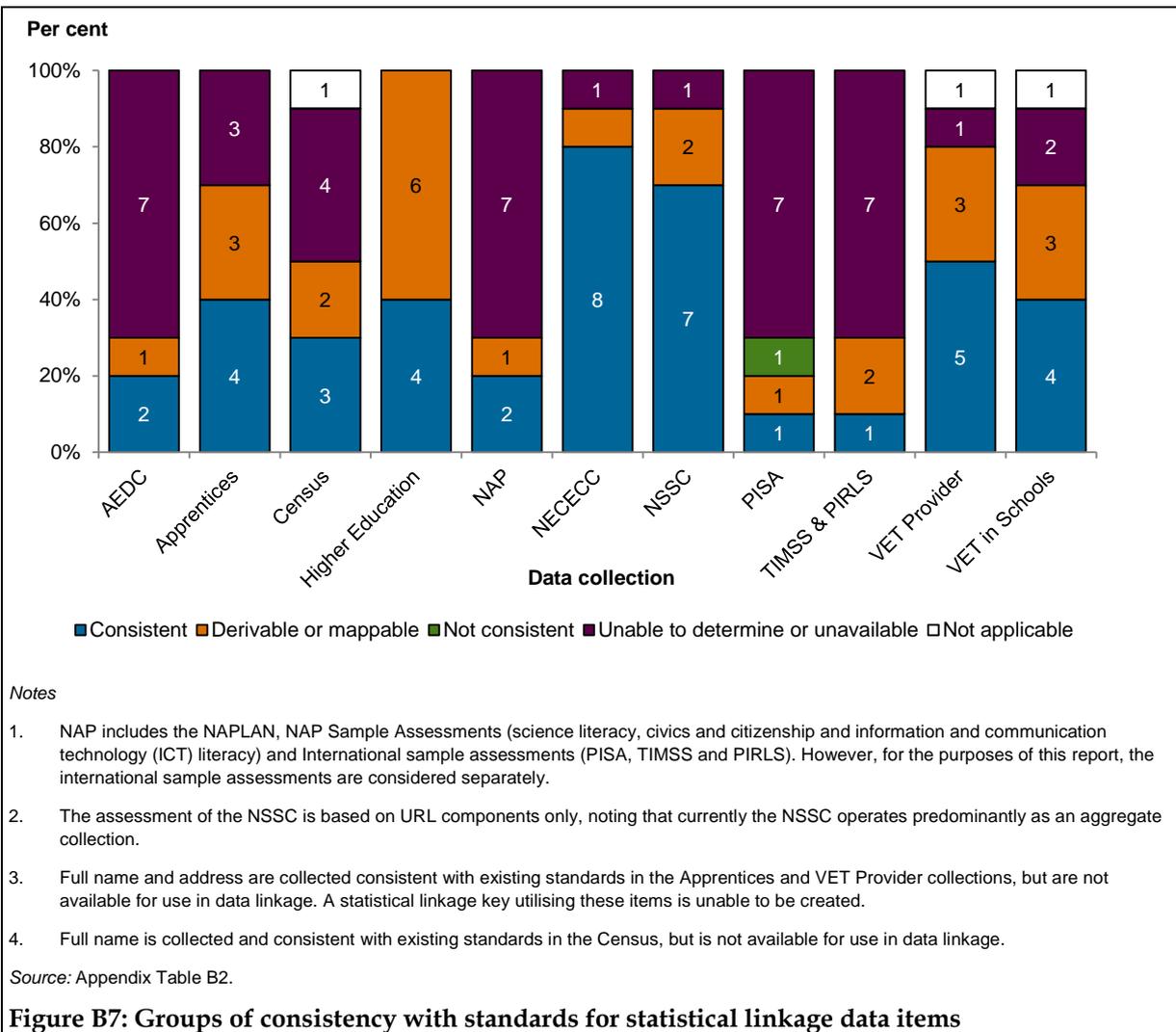
(b) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Statistical linkage data items

Ten data items were identified as statistical linkage data items across the data collections (110 data items in total across the 11 collections). Only the linkage key (statistical and non-statistical) and name (given and family name) data items are specific to the statistical linkage category, while the other items cut across the socio-demographic data item category (date of birth, sex, geography address – postcode, geography address – suburb).

Thirty-seven per cent (41) of the data items in this category were fully consistent with national standards, while a further 23% (25) were derivable or mappable (Figure B7 and Table B18).



As discussed in relation to socio-demographic data items, sex is consistent or mappable across all data collections, and date of birth appears to be consistent or mappable for all collections but the PISA, where information is available.

Among the data collections that collected the non-SLK data item, there was a high level of consistency with the national standard; however, the consistency of this data item between collections was non-existent. There is a project underway to implement a Unique Student Identifier into the National VET Provider Collection and National VET in Schools Collection, which may also be incorporated into the National Apprentices and Trainees Collection. This

would create consistency in this data item between these collections. This project is discussed further in Chapter 4 under 'Recent and future developments in data standards and linkage'.

SLK, given name and family name are collected and appear to be consistent, derivable or mappable to the national data standards in the Higher Education Statistics Collection, NECECC and NSSC. Despite all collections other than the Census collecting name items, this information may not be available for analysis or linkage (refer to information below regarding the inability to determine consistency from these collections).

Almost one-fifth (18%) of data items were graded as relevant but not collected (the PISA, TIMSS and PIRLS collections had the highest number of items in this category) (Table B19) and 18% lacked sufficient information to determine.

All of the data items in the statistical linkage category for which consistency was unable to be determined were from the AEDC, NAP, PISA, TIMSS and PIRLS collections. Consistency was unable to be determined based on the information in the publicly available metadata for these collections, and these data items have been graded accordingly in the mapping. It is expected that these data items are most likely consistent, derivable or mappable to the data standards. However, they are not held with the data collection and may not be available for analysis or data linkage purposes (dependent on the data collection).

Only 1 data item was assessed as not consistent with the national standard (date of birth for the PISA, Table B20) and 2 as not applicable (non-SLK for the Census of Population and Housing and Mesh Block for the National VET Provider Collection and National VET in Schools Collection) (Table B21).

The data collections with the highest levels of consistency with national standards (fully consistent, derivable or mappable) for the statistical linkage category are the Higher Education Statistics Collection (100%), followed by the the NECECC and NSSC (both 90%).

Table B18: Statistical linkage data items^(a) that are derivable or mappable, by collection

AEDC (1)	Apprentices (3)	Census (2)	Higher Education (6)	NAP ^(b) (1)	NECECC (1)	NSSC ^(c) (2)	PISA (1)	TIMSS & PIRLS (2)	VET Provider (3)	VET in Schools (3)
Sex	Family name	ASGS—SA1	Date of birth	Sex	Mesh Block	Address—postcode	Sex	Date of birth	Family name	Family name
	Given name	Sex	Identifier code—SLK			ASGS—SA1		Sex	Given name	Given name
	Sex		Address—suburb						Sex	Sex
			Mesh Block							
			ASGS—SA1							
			Sex							

(a) Includes some items that also appear in other categories.

(b) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(c) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B19: Statistical linkage data items^(a) that are relevant but do not exist in the collection

AEDC (0)	Apprentices (3)	Census (3)	Higher Education (0)	NAP ^(b) (0)	NECECC (1)	NSSC ^(c) (1)	PISA (5)	TIMSS & PIRLS (4)	VET Provider (1)	VET in Schools (2)
	Identifier code—SLK	Identifier code—SLK			Identifier code—non-SLK	Address—suburb	Identifier code—SLK	Address—suburb	Identifier code—SLK	Identifier code—SLK
	Mesh Block	Family name					Address—suburb	Address—postcode		ASGS—SA1
	ASGS—SA1	Given name					Address—postcode	Mesh Block		
							Mesh Block	ASGS—SA1		
							ASGS—SA1			

(a) Includes some items that also appear in other categories.

(b) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(c) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B20: Statistical linkage data items^(a) inconsistent with standards, by collection

AEDC (0)	Apprentices (0)	Census (0)	Higher Education (0)	NAP ^(b) (0)	NECECC (0)	NSSC ^(c) (0)	PISA (1)	TIMSS & PIRLS (0)	VET Provider (0)	VET in Schools (0)
							Date of birth			

(a) Includes some items that also appear in other categories.

(b) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(c) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Table B21: Statistical linkage data items^(a) not applicable, by collection

AEDC (0)	Apprentices (0)	Census (1)	Higher Education (0)	NAP ^(b) (0)	NECECC (0)	NSSC ^(c) (0)	PISA (0)	TIMSS & PIRLS (0)	VET Provider (1)	VET in Schools (1)
		Identifier code—non-SLK							Mesh Block	Mesh Block

(a) Includes some items that also appear in other categories.

(b) NAP includes the NAPLAN, NAP Sample Assessments (science literacy, civics and citizenship and information and communication technology (ICT) literacy) and International sample assessments (PISA, TIMSS and PIRLS). However, for the purposes of this report, the international sample assessments are considered separately.

(c) The assessment of the NSSC is based on URL components only, noting that currently the NSSC operates predominantly as an aggregate collection.

Source: Appendix Table B3.

Appendix C: Description of data collections and data items mapped

Australian Early Development Census

The Australian Early Development Census (AEDC) (formerly the Australian Early Development Index) is a measure of how young children are developing in Australian communities. The AEDC is a population measure of children's development as they enter their first year of formal school.

Reporting frequency

The AEDC is collected nationally every 3 years with the first collection in 2009 and the most recent in 2012.

Scope

The AEDC is a national collection across education sectors. The 2012 AEDC collected data on almost 290,000 children in their first year of formal full-time schooling, with almost 7,500 government and non-government schools participating.

Collection method

Information for the AEDC is collected through a teacher-completed checklist that measures 5 areas of early childhood development: physical health and wellbeing, social competence, emotional maturity, language and cognitive skills, and communication skills and general knowledge. The Australian Government Department of Education and Training and state and territory governments work in partnership with the Centre for Community Child Health, Royal Children's Hospital Centre Melbourne, the Murdoch Children's Research Institute and the Telethon Institute for Child Health Research in Perth to deliver the AEDC.

Further information

About the AEDC:

<<https://www.aedc.gov.au/about-the-aedc>>

Metadata

Data dictionary:

Currently under development. Available on request from AEDC Support.

Data user guide:

<<https://www.aedc.gov.au/researchers/resources-for-researchers/aedc-data-user-guide>>

Data quality statement

Refer to section 3 of the *Data user guide* (see above link).

Data items for mapping

Data item	Data description
Age	Age of child No further information available in the AEDC Data Dictionary
Country of birth	Indicates the child's country of birth Numeric 4 (SACC)
School ID	A unique 8-digit number identification code for each government school participating in the AEDC Numeric 8 NNNNNNNN Government school 99999999 Non-government school
School suburb	The suburb or town in which the school is located Text Length undisclosed
School state	To identify the location by state or territory of the school participating in the AEDC Text 4 ACT Australian Capital Territory NSW New South Wales NT Northern Territory QLD Queensland SA South Australia TAS Tasmania VIC Victoria WA Western Australia NULL Child can't be located
Dual placement	To identify if this child has dual placement Dual placement is when the child is co-enrolled at 2 or more educational institutions, or in 2 or more educational programs Numeric 1 0 No 1 Yes
Non-parental care/other educational program before entering school	To identify if the child was in regular, non-parental care or attended other educational programs in the year before entering school Numeric 2 0 No 1 Yes 88 Don't know
School type	To identify if the school participating in the AEDC is a government, independent or Catholic school Text 1 G Government I Independent C Catholic

Data item	Data description
Special needs status	To identify if the child has special needs status Numeric 1 0 Not special needs 1 Special needs
Proficiency in English	To indicate the child's proficiency in English Numeric 2 1 Poor/very poor 2 Average 3 Very good/good 88 Don't know 99 Not applicable
Behaviour problem Chronic illness Emotional problem Hearing impairment Learning disability Physical disability Neurodevelopmental disorder Speech impairment Visual impairment Other enduring problems	To identify if the child has a physical disability that affects their ability to do school work in a regular classroom Numeric 2 0 No 1 Yes 88 Don't know
Local Community ARIA category	The Accessibility and Remoteness Index of Australia (ARIA) category of the AEDC Local Community Text Major Cities of Australia Inner Regional Australia Outer Regional Australia Remote Australia Very Remote Australia
Local Community	AEDC Local Community in which the child resides Text Derived from Local Community ID Numeric 6 NNNNNN Assumed to be based on state suburb boundaries from 2006 (SSC-06) 999999 Child can't be located
State ID	Code for state in which child resides Numeric 1 1 New South Wales 2 Victoria 3 Queensland 4 South Australia 5 Western Australia 6 Tasmania 7 Northern Territory 8 Australian Capital Territory

Data item	Data description
	9 Offshore
SLA code	Code corresponding with SLA in which the child resides Numeric 9 NNNNNNNNN Valid NULL No data
Street address	Not a specified item but referred to in Data User Guide
Postcode	Not a specified item but referred to in Data User Guide
Student ID	An individual checklist is completed for each child participating in the AEDC. The Student ID uses a unique numerical code to distinguish each checklist without providing any identifying information. Numeric 6 NNNNNN A 6-digit number
ATSI type	To provide additional information if the child is of Aboriginal or Torres Strait Islander descent Numeric 1 1 Aboriginal but not Torres Strait Islander origin 2 Torres Strait Islander but not Aboriginal origin 3 Both Aboriginal and Torres Strait Islander origin 4 Neither Aboriginal nor Torres Strait Islander origin 9 Not stated or unknown
Language ID	Main language spoken at home (Only answer if Lang = Yes) Text 3 3-digit number AEDC Language reference file -1 iLanguageID selected NULL No Data (iLanguageID means that the child speaks a traditional Aboriginal or Torres Strait Islander language) The Data User Guide states that responses are coded using the Australian Standard Classification of Languages (ASCL) and that typically only the Broad Group of aggregated languages is available for research purposes.
Gender	The sex of the child Numeric 1 1 Male 2 Female
Local Community SEIFA score category	Quintile grouping variable for the LCABSSEIFAScore of the AEDC Local Community Numeric 4 1 SEIFA score <= 942.36 (most disadvantaged) 2 SEIFA score >= 942.37 and SEIFA score < 991.17 3 SEIFA score >= 991.18 and SEIFA score < 1029.49 4 SEIFA score >= 1029.50 and SEIFA score < 1065.41 5 SEIFA score > 1065.41 (least disadvantaged)

Source: Social Research Centre 2013.

Census of Population and Housing

The Census aims to accurately measure the number of and key characteristics of people in Australia on Census night, and the dwellings in which they live. This information provides a reliable basis for estimation of the population of each of the states, territories and local government areas, primarily for electoral purposes and the distribution of government funds.

Reporting frequency

Since 1911, the ABS has administered the Census and has conducted it every 5 years since 1961. The most recent Census was held on 9 August 2011.

Scope

The Census includes all persons and dwellings in Australia including visitors to Australia and the external territories of Christmas Island and Cocos (Keeling) islands on Census night. Australian residents out of the country on Census night are excluded along with diplomats, their families and diplomatic dwellings, and visitors from overseas who are not required to undergo migration formalities, such as foreign crews on ships.

Collection method

The Census is collected and published against geographic areas. Census forms are delivered to each dwelling in Australia. Since 2006, the Census has also been able to be completed and collected electronically.

Further information

How Australia takes a Census:

<<http://www.abs.gov.au/ausstats/abs@.nsf/lookup/2903.0Main%20Features22011>>

Metadata

2011 Census Dictionary:

<<http://www.abs.gov.au/ausstats/abs@.nsf/mf/2901.0>>

2011 reference and information:

<<http://www.abs.gov.au/websitedbs/censushome.nsf/home/2011information?opendocument&navpos=310>>

Data quality statement

Data quality information, including quality declaration, data quality statements and investigations:

<<http://www.abs.gov.au/websitedbs/censushome.nsf/home/dataquality?opendocument&navpos=400>>

Data items for mapping

Data item	Data description
Age	Person's age at last birthday. Age is calculated from date of birth when provided, otherwise stated age is used. Only age in years data are output.

Data item	Data description
	Numeric 3 Age range 0–115 years Available in 3 categories: Single year age group 116 categories Five year age group 21 categories Ten year age group 11 categories Imputation flag for age: Indicates if a person's age was imputed. 1 Age imputed 2 Age not imputed
Country of birth of person	Indicates in which country a person was born and is coded using the Standard Australian Classification of Countries (SACC) second edition, revision 1 Alphanumeric 4
Type of educational institution attending	Records the type of educational institution being attended. Alphanumeric 2 10 Preschool 21 Infants/Primary—Government 22 Infants/Primary—Catholic 23 Infants/Primary—Other Non-Government 31 Secondary—Government 32 Secondary—Catholic 33 Secondary—Other Non-Government 40 Technical or Further Educational Institution (including TAFE Colleges) 50 University or other Tertiary Institution 60 Other && Not stated @@ Not applicable VV Overseas visitor
Full-time/part-time student status	Describes the full-time or part-time status of students Alphanumeric 6 1 Not attending 2 Full-time student 3 Part-time student 4 Institution (TYPP) stated, full-time/part-time status (STUP) not stated & Both not stated—both institution (TYPP) and full-time/part-time status (STUP) not stated V Overseas visitor
Non-school qualification: Field of study	Describes the field of study of a person's highest completed non-school qualification Coded using the Australian Standard Classification of Education (ASCED), Field of Education Classification Alphanumeric 6 01–129999 Valid 000110 Field of study inadequately described &&&&& Field of study not stated @@@@ Not applicable VVVVVV Overseas visitor

Data item	Data description
Highest year of school completed	<p>This variable contains the highest level of primary or secondary schooling completed</p> <p>Coded using the Australian Standard Classification of Education (ASCED), 2001 (cat. no. 1272.0)</p> <p>Alphanumeric</p> <p>1</p> <p>1 Year 12 or equivalent</p> <p>2 Year 11 or equivalent</p> <p>3 Year 10 or equivalent</p> <p>4 Year 9 or equivalent</p> <p>5 Year 8 or below</p> <p>6 Did not go to school</p> <p>& Not stated</p> <p>@ Not applicable</p> <p>V Overseas visitor</p>
Non-school qualification: Level of Education	<p>Describes the level of a person's highest completed non-school qualification</p> <p>Coded using the Australian Standard Classification of Education (ASCED), Level of Education Classification</p> <p>Alphanumeric</p> <p>3</p> <p>1 Postgraduate Degree Level</p> <p>10 Postgraduate Degree Level, nfd</p> <p>11 Doctoral Degree Level</p> <p>12 Master Degree Level</p> <p>2 Graduate Diploma and Graduate Certificate Level</p> <p>20 Graduate Diploma and Graduate Certificate Level, nfd</p> <p>21 Graduate Diploma Level</p> <p>22 Graduate Certificate Level</p> <p>3 Bachelor Degree Level</p> <p>31 Bachelor Degree Level</p> <p>4 Advanced Diploma and Diploma Level</p> <p>40 Advanced Diploma and Diploma Level, nfd</p> <p>41 Advanced Diploma and Associate Degree Level</p> <p>42 Diploma Level</p> <p>5 Certificate Level</p> <p>50 Certificate Level, nfd</p> <p>500 Certificate Level, nfd</p> <p>51 Certificate III & IV Level</p> <p>510 Certificate III & IV Level, nfd</p> <p>511 Certificate IV</p> <p>514 Certificate III</p> <p>52 Certificate I & II Level</p> <p>520 Certificate I & II Level, nfd</p> <p>521 Certificate II</p> <p>524 Certificate I</p> <p>Supplementary codes</p> <p>011 Level of education inadequately described</p> <p>&&& Level of education not stated</p> <p>@@@ Not applicable</p> <p>VVV Overseas visitor</p>
Address line information	Information destroyed post statistical processing if provided

Data item	Data description
Suburb/locality	Alphanumeric 25
Postal area	Postal Areas are ABS approximations of Australia Post postcodes, created by allocating whole Statistical Areas Level 1 (SA1s) on a 'best fit' basis to postcodes This means that there are more Australia Post postcodes than Census Postal Areas. Alphanumeric 4
Mesh Block	Mesh Blocks are the smallest geographical area. Only limited Census data, i.e. total population and dwelling counts, is released at the Mesh Block level. Numeric 11 The 11-digit Mesh Block code comprises: S/T identifier (1 digit), Mesh Block identifier (10 digits). 0000000000–9999999999
Remoteness	Within the Australian Statistical Geography Standard (ASGS), the Remoteness structure comprises 6 categories, each of which identifies a non-contiguous region in Australia, being a grouping of Statistical Areas Level 1 (SA1s) sharing a particular degree of remoteness. The degrees of remoteness range from 'Major Cities' (highly accessible) to 'Very Remote'. The degree of remoteness of each SA1 was determined using the Accessibility/Remoteness Index of Australia (ARIA).
State/territory	The State/Territory is the largest spatial unit in the Australian Statistical Geography Standard (ASGS). Numeric 1 1 NSW 2 Vic 3 Qld 4 SA 5 WA 6 Tas 7 NT 8 ACT 9 Other territories Jervis Bay Territory and the Territories of Christmas Island and Cocos (Keeling) Islands are grouped as one spatial unit at the State/Territory level in the category of Other Territories.
Total personal income (weekly)	Indicates the total income that the person usually receives each week. Equivalent annual amounts appear in brackets. Alphanumeric 2 01 Negative income 02 Nil income 03 \$1–\$199 (\$1–\$10,399) 04 \$200–\$299 (\$10,400–\$15,599) 05 \$300–\$399 (\$15,600–\$20,799) 06 \$400–\$599 (\$20,800–\$31,199) 07 \$600–\$799 (\$31,200–\$41,599) 08 \$800–\$999 (\$41,600–\$51,999) 09 \$1,000–\$1,249 (\$52,000–\$64,999) 10 \$1,250–\$1,499 (\$65,000–\$77,999) 11 \$1,500–\$1,999 (\$78,000–\$103,999)

Data item	Data description
	12 \$2,000 or more (\$104,000 or more) && Not stated @@ Not applicable VV Overseas visitor
Indigenous status	Provides responses of persons who identified themselves as being of Australian Aboriginal and/or Torres Strait Islander origin Alphanumeric 1 1 Non-Indigenous 2 Aboriginal 3 Torres Strait Islander 4 Both Aboriginal and Torres Strait Islander & Not stated V Overseas visitor
Language spoken at home	Records responses to the Census question 'Does the person speak a language other than English at home?', and are coded using the Australian Standard Classification of Languages, second edition, revision 1 Alphanumeric 4 1000–9999 Valid 4-digit ASCL code 0000 Inadequately described 0002 Non-verbal, so described 0003 Swiss, so described 0004 Cypriot, so described 0005 Creole, nfd 0006 French Creole, nfd 0007 Spanish Creole, nfd 0008 Portuguese Creole, nfd 0009 Pidgin, nfd &&&& Not stated VVVV Overseas visitor
Occupation	Coded using the Australian and New Zealand Standard Classification of Occupations (ANZSCO), First Edition, Revision 1. The Occupation code assigned to a response is based on the occupation title and tasks of the main job held during the week prior to Census night. Alphanumeric 4 (6-digit level can be supplied on request) 1000–8999 Valid ANZSCO code 0998 Inadequately described &&&& Not stated @@@@ Not applicable VVVV Overseas visitor
Birthplace of male parent	Indicates whether a person's Father was born in Australia or overseas Alphanumeric 1 1 Born in Australia 2 Born overseas & Not stated V Overseas visitor
Birthplace of female parent	Indicates whether a person's Mother was born in Australia or overseas Alphanumeric

Data item	Data description
	1 1 Born in Australia 2 Born overseas & Not stated V Overseas visitor
Proficiency in spoken English	Classifies each person's self-assessed proficiency in spoken English Alphanumeric 1 Speaks English only 1 Speaks English only Speaks other language and speaks English 2 Very well 3 Well 4 Not well 5 Not at all Not stated 6 Not stated—both language (LANP) and proficiency (ENGP) not stated & Not stated—language (LANP) stated, proficiency (ENGP) not stated V Overseas visitor
Sex	Records each person's sex. If sex is not stated it is imputed Numeric 1 1 Female 2 Male
Year of arrival in Australia	For people born overseas, states the year they first arrived in Australia, with the intention of staying for at least 1 year The variable Country of Birth of Person (BPLP) is the basis for determining if a person was born in Australia or overseas. BPLP uses the Standard Australian Classification of Countries (SACC) Second Edition, Revision 1 to classify country of birth. In SACC, Australia is category 1101. Alphanumeric 4 1895–2011 1895 to 2011 singly &&&& Not stated @@@@ Not applicable VVVV Overseas visitor
Core activity need for assistance	People with a profound or severe disability are defined as those people needing help or assistance in one or more of the 3 core activity areas of self-care, mobility and communication, because of a disability, long-term health condition (lasting 6 months or more) or old age Alphanumeric 1 1 Has need for assistance with core activities 2 Does not have need for assistance with core activities & Not stated V Overseas visitor
Family composition	Family Composition classifies families into different types. When classifying families into different types, information about temporarily absent family members is used. Alphanumeric 4 1 COUPLE FAMILY WITH NO CHILDREN 12 Couple family with no children

Data item	Data description
	122 Couple family with no children
	1222 Couple family with no children
	2 COUPLE FAMILY WITH CHILDREN
	21 Couple family with children under 15
	211 Couple family with children under 15 and dependent students
	2111 Couple family with children under 15, dependent students and non-dependent children
	2112 Couple family with children under 15, dependent students and no non-dependent children
	212 Couple family with children under 15 and no dependent students
	2121 Couple family with children under 15, no dependent students and with non-dependent children
	2122 Couple family with children under 15, no dependent students and no non-dependent children
	22 Couple family with no children under 15
	221 Couple family with no children under 15 and with dependent students
	2211 Couple family with no children under 15, and with dependent students and non-dependent children
	2212 Couple family with no children under 15, and with dependent students and no non-dependent children
	222 Couple family with no children under 15 and no dependent students
	2221 Couple family with no children under 15, no dependent students and with non-dependent children
	3 ONE PARENT FAMILY
	31 One parent family with children under 15
	311 One parent family with children under 15 and dependent students
	3111 One parent family with children under 15, dependent students and non-dependent children
	3112 One parent family with children under 15, dependent students and no non-dependent children
	312 One parent family with children under 15 and no dependent students
	3121 One parent family with children under 15, no dependent students and with non-dependent children
	3122 One parent family with children under 15, no dependent students and no non-dependent children
	32 One parent family with no children under 15
	321 One parent family with no children under 15 and with dependent students
	3211 One parent family with no children under 15, with dependent students and non-dependent children
	3212 One parent family with no children under 15, with dependent students and no non-dependent children
	322 One parent family with no children under 15 and no dependent students
	3221 One parent family with no children under 15, no dependent students and with non-dependent children
	9 OTHER FAMILY
	92 Other family
	922 Other family
	9222 Other family
	@@@ Not applicable

Note: Surname and first name are collected in the Census but this information is destroyed post-statistical processing unless respondents have explicitly agreed to their name-identified information being retained. This information is not available for any purpose (including analysis or data linkage), within a 99 year closed access period.

Source: ABS 2011.

Higher Education Statistics Collection

The Higher Education Statistics Collection gathers statistics relating to students and staff from all Australian higher education institutions where funding is provided by the Australian Government. Data included in the Higher Education Statistics Collection relate to: courses conducted by higher education institutions; numbers and characteristics of students undertaking courses; student load; completion of units of study and courses; students' liabilities under the Higher Education Contribution Scheme (HECS); numbers and characteristics of staff in higher education institutions; and the educational profiles of higher education institutions.

Reporting frequency

Biannual (full year data in June, half year data in November).

Scope

The Higher Education Statistics Collection includes:

- courses and student information that relate to all units of study with census dates from 1 January to 31 December
- information about students who are in receipt of an OS-HELP or SA-HELP loan from 1 January to 31 December
- information about students who commence or continue a Commonwealth Scholarship in the reporting year.

Collection method

The higher education student submissions are made to the Department of Education and Training through the Higher Education Provider Client Assistance Tool. Commonwealth Scholarship submissions are made through Web Services and the HEIMS Administration system provides functionality to administer the allocation of Commonwealth Higher Education Student Support Numbers.

Metadata

Data elements:

<<http://heimshelp.education.gov.au/sites/heimshelp/dictionary/pages/data-element-dictionary>>

Scope and structure documents:

<<http://heimshelp.education.gov.au/sites/heimshelp/resources/pages/scope-structure-guidelines>>

Glossary:

<<http://heimshelp.education.gov.au/sites/heimshelp/resources/glossary/pages/glossary>>

Data quality statement

No data quality statement or quality declaration is available, but data providers are required to verify data reported twice a year as part of the annual reporting requirements issued under ministerial notice.

<<http://heimshelp.education.gov.au/sites/heimshelp/resources/toolkits/pages/verifying-data>>

Data items for mapping

Data item	Data description
Country of birth code	<p>A code representing the country of birth of a student. Codes are based on Standard Australian Classification of Countries</p> <p>Numeric 4</p> <p>1100 Student was born in Australia</p> <p>1200 to 9299 Overseas country code from countries classification for student who was born outside Australia</p> <p>9998 No information on country for student who was born outside Australia</p> <p>9999 No information on whether born in Australia or not</p>
Date of birth	<p>The day, month and year of birth of the student/applicant</p> <p>Numeric 8 YYYYMMDD</p> <p>Format: 1st 4 digits—Year sub-field (e.g. 1957) 2nd 2 digits—Month sub-field (e.g. 08) 3rd 2 digits—Day sub-field (e.g. 06)</p> <p>Where date of birth cannot be provided, report '19010101'</p>
Higher Education provider code	<p>A code which uniquely identifies the Higher Education Provider</p> <p>Numeric 4</p> <p>1000 to 9999 Higher Education Provider code issued by the Department of Education and Training</p> <p>0000 Providers not issued with a code</p>
Campus location	<p>Name of suburb/town/region where the campus is located</p> <p>Alphanumeric 27</p> <p>Any data Text provided by the university</p>
Postcode or overseas country code location of Higher Education/VET Provider campus/delivery location	<p>The Australian postcode or overseas country code location of the campus/delivery location of the Higher Education/VET Provider where the student is studying this unit</p> <p>Alphanumeric 5</p> <p>X1200 to X9299 Overseas country code from countries classification</p> <p>X9999 Overseas country, but not classifiable</p> <p>A0001 to A9998 Australian Postcode</p> <p>99999 No information on location of campus/delivery location</p>
Type of attendance code	<p>A code which identifies whether a student is classified as being a full-time or part-time</p> <p>Numeric 1</p> <p>0 For OUA and non-higher degree research student use only</p> <p>1 Full-time attendance for:</p> <ul style="list-style-type: none"> • higher degree research student • higher education course completions for all full-time students • VET student <p>2 Part-time attendance for:</p> <ul style="list-style-type: none"> • higher degree research student

Data item	Data description
	<ul style="list-style-type: none"> • higher education course completions for all part-time students • VET student
Geography— address Address of permanent home residence—Part 1	The first part of the address for the student's permanent home residence Text 38 (width) Any data provided by the student
Address of permanent home residence—Part 2	The second part of the address for the student's permanent home residence Text 38 (width) Any data provided by the student
Name of suburb/town/locality	Name of suburb/town/locality for the student Text 27 Any data provided by the student
Address of permanent home residence— postcode	The Australian postcode or overseas country code for the student's permanent home residence Alphanumeric 4 bbbb (4 spaces) Overseas country 0001 to 9999 Australian postcode
Residential address—state	A code which uniquely identifies the State/Territory or other area in which the student has their residential address Alphanumeric 3 NSW New South Wales VIC Victoria QLD Queensland WA Western Australia SA South Australia TAS Tasmania NT Northern Territory ACT Australian Capital Territory AAT Australian Antarctic Territories
Commonwealth Higher Education Student Support Number (CHESSN)	A code, allocated by the Department of Education and Training, which uniquely identifies the student within the Higher Education or VET Sector and remains constant from year to year Numeric 10 ZZZZZZZZZ No CHESSN required Any data Code assigned by HEIMS

Aboriginal and Torres Strait Islander Code	<p>A code which identifies whether or not the student/applicant identifies herself or himself as being of Aboriginal and/or Torres Strait Islander descent</p> <p>Numeric</p> <p>1</p> <p>2 Non-Indigenous—neither Aboriginal nor Torres Strait Islander origin</p> <p>3 Of Aboriginal origin but not Torres Strait Islander</p> <p>4 Of Torres Strait Islander origin but not Aboriginal</p> <p>5 Both Aboriginal and Torres Strait Islander origin</p> <p>9 No information</p>
Language spoken at home code	<p>A code indicating use of a language other than English at the student's/applicant's permanent home residence. The Australian Standard Classification of Languages is used.</p> <p>Numeric</p> <p>4</p> <p>0001 Student/Applicant speaks only English at permanent home residence</p> <p>1000 to 1199 Non-English language code from languages classification</p> <p>1300 to 9799 Non-English language code from languages classification</p> <p>9998 Non-English language spoken but no information on the language</p> <p>9999 No information on whether or not a non-English language is spoken at permanent home residence by a student/applicant</p>
Student surname	<p>The surname of the student/applicant</p> <p>Text</p> <p>30</p> <p>Any data provided by the student</p>
Student given name—first	<p>The first given name of the student/applicant</p> <p>Text</p> <p>15</p> <p>Any data provided by the student</p>
Highest level of educational attainment by parent/guardian 1	<p>A code which provides information about the highest educational attainment of the first parent or guardian as identified by the student</p> <p>Alphanumeric</p> <p>2</p> <p>01 Not a commencing student</p> <p>98 Not applicable</p> <p>99 No information provided by student</p> <p>A commencing student whose highest educational attainment of a male parent/guardian was:</p> <p>20 Postgraduate qualification (e.g. Postgraduate Diploma, Masters, PhD)</p> <p>21 Bachelor Degree</p> <p>22 Other post-school qualification (e.g. VET Certificate, Associate Degree or Diploma)</p> <p>23 Completed Year 12 schooling or equivalent</p> <p>24 Did not complete Year 12 schooling or equivalent</p> <p>25 Completed Year 10 schooling or equivalent</p> <p>26 Did not complete Year 10 schooling or equivalent</p> <p>49 Don't know</p>

	<p>A commencing student whose highest educational attainment of a female parent/guardian was:</p> <p>40 Postgraduate qualification (e.g. Postgraduate Diploma, Masters, PhD)</p> <p>41 Bachelor Degree</p> <p>42 Other post-school qualification (e.g. VET/TAFE Certificate, Associate Degree or Diploma)</p> <p>43 Completed Year 12 schooling or equivalent</p> <p>44 Did not complete Year 12 schooling or equivalent</p> <p>45 Completed Year 10 schooling or equivalent</p> <p>46 Did not complete Year 10 schooling or equivalent</p> <p>59 Don't know</p> <p><i>Note: Applies to commencing students only</i></p>
Highest level of educational attainment by parent/guardian 2	<p>A code which provides information about the highest educational attainment of a second parent or guardian as identified by the student</p> <p>01 Not a commencing student</p> <p>98 Not applicable</p> <p>99 No information provided by student</p> <p>A commencing student whose highest educational attainment of a male parent/guardian was:</p> <p>20 Postgraduate qualification (e.g. Postgraduate Diploma, Masters, PhD)</p> <p>21 Bachelor Degree</p> <p>22 Other post-school qualification (e.g. VET Certificate, Associate Degree or Diploma)</p> <p>23 Completed Year 12 schooling or equivalent</p> <p>24 Did not complete Year 12 schooling or equivalent</p> <p>25 Completed Year 10 schooling or equivalent</p> <p>26 Did not complete Year 10 schooling or equivalent</p> <p>49 Don't know</p> <p>A commencing student whose highest educational attainment of a female parent/guardian was:</p> <p>40 Postgraduate qualification (e.g. Postgraduate Diploma, Masters, PhD)</p> <p>41 Bachelor Degree</p> <p>42 Other post-school qualification (e.g. VET/TAFE Certificate, Associate Degree or Diploma)</p> <p>43 Completed Year 12 schooling or equivalent</p> <p>44 Did not complete Year 12 schooling or equivalent</p> <p>45 Completed Year 10 schooling or equivalent</p> <p>46 Did not complete Year 10 schooling or equivalent</p> <p>59 Don't know</p> <p><i>Note: Applies to commencing students only</i></p>
Gender code	<p>A code which identifies the sex of a student/applicant</p> <p>Alphanumeric</p> <p>1</p> <p>F Female</p> <p>M Male</p>

	<p>X Indeterminate/Intersex/Unspecified</p> <p>In response to the Australian Government Guidelines on the Recognition of Sex and Gender, the Department of Education and Training has updated the HEIMS systems to allow the reporting of code X (indeterminate/Intersex/Unspecified). This facilitates the progressive implementation of the guidelines by all providers by 1 July 2016.</p>
Year of arrival in Australia	<p>The year in which a student/applicant not born in Australia first arrived in Australia</p> <p>Alphanumeric 4</p> <p>0000 Student/Applicant never arrived in Australia</p> <p>0001 Student/Applicant was born in Australia</p> <p>Student/Applicant not born in Australia and year of arrival in Australia:</p> <p>1900 to 2099 Year of arrival in YYYY format</p> <p>A998 No information on year of arrival</p> <p>A999 No information on whether student/applicant was born in Australia or not</p>
Field of education code	<p>A code which identifies the field of education to which the course is classified</p> <p>Coded using the Australian Standard Classification of Education (ASCED), Field of Education Classification</p> <p>Numeric 6</p> <p>000000 Non-award course , BOTP, or OUA unit</p> <p>010000 to 129999 Field of education code</p>
Unit of study completion status	<p>A code which categorises a students unit of study completion status</p> <p>Numeric 1</p> <p>1 Withdrew without penalty</p> <p>2 Failed</p> <p>3 Successfully completed all the requirements</p> <p>4 Unit of study to be commenced later in the year or still in process of completing or completion status not yet determined</p> <p>5 Recognition of prior learning (VET only)</p>
Highest educational participation prior to commencement	<p>A code which indicates the student's or applicant's highest educational participation and last year of that participation prior to the first enrolment in the course of study</p> <p>Numeric 6</p> <p>000000 Overseas student</p> <p>010000 Not a commencing student</p> <p>A commencing domestic student for whom highest educational participation and last year of participation prior to commencement is OR an applicant's highest educational participation as reported on their application is:</p> <p>02YYYY A complete Higher education postgraduate level course</p> <p>03YYYY A complete Higher education bachelor level course</p> <p>04YYYY A complete Higher education sub-degree level course</p> <p>05YYYY An incomplete Higher education course</p> <p>07YYYY A complete final year of secondary education course at school or through a registered training organisation</p> <p>08YYYY Other qualification, complete or incomplete</p> <p>090000 No prior educational attainment</p> <p>10YYYY A complete VET award course</p> <p>11YYYY An incomplete VET award course</p>
Disability	<p>An 8-character code indicating responses to 3 questions about disabilities</p> <p>Numeric</p>

8	<p>1st character—disability, impairment or long-term medical condition</p> <p>0 Student did not indicate the existence of a disability, impairment or long-term medical condition</p> <p>1 Student indicated that she/he has a disability, impairment or long-term medical condition</p> <p>2 Student indicated that she/he does not have a disability, impairment or long-term medical condition</p> <p>2nd character—hearing disability</p> <p>0 Student did not indicate she/he has a hearing disability</p> <p>1 Student indicated that she/he has a disability, impairment or long-term medical condition and that she/he has a hearing disability</p> <p>3rd character—learning disability</p> <p>0 Student did not indicate she/he has a learning disability</p> <p>1 Student indicated that she/he has a disability, impairment or long-term medical condition and that she/he has a learning disability</p> <p>4th character—mobility disability</p> <p>0 Student did not indicate she/he has a mobility disability</p> <p>1 Student indicated that she/he has a disability, impairment or long-term medical condition and that she/he has a mobility disability</p> <p>5th character—visual disability</p> <p>0 Student did not indicate she/he has a visual disability</p> <p>1 Student indicated that she/he has a disability, impairment or long-term medical condition and that she/he has a visual disability</p> <p>6th character—medical disability</p> <p>0 Student did not indicate she/he has a medical disability</p> <p>1 Student indicated that she/he has a disability, impairment or long-term medical condition and that she/he has a medical disability</p> <p>7th character—other disability</p> <p>0 Student did not indicate she/he has an other disability</p> <p>1 Student indicated that she/he has a disability, impairment or long-term medical condition and that she/he has an other disability</p> <p>8th character—support services, equipment and facilities</p> <p>0 Student did not indicate that she/he has a disability</p> <p>1 Student indicated that she/he has a disability, impairment or long-term medical condition and that she/he would like to receive advice on support services, equipment and facilities</p> <p>2 Student indicated that she/he has a disability, impairment or long-term medical condition and that she/he does not wish to receive advice on support services, equipment and facilities</p>
---	--

Source: Australian Government Department of Education 2013.

National Apprentice and Trainee Collection

The purpose of the apprentice and trainee collection is to provide data on all people employed under a training contract. This includes apprentices and trainees. Publications provide a summary of training activity in apprenticeships and traineeships in Australia, including information on training rates and duration of training.

Reporting frequency

Information is collected quarterly and released annually in August the following year.

Scope

The collection includes all apprentices and trainees with an apprenticeship/traineeship training contract.

Collection method

Administrative data collection. State and territory training authorities report data to the NCVER.

Metadata

Apprentice and Trainee data standard

<<http://www.ncver.edu.au/publications/1965.html>>

Terms and definitions (refer to additional notes)

<http://www.ncver.edu.au/wps/poc?urile=wcm:path:/wps/wcm/connect/NCVER_Shared/Collections/apprentice-and-trainee-quarterly-collection>

Data quality statement

There are plans to publish a data quality statement in line with other collections.

The notes on tables accompanying every publication also contain data quality information. Refer to the terms and definitions and notes on tables in related documents in:

<http://www.ncver.edu.au/wps/poc?urile=wcm:path:/wps/wcm/connect/NCVER_Shared/Collections/apprentice-and-trainee-quarterly-collection>

Data items for mapping

Data item	Data description
Country identifier	Country identifier classification is based on the Standard Australian Classification of Countries (SACC), ABS cat. no.1269.0, 2011 (second edition, revision 1) Alphanumeric 4 0000–9999 Valid 4-digit SACC code
Date of birth	The date a client was born Alphanumeric 8 DDMMYYYY Valid date @@MMYYYY Valid month and year but day not specified @@@YYYY Valid year but day and month not specified
Training organisation identifier	A unique code used to identify a training organisation Alphanumeric

Data item	Data description
	10 text Unique alphanumeric identifier If the training provider is a registered training organisation, the training organisation identifier must be the valid national code listed on the National Training Register <www.training.gov.au>.
Address location—suburb, locality or town	The name of a suburb, locality or town of a physical geographic location Alphanumeric 50 text Name of suburb, locality or town
Postcode	Identifies the Australia Post postcode of a physical location or a postal address Alphanumeric 4 0001–9999 4-digit Australia Post postcode valid during the collection year OSPC Overseas address location 0000 Postcode unknown
State identifier	Uniquely identifies the state or territory of a physical location or postal address Numeric 2 01 New South Wales 02 Victoria 03 Queensland 04 South Australia 05 Western Australia 06 Tasmania 07 Northern Territory 08 Australian Capital Territory 09 Other Australian territories or dependencies 99 Other (overseas but not an Australian territory or dependency)
Client identifier—apprenticeships	Identifies a client with an apprenticeship/traineeship training contract Alphanumeric 10 text Unique apprenticeship client identifier
Client identifier—TYIMS	Uniquely identifies a client on Department of Industry, Innovation, Climate Change, Science, Research and Tertiary Education's (DIICCSRTE) Training Youth Information Management System (TYIMS) Alphanumeric 10 1–9999999999 Client identifier issued by DIICCSRTE and recorded in TYIMS
Indigenous status identifier	Indicates a client who self-identifies as being of Australian Aboriginal or Torres Strait Islander descent The classification is based on the Standards for Statistics on Cultural and Language Diversity, ABS cat. no.1289.0, 1999 Alphanumeric 1 1 Yes, Aboriginal 2 Yes, Torres Strait Islander 3 Yes, Aboriginal AND Torres Strait Islander 4 No, Neither Aboriginal nor Torres Strait Islander The value below is not valid for the National VET Provider Collection or clients with apprenticeship contracts commenced or recommenced after 1/1/02 9 Yes—Client is of Aboriginal and/or Torres Strait Islander origin

Data item	Data description												
Language identifier	<p>Uniquely identifies the main language other than English spoken at home by the client</p> <p>This classification is based on the Australian Standard Classification of Languages (ASCL), ABS cat. no.1267.0, 2011 (second edition, revision 1)</p> <p>If English is the only language spoken at home, the Language identifier must be '1201 — English'</p> <p>In AVETMISS, '@@@@' is used for 'Not stated' rather than the Australian Bureau of Statistics code of '0002'</p> <p>Alphanumeric 4</p> <table border="0"> <tr> <td>0000–9999</td> <td>Valid 4-digit ASCL code</td> </tr> <tr> <td>0000</td> <td>Unknown</td> </tr> <tr> <td>0001</td> <td>Non-verbal</td> </tr> </table>	0000–9999	Valid 4-digit ASCL code	0000	Unknown	0001	Non-verbal						
0000–9999	Valid 4-digit ASCL code												
0000	Unknown												
0001	Non-verbal												
Name for encryption	<p>Contains the client's full name in the defined format. A non-reversible encryption process encrypts the name, facilitating analysis of unit-record data while ensuring the anonymity of individual clients</p> <p>Family name (comma) (space) first given name (space) second given name. If the name for encryption with comma and spaces exceeds 60 characters, enter client's full name in the above order and truncate at 60 characters</p> <p>Alphanumeric 60</p> <table border="0"> <tr> <td>text</td> <td>Client's full name in the specified format, e.g. family name (comma)(space)first given name (space) second given name</td> </tr> </table>	text	Client's full name in the specified format, e.g. family name (comma)(space)first given name (space) second given name										
text	Client's full name in the specified format, e.g. family name (comma)(space)first given name (space) second given name												
ANZSCO identifier	<p>Uniquely identifies the type of occupation that may be expected for those undertaking a program of study</p> <p>The classification is based on the Australian and New Zealand Standard Classification of Occupations (ANZSCO), ABS cat. no.1220.0, 2009 (first edition, revision 1)</p> <p>Alphanumeric 6</p> <table border="0"> <tr> <td>100000–899999</td> <td>Valid 6-digit ANZSCO code</td> </tr> <tr> <td>GEN19</td> <td>Occupational non-specific—general education</td> </tr> <tr> <td>GEN20</td> <td>Non-industry specific training</td> </tr> <tr> <td>NONVET</td> <td>Non-VET course—no occupational outcome</td> </tr> </table>	100000–899999	Valid 6-digit ANZSCO code	GEN19	Occupational non-specific—general education	GEN20	Non-industry specific training	NONVET	Non-VET course—no occupational outcome				
100000–899999	Valid 6-digit ANZSCO code												
GEN19	Occupational non-specific—general education												
GEN20	Non-industry specific training												
NONVET	Non-VET course—no occupational outcome												
Sex	<p>Identifies the gender of a client (male or female)</p> <p>Alphanumeric 1</p> <table border="0"> <tr> <td>F</td> <td>Female</td> </tr> <tr> <td>M</td> <td>Male</td> </tr> </table>	F	Female	M	Male								
F	Female												
M	Male												
Full-time identifier	<p>Identifies the basis on which the training contract is being undertaken</p> <p>Alphanumeric 1</p> <table border="0"> <tr> <td>1</td> <td>Full-time</td> </tr> <tr> <td>2</td> <td>Part-time</td> </tr> </table>	1	Full-time	2	Part-time								
1	Full-time												
2	Part-time												
Highest school level completed identifier	<p>Identifies the highest level of school that a client has completed</p> <p>Alphanumeric 2</p> <table border="0"> <tr> <td>02</td> <td>Did not go to school</td> </tr> <tr> <td>08</td> <td>Year 8 or below</td> </tr> <tr> <td>09</td> <td>Year 9 or equivalent</td> </tr> <tr> <td>10</td> <td>Completed Year 10</td> </tr> <tr> <td>11</td> <td>Completed Year 11</td> </tr> <tr> <td>12</td> <td>Completed Year 12</td> </tr> </table>	02	Did not go to school	08	Year 8 or below	09	Year 9 or equivalent	10	Completed Year 10	11	Completed Year 11	12	Completed Year 12
02	Did not go to school												
08	Year 8 or below												
09	Year 9 or equivalent												
10	Completed Year 10												
11	Completed Year 11												
12	Completed Year 12												
Prior educational achievement	<p>Uniquely identifies the level of non-schooling sector prior educational achievement</p>												

Data item	Data description
identifier	<p>successfully completed by a client</p> <p>This classification is based on a subset of the Australian Standard Classification of Education (ASCED), ABS cat. no.1272.0, 2001.</p> <p>Numeric</p> <p>3</p> <p>008 Bachelor degree or higher degree level (defined for AVETMISS use only)</p> <p>410 Advanced diploma or associate degree level</p> <p>420 Diploma level</p> <p>511 Certificate IV</p> <p>514 Certificate III</p> <p>521 Certificate II</p> <p>524 Certificate I</p> <p>990 Miscellaneous education</p>
Disability flag	<p>Indicates whether clients consider themselves to have a disability, impairment or long-term condition</p> <p>Alphanumeric</p> <p>1</p> <p>Y Yes—the client has a disability, impairment or long-term condition</p> <p>N No—the client does not have a disability, impairment or long-term condition</p>

Source: NCVET 2013a.

National Assessment Program—Literacy and Numeracy (NAPLAN)

NAPLAN tests provide rich, nationally comparable data on student performance across the range of student achievement, providing an indication of how students are performing. The same tests are used in each state and territory. ACARA is responsible for the management, development and reporting of the NAPLAN.

Reporting frequency

An annual assessment, first conducted in May 2008.

Scope

The program is a national assessment of all Year 3, 5, 7 and 9 students on tests covering numeracy, reading, writing and language conventions (spelling, punctuation and grammar).

Collection method

Test Administration Authorities are responsible for the implementation and administration of the NAPLAN tests in their jurisdiction, in accordance with nationally agreed protocols.

Metadata

Data catalogue:

<http://www.acara.edu.au/verve/_resources/Data_Catalogue.pdf>

Data quality statement

ACARA provides data quality statements to the Productivity Commission as part of its coordination of national reports. An example of this can be found at p. 269 of the 2012 NEA Performance Report at:

<http://www.pc.gov.au/data/assets/pdf_file/0006/128670/national-education-agreement-2012.pdf>

Data items for mapping

Data item	Data description
DOB	Date of Birth in DDMMYYYY format including leading 0s Numeric 8
School ID	School code, left zero-padded to 9 digits Alphanumeric 9
School name	School - Official Name (of Campus) School name; e.g., Sydney Church of England Co-educational Grammar School, Redlands Campus No further information available in ACARA data catalogue
Geolocation	Numeric 1 Derived from MCEECDYA Geographical Location Classification School Geo-location: 1=Metropolitan 2=Provincial 3=Remote 4=Very remote

Data item	Data description
Student ID	Within each school, each student should have a unique ID [note that these identifiers are not consistent over time] Alphanumeric 15
Indigenous status	Numeric 1 1=Aboriginal but not Torres Strait Islander Origin 2=Torres Strait Islander but not Aboriginal Origin 3=Both Aboriginal and Torres Strait Islander Origin 4=Neither Aboriginal nor Torres Strait Islander Origin 9=Not Stated/Unknown
Main language other than English spoken at home	The ABS coding index to link responses to the 'Main language other than English spoken at home' question to the 1267.0—Australian Standard Classification of Languages (ASCL), 2011 Numeric 4 0000 = Inadequately described 0001 = Non-verbal, so described 0002 = Not stated 1201 = English All other 4-digit codes as specified in the ASCL (2011)
Parent/guardian 1 school education	The highest level of primary or secondary education attained by the student's parent/guardian 1 Numeric 1 1 = Year 9 or equivalent or below 2 = Year 10 or equivalent 3 = Year 11 or equivalent 4 = Year 12 or equivalent 9 = Not stated/Unknown
Parent/guardian 2 school education	The highest level of primary or secondary education attained by the student's parent/guardian 2 Numeric 1 1 = Year 9 or equivalent or below 2 = Year 10 or equivalent 3 = Year 11 or equivalent 4 = Year 12 or equivalent 9 = Not stated/Unknown
Parent/guardian 1 non-school education	The highest level of post-school qualification attained by the student's parent/guardian 1 Numeric 1 5 = Certificate I to IV (including trade certificate) 6 = Advanced diploma/Diploma 7 = Bachelor degree or above 8 = No non-school qualification 9 = Not stated/unknown
Parent/guardian 2 non-school education	The highest level of post-school qualification attained by the student's parent/guardian 1 Numeric 1 5 = Certificate I to IV (including trade certificate) 6 = Advanced diploma/Diploma 7 = Bachelor degree or above 8 = No non-school qualification

Data item	Data description
	9 = Not stated/unknown
Parent/guardian 1 occupation group	The current occupation of employed work of parent/guardian 1 Numeric 1 1 = Senior management in large business organisation, government administration and defence, and qualified professionals 2 = Other business managers, arts/media/sportspersons and associate professionals 3 = Tradesmen/women, clerks and skilled office, sales and service staff 4 = Machine operators, hospitality staff, assistants, labourers and related workers 8 = Not in paid work in last 12 months 9 = Not stated/unknown
Parent/guardian 2 occupation group	The current occupation of employed work of parent/guardian 2 Numeric 1 1 = Senior management in large business organisation, government administration and defence, and qualified professionals 2 = Other business managers, arts/media/sportspersons and associate professionals 3 = Tradesmen/women, clerks and skilled office, sales and service staff 4 = Machine operators, hospitality staff, assistants, labourers and related workers 8 = Not in paid work in last 12 months 9 = Not stated/unknown
Sex	Numeric 1 1=Male 2=Female
Street name	Street address of the school No other information available in ACARA data catalogue
Suburb	School—Address Suburb of the school e.g. Garran No other information available in ACARA data catalogue
State	Refers to the location state of the school NSW VIC QLD SA WA NT ACT TAS
Postcode	School—Address Postcode of the school e.g. 2000 No other information available in ACARA data catalogue
Sector	School sector. Contained within Student Level Data Alphanumeric 1 I Independent C Catholic G Government H Home School
Student grade level	Refers to the data for the respective year level Year 3 Year 5

Data item	Data description
	Year 7 Year 9

Source: ACARA 2013.

National Assessment Program Sample Assessments

NAP Sample Assessments provide nationally comparable data on student performance across the domains of science literacy, civics and citizenship and ICT literacy. The same tests are used in each state and territory. ACARA is responsible for the management, development and reporting of the NAP Sample Assessments.

Reporting frequency

Rolling triennial collections from 2003.

Scope

The program is an assessment of selected groups of students in Years 6 and 10 covering students' skills and understanding in science literacy, civics and citizenship and ICT literacy.

Collection method

Test Administration Authorities are responsible for the implementation and administration of the NAP Sample Assessment tests in their jurisdiction, in accordance with nationally agreed protocols.

Metadata

Data catalogue:

<http://www.acara.edu.au/verve/_resources/Data_Catalogue.pdf>

Data quality statement

ACARA provides data quality statements to the Productivity Commission as part of its coordination of national reports. An example of this can be found at p. 269 of the 2012 NEA Performance Report at:

<http://www.pc.gov.au/data/assets/pdf_file/0006/128670/national-education-agreement-2012.pdf>

Data items for mapping

Data item	Data description
DOB	Date of Birth in DDMMYYYY format including leading 0s Numeric 8
School ID	School code, left zero-padded to 9 digits Alphanumeric 9
School name	School - Official Name (of Campus) School name; e.g., Sydney Church of England Co-educational Grammar School, Redlands Campus No further information available in ACARA data catalogue
Geolocation	Numeric 1 Derived from MCEECDYA Geographical Location Classification School Geo-location: 1=Metropolitan 2=Provincial 3=Remote 4=Very remote

Data item	Data description
Student ID	Within each school, each student should have a unique ID [note that these identifiers are not consistent over time] Alphanumeric 15
Indigenous status	Numeric 1 1=Aboriginal but not Torres Strait Islander Origin 2=Torres Strait Islander but not Aboriginal Origin 3=Both Aboriginal and Torres Strait Islander Origin 4=Neither Aboriginal nor Torres Strait Islander Origin 9=Not Stated/Unknown
Main language other than English spoken at home	The ABS coding index to link responses to the 'Main language other than English spoken at home' question to the 1267.0—Australian Standard Classification of Languages (ASCL), 2011 Numeric 4 0000 = Inadequately described 0001 = Non-verbal, so described 0002 = Not stated 1201 = English All other 4-digit codes as specified in the ASCL (2011)
Parent/guardian 1 school education	The highest level of primary or secondary education attained by the student's parent/guardian 1 Numeric 1 1 = Year 9 or equivalent or below 2 = Year 10 or equivalent 3 = Year 11 or equivalent 4 = Year 12 or equivalent 9 = Not stated/Unknown
Parent/guardian 2 school education	The highest level of primary or secondary education attained by the student's parent/guardian 2 Numeric 1 1 = Year 9 or equivalent or below 2 = Year 10 or equivalent 3 = Year 11 or equivalent 4 = Year 12 or equivalent 9 = Not stated/Unknown
Parent/guardian 1 non-school education	The highest level of post-school qualification attained by the student's parent/guardian 1 Numeric 1 5 = Certificate I to IV (including trade certificate) 6 = Advanced diploma/Diploma 7 = Bachelor degree or above 8 = No non-school qualification 9 = Not stated/unknown
Parent/guardian 2 non-school education	The highest level of post-school qualification attained by the student's parent/guardian 1 Numeric 1 5 = Certificate I to IV (including trade certificate) 6 = Advanced diploma/Diploma 7 = Bachelor degree or above 8 = No non-school qualification 9 = Not stated/unknown

Data item	Data description
Parent/guardian 1 occupation group	The current occupation of employed work of parent/guardian 1 Numeric 1 1 = Senior management in large business organisation, government administration and defence, and qualified professionals 2 = Other business managers, arts/media/sportspersons and associate professionals 3 = Tradesmen/women, clerks and skilled office, sales and service staff 4 = Machine operators, hospitality staff, assistants, labourers and related workers 8 = Not in paid work in last 12 months 9 = Not stated/unknown
Parent/guardian 2 occupation group	The current occupation of employed work of parent/guardian 2 Numeric 1 1 = Senior management in large business organisation, government administration and defence, and qualified professionals 2 = Other business managers, arts/media/sportspersons and associate professionals 3 = Tradesmen/women, clerks and skilled office, sales and service staff 4 = Machine operators, hospitality staff, assistants, labourers and related workers 8 = Not in paid work in last 12 months 9 = Not stated/unknown
Sex	Numeric 1 1=Male 2=Female
Suburb	School—Address Suburb of the school e.g. Garran No other information available in ACARA data catalogue
State	Refers to the location state of the school. NSW VIC QLD SA WA NT ACT TAS
Postcode	School—Address Postcode of the school e.g. 2000 No other information available in ACARA data catalogue
Sector	School sector. Contained within Student Level Data Alphanumeric 1 I Independent C Catholic G Government H Home School
Student grade level	Refers to the data for the respective year level Year 6 Year 10

Source: ACARA 2013.

National Early Childhood Education and Care Collection

This collection was established to provide nationally comparable statistics on early childhood education and care (ECEC). The collection also enables improvements in the quality and accessibility of national ECEC data and assists in the progress reporting for the National Partnership Agreement on Early Childhood Education.

Reporting frequency

An annual collection with the collection of data undertaken on the first Friday in August each year.

Scope

All service providers delivering a preschool program during a reference period that includes the first Friday in August.

Collection method

Jurisdictions collect and report data for the National ECEC Collection using either a URL collection methodology or a combination of aggregate and URL collection methodologies.

Metadata

Data set specification:

Early Childhood Education and Care DSS 2014
<<http://meteor.aihw.gov.au/content/index.phtml/itemId/555371>>

Data collection guide:

<<http://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/4240.0.55.002?OpenDocument>>

Concepts, sources and methods:

<<http://www.abs.gov.au/ausstats/abs@.nsf/mf/4240.0.55.001>>

Data quality statement

Available annually from 2010. Latest available at:
<<http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/4240.0Quality%20Declaration02014?opendocument&tabname=Notes&prodno=4240.0&issue=2014&num=&view=>>>

Data items for mapping

Data item	Data description
Age	The age of the person in (completed) years at a specific point in time Numeric 3 Value as a number with a maximum of 3 characters 999 Unknown/not stated
Date of birth	DDMMYYYY If date of birth is not known or cannot be obtained, provision should be made to collect or estimate age. Collected or estimated age would usually be in years for adults, and to the nearest

Data item	Data description
	3 months (or less) for children aged less than 2 years.
Address line	A composite of one or more standard address components, as represented by text Alphanumeric 180
Suburb/town/locality name	The name of the locality/suburb of the address, as represented by text This item is collected twice: once for the child and once for the early childhood education and care service Alphanumeric 50
Australian postcode	The Australian numeric descriptor for a postal delivery area for an address This item is collected twice: once for the child and once for the early childhood education and care service Numeric 4
Statistical area level 1	A designated region representing the smallest unit for the release of Census data used for describing a location, as represented by a code This item is collected twice: once for the child and once for the early childhood education and care service Numeric 11
Area of usual residence (until 2013)	The geographical location of the usual residence of the person, using a 7-digit numerical code which indicates the Collection District (CD) within the state or territory of Australia The single-digit codes for the states and territories and the 6-digit codes for the CDs are as defined in the Australian Standard Geographical Classification (ASGC). Numeric 7 NNNNNNN
Australian state/territory identifier	The Australian state or territory where a person can be located, as represented by a code Numeric 1 1 New South Wales 2 Victoria 3 Queensland 4 South Australia 5 Western Australia 6 Tasmania 7 Northern Territory 8 Australian Capital Territory 9 Other territories (Cocos (Keeling) Islands, Christmas Island and Jervis Bay Territory)
Identifier code—SLK	A key that enables 2 or more records belonging to the same individual to be brought together. It is represented by a code consisting of the second, third and fifth characters of a person's family name, the second and third letters of the person's given name, the day, month and year when the person was born and the sex of the person, concatenated in that order. Alphanumeric 14 XXXXDDMMYYYYN
Indigenous status	Whether a person identifies as being of Aboriginal or Torres Strait Islander origin, as represented by a code Numeric 1 1 Aboriginal but not Torres Strait Islander origin 2 Torres Strait Islander but not Aboriginal origin 3 Both Aboriginal and Torres Strait Islander origin

Data item	Data description
	4 Neither Aboriginal nor Torres Strait Islander origin 9 Not stated/inadequately described
Letters of family name	The combination of 2nd, 3rd and 5th letters of a person's family name Alphanumeric 3 XXX
Letters of given name	The combination of the 2nd and 3rd letters of a person's given name Alphanumeric 2 XX
Sex	The biological distinction between male and female, as represented by a code Numeric 1 1 Male 2 Female 3 Intersex or indeterminate 9 Not stated/inadequately described
Organisation name	The full title of an organisation's name by which it trades or is recognised, as represented by text Alphanumeric 200
Service provider organisation—Address line	A composite of one or more standard address components, as represented by text Alphanumeric 180
Preschool program repeater indicator	An indicator of whether the child was enrolled in their second or third year of a preschool program, as represented by a code Numeric 1 1 Yes 2 No 3 Unknown 9 Not stated/inadequately described
Service activity type	The main type of activity available from or provided by an early childhood education and care service provider, as represented by a code Numeric 2 1 Long day care 2 Occasional care 3 Outside school hours care 4 Vacation care 5 Family day care 6 In-home care 7 Preschool program: stand-alone 8 Preschool program: as part of a school 99 Not stated/inadequately described
Preschool program attendance indicator	An indicator of whether a child attended a preschool program, as represented by a code Numeric 1 1 Yes 2 No

Data item	Data description
Preschool program enrolment indicator	<p>An indicator of whether the child was enrolled in a preschool program, as represented by a code</p> <p>Numeric</p> <p>1</p> <p>1 Yes</p> <p>2 No</p>
Service provider organisation—Australian state/territory identifier code	<p>An identifier of the Australian state or territory where an organisation or agency can be located, as represented by a code</p> <p>Numeric</p> <p>1</p> <p>1 New South Wales</p> <p>2 Victoria</p> <p>3 Queensland</p> <p>4 South Australia</p> <p>5 Western Australia</p> <p>6 Tasmania</p> <p>7 Northern Territory</p> <p>8 Australian Capital Territory</p> <p>9 Other territories (Cocos (Keeling) Islands, Christmas Island and Jervis Bay Territory)</p>

Source: AIHW 2014b.

National Early Childhood Education and Care Workforce Census

The National Workforce Census helps to inform planning for the early childhood education and care sector, which is important to ensure that high-quality early childhood services are provided to young children and their families.

The census consists of 2 surveys:

- The Services Survey for child care services collects information on service usage, children with additional needs, access to preschool programs and staff details.
- The Staff Survey collects information on the roles, qualifications, employment status, training activities, pay and conditions and career intentions of staff.

Reporting frequency

The census is conducted 3-yearly

Scope

The scope of the collection includes all child care providers that are Department of Education approved to receive Child Care Benefit:

- long day care services
- family day care services
- in-home care services
- occasional care services
- out of school hours care services
- vacation care services
- all government/non-government preschool/kindergarten services, including stand-alone and those located with a school.

Collection method

Data are collected in 2 parts via self-completion. The first part involves the collection of site-level information from preschool services and approved child care services in a reference week (services survey). The second part involves the collection of information from individual staff through the voluntary staff survey. Both surveys collect information about staff.

Metadata

There are 2 data dictionaries for the National Workforce Census Services Survey and Staff Survey.

There are also protocols and standards that govern the use of the National Workforce Census. These include:

- National Workforce Census 2013 Data Protocol
- Early Childhood Education and Care National Minimum Data Set.

Data quality statement

There is no data quality statement for this collection.

Issues relating to data quality are documented in the national report and footnotes of any data tables that are produced.

Data items for mapping

Data item	Data description
Age	15–17 years 18–19 years 20–24 years 25–29 years 30–34 years 35–39 years 40–44 years 45–49 years 50–54 years 55 years and over Not answered No other information available in the Staff Data Variables list
State	New South Wales Victoria Queensland South Australia Western Australia Tasmania Northern Territory ACT No other information available in the Staff Data Variables list
Worker sex	Male=1 Female=2 Intersex or indeterminate Not answered No other information available in the Staff Data Variables list
Remoteness area name	Inner Regional Australia Major Cities of Australia Outer Regional Australia Remote Australia Very Remote Australia No other information available in the Staff Data Variables list
Unique staff ID	Unique ID No other information available in the Staff Data Variables list
Indigenous status	1=Aboriginal 2=Torres Strait Islander 3=Both 1 & 2 4=Non-Indigenous No other information available in the Staff Data Variables list
Income	\$2,000 or more per week (\$104,000 or more per year) \$1,500–\$1,999 per week (\$78,000–\$103,999 per year)

Data item	Data description
	\$1,250–\$1,499 per week (\$65,000–\$77,999 per year) \$1,000–\$1,249 per week (\$52,000–\$64,999 per year) \$800–\$999 per week (\$41,600–\$51,999 per year) \$600–\$799 per week (\$31,200–\$41,599 per year) \$400–\$599 per week (\$20,800–\$31,199 per year) \$300–\$399 per week (\$15,600–\$20,799 per year) \$200–\$299 per week (\$10,400–\$15,599 per year) \$100–\$199 per week (\$5,200–\$10,399 per year) \$99 or less per week (less than \$5,199 per year) Unknown Not answered No other information available in the Staff Data Variables list
Highest level of relevant qualification (teaching)/ Highest attainment non-teaching qualification	Postgraduate degree Graduate diploma or graduate certificate Bachelor degree honours Bachelor degree pass (4 years or equivalent) Bachelor degree pass (3 years or equivalent) Advanced diploma Diploma Certificate level IV Certificate level III Certificate level I or II Other certificate Not answered No other information available in the Staff Data Variables list
Unique service identifier	Unique ID No other information available in the Services Data Variables list

Source: Australian Government Department of Education (unpublished).

National Schools Attendance Collection

The National Schools Attendance Collection is an aggregate collection of attendance rates of students in Years 1 to 10 in government and non-government schools.

Attendance rates are available by the following disaggregation levels: school sector, jurisdiction, year level, sex, and Indigenous status.

Reporting frequency

Annual collection from 2007 (conducted by ACARA since 2010, and prior to that by the SCSEEC Secretariat).

Scope

All students in Years 1–10 in government and non-government schools.

Collection method

Data for government schools are collected by ACARA from the state and territory education departments. For non-government schools (independent and Catholic), the collection is carried out by the Australian Government Department of Education and Training, and the data are then provided to ACARA.

Metadata

Data catalogue:

Refer to Appendix 3.4: <http://www.acara.edu.au/verve/_resources/Data_Catalogue.pdf>

National Standards for Student Attendance Data Reporting:

<http://www.acara.edu.au/verve/_resources/D13_20398__National_Standards_for_Student_Attendance_Data_Reporting_-_UPDATED_November_2013.pdf>

Data quality statement

ACARA provides data quality statements to the Productivity Commission as part of its coordination of national reports. An example of this can be found at p. 269 of the 2012 NEA Performance Report at:

<http://www.pc.gov.au/data/assets/pdf_file/0006/128670/national-education-agreement-2012.pdf>.

Data items for mapping

Aggregate collection, unable to be mapped.

National Schools Statistics Collection

The National Schools Statistics Collection (NSSC) is a census, conducted annually as a collaborative arrangement between state, territory and Australian Government education authorities and the ABS. School, student and staff data for the government sector are compiled administratively by state and territory education authorities. The NSSC is primarily based on aggregate data with some states and territories supplying government school data in unit record form. The ABS engages bilaterally with data custodians to increase the proportion of data supplied in unit record form.

Reporting frequency

The first collection was conducted in 1981 and it is now conducted annually on the first Friday in August each year.

Scope

Data are collected from the relevant authorities on a range of issues relating to schools, students and staff in primary and secondary schools throughout Australia, from both the government and non-government sectors.

Collection method

Collection methodologies vary between the different state and territory departments of education. Data may be accessed from central administrative records or collected directly from education establishments. The Australian Government Department of Education and Training collects data directly from establishments in the Independent sector for all states and territories for administrative purposes. The Catholic sector collects data for catholic schools for administrative purposes and provides this to the ABS via the Australian Government Department of Education and Training. The non-government sector statistics in this publication are a summary of results from these collections.

Metadata

Notes, Instructions and Tabulations document

Available on request from the ABS

Data Collection Manual

Available on request from the ABS

Data quality statement

Available annually from 2011. Latest available at:

<http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/4221.0Quality%20Declaration02014?opendocument&tabname=Notes&prodno=4221.0&issue=2014&num=&view=>

Data items for mapping

Some of the following data items are not currently collected by the ABS through the NSSC at the unit record level for all states and territories. Addressing the full access and standards improvements required for school enrolment data requires consultation with both the ABS and the relevant state and territory custodians.

Data item	Data description
Student's country of birth	The country in which the student was born Numeric 4 The country of birth field is a 4-digit, 3-level hierarchical coding structure using the ABS: Standard Australian Classification of Countries (SACC), Second Edition (cat. no. 1269.0).
Date of birth	The day, month and year on which the student was born Numeric 8 DDMMYYYY 00 = missing day value 00 = missing month value 0000 = missing year value
System school ID	A unique authorisation number allocated to registered schools by the system administering those schools Alphanumeric
Area of usual residence Mesh Block code	The place where the student usually resides for the majority of their time during the given enrolment as represented by the ASGS geographical area code Mesh Block (MB). This is assumed to be based on where the primary carer for the student (parent/guardian 1 or 2) resides. It is represented by a geographical area code derived from the Australian Statistical Geography Standard (ASGS) 2011. Numeric 11 The Area of Usual Residence ASGS MB is a 11-digit hierarchical coding structure using the ABS: Australian Statistical Geography Standard (ASGS) 2011. Where this item is not available the following codes should be used: 10000009899—New South Wales: undefined MB 20000009899—Victoria: undefined MB 30000009899—Queensland: undefined MB 40000009899—South Australia: undefined MB 50000009899—Western Australia: undefined MB 60000009899—Tasmania: undefined MB 70000009899—Northern Territory: undefined MB 80000009899—Australian Capital Territory: undefined MB 90000009899—Other Territories: undefined MB
System student ID	The ID used by the enrolling system for administration of the student and their enrolment(s). String various alphanumeric length Length and composition of the System Student ID may vary between enrolling education systems
Statistical linkage key	A string containing the letters of the student's family (last) name and given (first) name, and the digits of the student's date of birth and sex String 14 This item is derived by combining 4 student-level items: letters of family (last) name,

Data item	Data description
	letters of given (first) name, date of birth and sex.
ATSI status	<p>The status of an enrolment showing whether the enrolled person is a person of Aboriginal and/or Torres Strait Islander descent who identifies as an Aboriginal and/or Torres Strait Islander and is accepted as such by the community in which he or she lives</p> <p>Numeric</p> <p>1</p> <p>1 = Aboriginal but not Torres Strait Islander origin</p> <p>2 = Torres Strait Islander but not Aboriginal origin</p> <p>3 = Both Aboriginal and Torres Strait Islander origin</p> <p>4 = Neither Aboriginal nor Torres Strait Islander origin</p> <p>9 = Not Stated/inadequately described</p>
Student's main language other than English spoken at home	<p>The main language other than English, spoken in the home by the student</p> <p>Numeric</p> <p>4</p> <p>The 'Main Language other than English' is a 4-digit, 3-level hierarchical coding structure using the ABS: Australian Standard Classification of Languages (ASCL) second edition (ABS cat. no. 1267.0).</p> <p>Common codes:</p> <p>0000—Inadequately described</p> <p>0001—Non verbal, so described</p> <p>0002—Not stated</p> <p>1201—English</p>
Letters of student's enrolled family name	<p>The second, third and fifth letters of the student's family name (or surname)</p> <p>Alphanumeric</p> <p>3</p> <p>XXX</p> <p>999 = Not stated/inadequately described</p> <p>Where letters are missing (i.e. family name (or surname) is less than 5 letters), replace the missing letters with '2'</p>
Letters of student's enrolled given name	<p>The second and third letters in the student's first (or given) name</p> <p>Alphanumeric</p> <p>2</p> <p>XX</p> <p>99 = Not stated/inadequately described</p> <p>Where letters are missing (i.e. first (or given) name is less than 3 letters), replace the missing letters with '2'</p>
Parent/guardian 1 school education	<p>The highest level of primary or secondary education attained by the student's parent/guardian 1</p> <p>Numeric</p> <p>1</p> <p>1 = Year 9 or equivalent or below</p> <p>2 = Year 10 or equivalent</p> <p>3 = Year 11 or equivalent</p> <p>4 = Year 12 or equivalent</p> <p>9 = Not stated/Unknown</p>
Parent/guardian 2 school education	<p>The highest level of primary or secondary education attained by the student's parent/guardian 2</p> <p>Numeric</p> <p>1</p> <p>1 = Year 9 or equivalent or below</p> <p>2 = Year 10 or equivalent</p> <p>3 = Year 11 or equivalent</p>

Data item	Data description
	4 = Year 12 or equivalent 9 = Not stated/Unknown
Parent/guardian 1 highest non-school qualification	The highest level of post-school qualification attained by the student's parent/guardian 1 Numeric 1 5 = Certificate I to IV (including trade certificate) 6 = Advanced diploma/Diploma 7 = Bachelor degree or above 8 = No non-school qualification 9 = Not stated/unknown
Parent/guardian 2 highest non-school qualification	The highest level of post-school qualification attained by the student's parent/guardian 2 Numeric 1 5 = Certificate I to IV (including trade certificate) 6 = Advanced diploma/Diploma 7 = Bachelor degree or above 8 = No non-school qualification 9 = Not stated/unknown
Parent/guardian 1 occupation group	The current occupation of employed work of parent/guardian 1 Numeric 1 1 = Senior management in large business organisation, government administration and defence, and qualified professionals 2 = Other business managers, arts/media/sportspersons and associate professionals 3 = Tradesmen/women, clerks and skilled office, sales and service staff 4 = Machine operators, hospitality staff, assistants, labourers and related workers 8 = Not in paid work in last 12 months 9 = Not stated/unknown
Parent/guardian 2 occupation group	The current occupation of employed work of parent/guardian 2 Numeric 1 1 = Senior management in large business organisation, government administration and defence, and qualified professionals 2 = Other business managers, arts/media/sportspersons and associate professionals 3 = Tradesmen/women, clerks and skilled office, sales and service staff 4 = Machine operators, hospitality staff, assistants, labourers and related workers 8 = Not in paid work in last 12 months 9 = Not stated/unknown
Sex	The biological sex of the student Numeric 1 1 = Male 2 = Female 3 = Intersex or Indeterminate 9 = Not stated/inadequately described
School address—Apartment/flat number	The apartment/flat number of the school's physical address Alphanumeric 5 'Not Applicable' or 'Not provided' responses left blank
School address—Street number	The street number of the school's physical address

Data item	Data description
	Alphanumeric 5 'Not Applicable' or 'Not provided' responses left blank
School address—Street name	The street name of the school's physical address Alphanumeric 18 'Not Applicable' or 'Not provided' responses left blank
School address—suburb	The suburb of the school's physical address Alphanumeric 18 'Not Applicable' or 'Not provided' responses left blank
State/territory	The state or territory that is responsible for administration of a school School state/territory will be derived by the ABS based on the source of a data file or information provided on the Department of Education and Training non-government data provision Numeric 1 1 NSW 2 Vic. 3 Qld 4 SA 5 WA 6 Tas. 7 NT 8 ACT
School address—Postcode	The postcode of the school's physical address Numeric 4
School affiliation	The system that administers the school Government: An establishment administered by the department/ministry of education under directors-general of education (or equivalent) (as defined by membership of the Conference of Education Systems Chief Executive Officers (CESCO)) Non-Government: Any establishment not administered by the departments of education, including those establishments administered by any other government authority Subcategories of non-government, as provided by the Department of Education and Training through the non-government NSSC data submission and as self-reported by the schools include: Catholic—including systemic and non-systemic Independent—including Anglican and all other non-government affiliated schools Numeric 2 1 Government 21 Catholic 22 Independent
Full-time or part-time status of enrolment	A flag showing whether the student's enrolment is full-time or part-time This item is derived from the Full-Time Equivalent (FTE) Value of Enrolment (EN_FTEVAL). Numeric 1 1 Full-time 2 Part-time
Grade of student's enrolment	The incremental level of formal schooling that the enrolment relates to on a Pre-Year 1

Data item	Data description
	to Year 12 scale Numeric 3 100 Pre-Year 1 101 Year 1 102 Year 2 103 Year 3 104 Year 4 105 Year 5 106 Year 6 107 Year 7 primary 199 Ungraded primary 207 Year 7 secondary 208 Year 8 209 Year 9 210 Year 10 211 Year 11 212 Year 12 299 Ungraded secondary

Source: ABS 2013b.

Programme for International Student Assessment (PISA)

PISA is an international survey which is an initiative of the Organisation for Economic Co-operation and Development.

Reporting frequency

ACER is responsible for undertaking the collection in Australian schools every 3 years, with the first in 2000 and the most recent being conducted in 2012.

Scope

PISA collects information on student achievement and contextual information about students, teachers and schools, and derives educational indicators that can monitor differences (and similarities) over time. PISA assesses the skills and knowledge of a nationally representative sample of 15-year-old students in 3 core areas: reading, mathematics and science.

Collection method

Fifteen-year-old students from randomly selected schools worldwide take tests in the 3 subjects. Students and school principals also provide background information on the students' family background, approaches to learning and the way their schools are run.

Metadata

Data collection manuals

Not publicly available

Technical report

<<http://www.oecd.org/pisa/pisaproducts/pisa2012technicalreport.htm>>

Codebooks

<<http://pisa2012.acer.edu.au/downloads.php>>

Data analysis manual

<<http://www.oecd.org/pisa/pisaproducts/pisadataanalysismanualspssandsassecondeditio.htm>>

Data quality statement

Refer to Technical report:

<<http://www.oecd.org/pisa/pisaproducts/pisa2012technicalreport.htm>>

Data items for mapping

Data item	Data description
Country of birth national categories—self	Alphanumeric 3
Birth month	Alphanumeric 2 01 January 02 February 03 March 04 April 05 May 06 June 07 July 08 August 09 September 10 October 11 November 12 December 99 Missing
Birth year	Alphanumeric 4 1996 1997
School ID	Unique 5-digit school ID Alphanumeric 5
School sector	Numeric 1 1 Catholic 2 Government 3 Independent
Geographic location	Numeric 1 1 Metropolitan 2 Provincial 3 Remote MCEECDYA schools location classification
School address	Information not available
State	1 ACT 2 NSW 3 VIC 4 QLD 5 SA 6 WA 7 TAS 8 NT
Student ID	Alphanumeric 5
Indigenous background	Numeric 1

Data item	Data description
	1 No 2 Yes, Aboriginal 3 Yes, Torres Strait Islander [4 Both]
Language at home	Alphanumeric 3 105–999 valid
Family name	Information not available
Given name	Information not available
Country of birth International— Mother	Numeric 1 1 Country of test 2 Other country 7 N/A 8 Invalid 9 Missing
Country of birth International— Father	Numeric 1 1 Country of test 2 Other country 7 N/A 8 Invalid 9 Missing
Educational level of mother	International Standard Classification of Education Numeric 1 0 None 1 ISCED 1 2 ISCED 2 3 ISCED 3B, C 4 ISCED 3A, ISCED 4 5 ISCED 5B 6 ISCED 5A, 6 7 N/A 9 Missing Australia's responses based on: Completed PhD or equivalent doctoral program Completed a Masters Completed a university degree Completed TAFE Diploma Completed TAFE Certificate Completed post-secondary education but not university Completed Year 12 Completed Year 10 or 11 Completed some secondary school, but did not complete Year 10 Completed primary school only Did not complete primary school
Educational level of father	International Standard Classification of Education Numeric

Data item	Data description
	1 0 None 1 ISCED 1 2 ISCED 2 3 ISCED 3B, C 4 ISCED 3A, ISCED 4 5 ISCED 5B 6 ISCED 5A, 6 7 N/A 9 Missing Australia's responses based on: Completed PhD or equivalent doctoral program Completed a Masters Completed a university degree Completed TAFE Diploma Completed TAFE Certificate Completed post-secondary education but not university Completed Year 12 Completed Year 10 or 11 Completed some secondary school, but did not complete Year 10 Completed primary school only Did not complete primary school
Mother occupation	International Standard Classification of Occupation Alphanumeric 4 0000–9999 valid
Father occupation	International Standard Classification of Occupation Alphanumeric 4 0000–9999 valid
Sex	Numeric 1 1 female 2 male
Index of Economic, Social and Cultural Status	Derived from 3 variables: Index of highest level of parental education in number of years of education, index of highest parental occupation status and the index of home possessions Numeric 8.2 -5.95–9999.00 valid
Grade	Numeric 2 7–96 Valid PISA report includes the following grade levels: 7 8 9 10 11 12

Data item	Data description
Family structure	Numeric 1 1 Single parent (natural or otherwise) 2 Two parents (natural or otherwise) 3 Other 7 N/A 9 Missing

Source: ACER 2012.

Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS)

TIMSS is a large-scale assessment providing an international perspective on teaching and learning in mathematics and science.

PIRLS is a large-scale assessment providing an international perspective on teaching and learning in reading literacy.

TIMSS and PIRLS are projects of the International Association for the Evaluation of Educational Achievement (IEA).

Reporting frequency

TIMSS reports every 4 years on the achievement of Year 4 and Year 8 students worldwide.

PIRLS reports every 5 years on the achievement of Year 4 students worldwide. Australia participated for the first time in 2011.

Scope

TIMSS involves assessment of mathematics and science in Year 4 and Year 8.

PIRLS involves assessment of reading in Year 4.

Collection method

ACER is responsible for undertaking the data collections in Australian schools.

Metadata

Data collection manuals:

Not publicly available

Technical report (most recent cycle):

<<http://timssandpirls.bc.edu/methods/index.html>>

Codebooks (most recent cycle):

<<http://timssandpirls.bc.edu/timss2011/international-database.html>>

<<http://timssandpirls.bc.edu/pirls2011/international-database.html>>

Data quality summary

Refer to Technical report:

<<http://timssandpirls.bc.edu/methods/index.html>>

Data items for mapping

Data item	Data description
Born in country	Born in country of testing (Australia) 1 yes 2 no 9 omitted or invalid
School ID	Numeric 4 1–9996
School sector	Numeric 1 1 = Catholic 2 = Government 3 = Independent
Date of students birth day	Numeric 2 1–31 valid 99 omitted or invalid 98 not admin
Date of students birth month	Numeric 2 1–12 valid 99 omitted or invalid 98 not admin
Date of students birth year	Numeric 4 1993–2001 valid 9999 omitted or invalid 9998 not admin
Geographic location	Geographic location of the school—applied from the school address Numeric 1 1 Metropolitan 2 Provincial 3 Remote
School address	Information not available
State	1 ACT 2 NSW 3 VIC 4 QLD 5 SA 6 WA 7 TAS 8 NT
Student ID	Numeric 8 10101–99999996 valid
Indigenous background	Numeric 1 1 No

Data item	Data description
	2 Yes, Aboriginal 3 Yes, Torres Strait Islander [4 Both]
Often speak English at home	Alphanumeric 1 1 Always 2 Almost always 3 Sometimes 4 Never 9 omitted or invalid 8 not admin.
Family name	Information not available
Given name	Information not available
Mother/father born in Australia	Alphanumeric 1 1 Yes 2 No 9 Omitted or invalid
Highest level of education mother/father	Alphanumeric 1 ISCED Coding: 1 = Some <ISCED Level 1 or 2 > or did not go to school 2 = <ISCED Level 2> 3 = <ISCED Level 3> 4 = <ISCED Level 4> 5 = <ISCED Level 5B> 6 = <ISCED Level 5A, first degree> 7 = Beyond <ISCED Level 5A, first degree> 8 = I don't know Australian coding: 1 Some Primary school or did not go to school 2 Primary school 3 Secondary school 4 TAFE training certificate, apprenticeship or traineeship 5 TAFE or College Diploma (e.g. Diploma in Information Technology, Diploma in Veterinary Nursing) 6 Undergraduate degree (e.g. Bachelor of Arts, BA, Bachelor of Commerce) 7 Postgraduate degree (e.g. Honours, Master of Arts, PhD) 8 I don't know
Mother/father occupation	1 Has never worked for pay 2 Small Business Owner 3 Clerk 4 Service or Sales Worker 5 Skilled Agricultural or Fishery Worker 6 Craft or Trade Worker 7 Plant or Machine Operator 8 General Laborers 9 Corporate Manager or Senior Official 10 Professional 11 Technician or Associate Professional

Data item	Data description
	12 Not applicable
Sex of student	Alphanumeric 1 1 Girl 2 Boy
Economic disadvantage	Alphanumeric 1 1 0–10% 2 11–25% 3 26–50% 4 more than 50%
Grade ID	Numeric 2 7 Grade 7 8 Grade 8 9 Grade 9 99 Omitted or invalid

Source: ACER 2014; IEA 2013a; IEA 2013b.

VET Provider Collection

The collection contains information on publicly funded training programs delivered by government-funded and privately operated training providers.

Reporting frequency

Annual.

Scope

Information collected includes VET student characteristics, VET programs such as courses, and VET training provider details such as location. The Australian Vocational Education and Training Management Information Statistical Standard (AVETMISS) for VET Provider is the data standard for the National VET Provider Collection.

Collection method

Administrative data collection. Registered training providers and state and territory training authorities report data based on student enrolment records to the NCVER.

Metadata

AVETMISS data standard

<www.ncver.edu.au/avetmiss/>

Terms and definitions

Refer to 'Additional information' in:

<http://www.ncver.edu.au/wps/poc?urile=wcm:path:/wps/wcm/connect/NCVER_Shared/Collections/students-and-courses-2012>

Data quality statement

Data quality statements are produced for COAG Reform Council and Report on Government Services reporting.

The notes on tables accompanying every publication also contain data quality information.

Refer to 'Additional information' in:

<http://www.ncver.edu.au/wps/poc?urile=wcm:path:/wps/wcm/connect/NCVER_Shared/Collections/students-and-courses-2012>

Data items for mapping

Data item	Data description
Country identifier	Classification is based on the Standard Australian Classification of Countries (SACC), ABS cat. no.1269.0, 2011 (second edition, revision 1) Alphanumeric 4 0000–9999 Valid 4-digit SACC code
Date of birth	The date a client was born Alphanumeric 8 DDMMYYYY Valid date @@MMYYYY Valid month and year but day not specified @@@YYYY Valid year but day and month not specified

Data item	Data description
Training organisation identifier	A unique code used to identify a training organisation Alphanumeric 10 text Unique alphanumeric identifier If the training provider is a registered training organisation, the Training organisation identifier must be the valid national code listed on the National Training Register <www.training.gov.au>.
VET in Schools flag	Identifies whether a unit of competency/module enrolment is part of a VET in Schools program Alphanumeric 1 Y VET in Schools program N Not a VET in Schools program
Address building/property name	The official place name or common usage name for an address site, including the name of a building, Aboriginal community, homestead, building complex, agricultural property, park or unbounded address site Alphanumeric 50 text Name of building or property
Address flat/unit details	Identifies an address within a building/sub-complex Alphanumeric 30 text Flat, unit or apartment address details
Address street number	Identifies the number of the address in the street Alphanumeric 15 text Street or lot number
Address street name	Identifies the name and type of the street to the address site Alphanumeric 70 text Street name, type and suffix
Address first line	The first line of an address—provides a valid street number and name (including floor and building name if relevant) or post office box information Alphanumeric 50 text Street number and name (and floor and building name if relevant) or post office box information
Address second line	The second line of an address—provides a continuation of the valid name and number of the street (including floor and business name if relevant) or post office box information Alphanumeric 50 text Continuation of street number and name
Address location – suburb, locality or town	The name of a suburb, locality or town of a physical geographic location Alphanumeric 50 text Name of suburb, locality or town
Postcode	Identifies the Australia Post postcode of a physical location or a postal address Alphanumeric 4 0001–9999 4-digit Australia Post postcode valid during the collection year OSPC Overseas address location 0000 Postcode unknown

Data item	Data description
State identifier	<p>Uniquely identifies the state or territory of a physical location or postal address</p> <p>Numeric 2</p> <p>01 New South Wales 02 Victoria 03 Queensland 04 South Australia 05 Western Australia 06 Tasmania 07 Northern Territory 08 Australian Capital Territory 09 Other Australian territories or dependencies 99 Other (overseas but not an Australian territory or dependency)</p>
Statistical area level 1 identifier	<p>Identifies a geographic-based population group at the lowest level for which census data is reported</p> <p>Used to collect aggregated client usual residential address information that can be used to derive client socioeconomic status while protecting client privacy</p> <p>Based on the Australian Statistical Geography Standard (ASGS), ABS cat. no.1270.0, 2011</p> <p>Alphanumeric 11</p> <p>0000000001–9999999999 Valid ASGS Statistical area level 1 code</p>
Statistical area level 2 identifier	<p>Identifies a geographic-based population group at the second lowest level for which census data is reported</p> <p>Based on the Australian Statistical Geography Standard (ASGS), ABS cat. no.1270.0, 2011</p> <p>Used to collect aggregated client usual residential address information that can be used to derive client socioeconomic status while protecting client privacy</p> <p>Alphanumeric 9</p> <p>000000001–999999999 Valid ASGS Statistical area level 2 code</p>
Client identifier	<p>Uniquely distinguishes an individual within a training organisation</p> <p>Alphanumeric 10</p> <p>text Unique client identifier within the collection</p>
Client identifier-apprenticeships	<p>Identifies a client with an apprenticeship/traineeship training contract</p> <p>Alphanumeric 10</p> <p>text Unique apprenticeship client identifier</p>
Indigenous status identifier	<p>Indicates a client who self-identifies as being of Australian Aboriginal or Torres Strait Islander descent</p> <p>The classification is based on the Standards for Statistics on Cultural and Language Diversity, ABS cat. no.1289.0, 1999</p> <p>Alphanumeric 1</p> <p>1 Yes, Aboriginal 2 Yes, Torres Strait Islander 3 Yes, Aboriginal AND Torres Strait Islander 4 No, Neither Aboriginal nor Torres Strait Islander</p>
Language identifier	<p>Uniquely identifies the main language other than English spoken at home by the client</p> <p>This classification is based on the Australian Standard Classification of Languages (ASCL), ABS cat. no.1267.0, 2011 (second edition, revision 1).</p>

Data item	Data description
	<p>If English is the only language spoken at home, the Language identifier must be '1201—English'.</p> <p>In AVETMISS, '@@@@' is used for 'Not stated' rather than the Australian Bureau of Statistics code of '0002'.</p> <p>Alphanumeric 4</p> <p>0000–9999 Valid 4-digit ASCL code</p> <p>0000 Unknown</p> <p>0001 Non-verbal</p>
Name for encryption	<p>Contains the client's full name in the defined format. A non-reversible encryption process encrypts the name, facilitating analysis of unit-record data while ensuring the anonymity of individual clients.</p> <p>Family name (comma) (space) first given name (space) second given name. If the name for encryption with comma and spaces exceeds 60 characters, enter client's full name in the above order and truncate at 60 characters.</p> <p>Alphanumeric 60</p> <p>text Client's full name in the specified format, e.g. family name (comma)(space) first given name (space) second given name</p>
ANZSCO identifier	<p>Uniquely identifies the type of occupation that may be expected for those undertaking a program of study</p> <p>The classification is based on the Australian and New Zealand Standard Classification of Occupations (ANZSCO), ABS cat. no.1220.0, 2009 (first edition, revision 1)</p> <p>Alphanumeric 6</p> <p>100000–899999 Valid 6-digit ANZSCO code</p> <p>GEN19 Occupational non-specific—general education</p> <p>GEN20 Non-industry specific training</p> <p>NONVET Non-VET course—no occupational outcome</p>
Proficiency in spoken English	<p>The self-assessed level of ability to speak English asked of people who speak a language other than English at home</p> <p>Proficiency in spoken English identifier is based on the Language Standards, ABS cat. no.1200.0.55.005, 2012 (version 1.1).</p> <p>Alphanumeric 1</p> <p>1 Very well</p> <p>2 Well</p> <p>3 Not well</p> <p>4 Not at all</p>
Sex	<p>Identifies the gender of a client (male or female)</p> <p>Alphanumeric 1</p> <p>F Female</p> <p>M Male</p>
Highest school level completed identifier	<p>Identifies the highest level of school that a client has completed</p> <p>Alphanumeric 2</p> <p>02 Did not go to school</p> <p>08 Year 8 or below</p> <p>09 Year 9 or equivalent</p> <p>10 Completed Year 10</p> <p>11 Completed Year 11</p> <p>12 Completed Year 12</p>

Data item	Data description
Prior educational achievement identifier	<p>Uniquely identifies the level of non-schooling sector prior educational achievement successfully completed by a client</p> <p>This classification is based on a subset of the Australian Standard Classification of Education (ASCED), ABS cat. no.1272.0, 2001.</p> <p>Numeric 3</p> <p>008 Bachelor degree or higher degree level (defined for AVETMISS use only)</p> <p>410 Advanced diploma or associate degree level</p> <p>420 Diploma level</p> <p>511 Certificate IV</p> <p>514 Certificate III</p> <p>521 Certificate II</p> <p>524 Certificate I</p> <p>990 Miscellaneous education</p>
Subject field of education identifier	<p>Based on the field of education (FOE) at the detailed level (6-digit), which is 1 part of the Australian Standard Classification of Education (ASCED), ABS cat. no.1272.0, 2001</p> <p>The FOE is the subject matter of the unit of competency or module.</p> <p>Alphanumeric 6</p> <p>BBNDD Valid 6-digit detailed level FOE code</p>
Program field of education identifier	<p>Identifies the subject matter that is the ultimate aim of the skills and knowledge gained in a qualification, course or skill set</p> <p>Based on the FOE at the narrow level (4-digit), which is one part of the Australian Standard Classification of Education (ASCED), ABS cat. no.1272.0, 2001</p> <p>Numeric 4</p> <p>BBNN Valid 4-digit narrow level FOE code</p>
Outcome identifier-national	<p>Identifies the result or outcome of a client's participation in a unit of competency or module</p> <p>Numeric Assessable</p> <p>20 Competency achieved/pass</p> <p>30 Competency not achieved/fail</p> <p>40 Withdrawn/discontinued</p> <p>51 Recognition of prior learning granted</p> <p>52 Recognition of prior learning not granted</p> <p>60 Credit transfer/national recognition</p> <p>70 Continuing enrolment</p> <p>Non-assessable</p> <p>81 Non-assessable enrolment—satisfactorily completed</p> <p>82 Non-assessable enrolment—withdrawn or not satisfactorily completed</p> <p>Other</p> <p>90 Not yet available</p>
Disability flag	<p>Indicates whether clients consider themselves to have a disability, impairment or long-term condition</p> <p>Alphanumeric 1</p> <p>Y Yes—the client has a disability, impairment or long-term condition</p> <p>N No—the client does not have a disability, impairment or long-term condition</p>
Disability type identifier	<p>Uniquely identifies the type(s) of disability, impairment or long-term condition that a client indicates</p> <p>Classification is intended to be consistent with the International Classification of</p>

Data item	Data description
	Impairments, Disabilities and Handicaps as published by the World Health Organization (Geneva, 1980) Numeric 2 11 Hearing/deaf 12 Physical 13 Intellectual 14 Learning 15 Mental illness 16 Acquired brain impairment 17 Vision 18 Medical condition 19 Other 99 Not specified

Source: NCVET 2013b.

VET in Schools Collection

The purpose of the VET in Schools Collection is to provide data on nationally recognised VET programs undertaken as part of a student's senior secondary certificate.

Reporting frequency

Annual.

Scope

The collection provides information on school-based apprenticeships and traineeships or VET subjects and courses undertaken as part of a student's senior secondary certificate that provide credit towards a recognised VET qualification.

Collection method

Administrative data collection compiled annually by the NCVER

Metadata

AVETMISS data standard:

<<http://www.ncver.edu.au/wps/portal/vetdataportal/ncversupport/supportsubsupport/NCVER-Standard-Support!/ut/p/a1/pVNNb4JAEP0tPXg>>

Terms and definitions:

Refer to 'Additional information' in:

<http://www.ncver.edu.au/wps/poc?urile=wcm:path:/wps/wcm/connect/NCVER_Shared/Collections/VET-in-schools>

Data quality statement

Data quality statements are produced for COAG Reform Council and Report on Government Services reporting.

The notes on tables accompanying every publication also contain data quality information.

Refer to the notes in the report, *Young people in education and training 2012*:

<http://www.ncver.edu.au/wps/poc?urile=wcm:path:/wps/wcm/connect/NCVER_Shared/Collections/VET-in-schools>

Data items for mapping

Data item	Data description
Country identifier	Classification is based on the Standard Australian Classification of Countries (SACC), ABS cat. no.1269.0, 2011 (second edition, revision 1) Alphanumeric 4 0000–9999 Valid 4-digit SACC code
Date of birth	The date a client was born Alphanumeric 8 DDMMYYYY Valid date @@MMYYYY Valid month and year but day not specified @@@YYYY Valid year but day and month not specified

Data item	Data description
Training organisation identifier	<p>A unique code used to identify a training organisation</p> <p>Alphanumeric 10</p> <p>text Unique alphanumeric identifier</p> <p>If the training provider is a registered training organisation, the Training organisation identifier must be the valid national code listed on the National Training Register <www.training.gov.au>.</p>
School type (parent school type)	<p>The home school of the student and not where the vocational training takes place. 'Other' includes TAFE institutes, community education providers, Australian technical colleges, and students attending mixed school types.</p> <p>Alphanumeric 2</p> <p>21 = School—Government 25 = School—Catholic 27 = School—Independent 31 = Technical and Further Education institute 61 = Community-based adult education provider</p>
Address first line	<p>The first line of an address and provides a valid street number and name (including floor and building name if relevant) or post office box information</p> <p>Alphanumeric 50</p> <p>text Street number and name (and floor and building name if relevant) or post office box information</p>
Address second line	<p>The second line of an address and provides a continuation of the valid name and number of the street (including floor and business name if relevant) or post office box information</p> <p>Alphanumeric 50</p> <p>text Continuation of street number and name</p>
Address location – suburb, locality or town	<p>The name of a suburb, locality or town of a physical geographic location</p> <p>Alphanumeric 50</p> <p>text Name of suburb, locality or town</p>
Postcode	<p>Identifies the Australia Post postcode of a physical location or a postal address</p> <p>Alphanumeric 4</p> <p>0001–9999 4-digit Australia Post postcode valid during the collection year OSPC Overseas address location 0000 Postcode unknown</p>
State identifier	<p>Uniquely identifies the state or territory of a physical location or postal address</p> <p>Numeric 2</p> <p>01 New South Wales 02 Victoria 03 Queensland 04 South Australia 05 Western Australia 06 Tasmania 07 Northern Territory 08 Australian Capital Territory 09 Other Australian territories or dependencies 99 Other (overseas but not an Australian territory or dependency)</p>

Data item	Data description
Remoteness (ARIA+) region	<p>Student remoteness (ARIA+) based on ABS ASGS SA2 regions (which are derived from the AVETMISS fields Postcode and Suburb from the Client file)</p> <p>The ASGC divides Australia into 6 Remoteness Areas and is used for collection and dissemination of geographically classified statistics. It groups locations together into comparative classes of remoteness so that data can be collected, analysed and disseminated for broad regions which are more or less remote.</p> <p>Major cities Inner regional Outer regional Remote Very remote Outside Australia Not known</p>
Client identifier	<p>Uniquely distinguishes an individual within a training organisation</p> <p>Alphanumeric 10 text Unique client identifier within the collection</p>
Indigenous status identifier	<p>Indicates a client who self-identifies as being of Australian Aboriginal or Torres Strait Islander descent</p> <p>The classification is based on the Standards for Statistics on Cultural and Language Diversity, ABS cat. no.1289.0, 1999.</p> <p>Alphanumeric 1 1 Yes, Aboriginal 2 Yes, Torres Strait Islander 3 Yes, Aboriginal AND Torres Strait Islander 4 No, Neither Aboriginal nor Torres Strait Islander</p>
Language identifier	<p>Uniquely identifies the main language other than English spoken at home by the client</p> <p>This classification is based on the Australian Standard Classification of Languages (ASCL), ABS cat. no.1267.0, 2011 (second edition, revision 1).</p> <p>If English is the only language spoken at home, the Language identifier must be '1201—English'.</p> <p>'@@@@" is used for 'Not stated' rather than the Australian Bureau of Statistics code of '0002'.</p> <p>Alphanumeric 4 0000–9999 Valid 4-digit ASCL code 0000 Unknown 0001 Non verbal</p>
ANZSCO identifier	<p>Uniquely identifies the type of occupation that may be expected for those undertaking a program of study</p> <p>The classification is based on the Australian and New Zealand Standard Classification of Occupations (ANZSCO), ABS cat. no.1220.0, 2009 (first edition, revision 1)</p> <p>Alphanumeric 6 100000–899999 Valid 6-digit ANZSCO code GEN19 Occupational non-specific—general education GEN20 Non-industry specific training NONVET Non-VET course—no occupational outcome</p>
Name for encryption	<p>Contains the client's full name in the defined format. A non-reversible encryption process encrypts the name, facilitating analysis of unit-record data while ensuring the anonymity of individual clients.</p> <p>Family name (comma) (space) first given name (space) second given name. If the name</p>

Data item	Data description
	<p>for encryption with comma and spaces exceeds 60 characters, enter client's full name in the above order and truncate at 60 characters.</p> <p>Alphanumeric 60</p> <p>text Client's full name in the specified format, e.g. family name (comma)(space) first given name (space) second given name</p>
Proficiency in spoken English	<p>The self-assessed level of ability to speak English asked of people who speak a language other than English at home</p> <p>Based on the Language Standards, ABS cat. no.1200.0.55.005, 2012 (version 1.1)</p> <p>Alphanumeric 1</p> <p>1 Very well 2 Well 3 Not well 4 Not at all</p>
Sex	<p>Identifies the gender of a client (male or female)</p> <p>Alphanumeric 1</p> <p>F Female M Male</p>
Highest school level completed identifier	<p>Identifies the highest level of school that a client has completed</p> <p>Alphanumeric 2</p> <p>02 Did not go to school 08 Year 8 or below 09 Year 9 or equivalent 10 Completed Year 10 11 Completed Year 11 12 Completed Year 12</p>
Prior educational achievement identifier	<p>Uniquely identifies the level of non-schooling sector prior educational achievement successfully completed by a client</p> <p>This classification is based on a subset of the Australian Standard Classification of Education (ASCED), ABS cat. no.1272.0, 2001.</p> <p>Numeric 3</p> <p>008 Bachelor degree or higher degree level (defined for AVETMISS use only) 410 Advanced diploma or associate degree level 420 Diploma level 511 Certificate IV 514 Certificate III 521 Certificate II 524 Certificate I 990 Miscellaneous education</p>
Subject field of education identifier	<p>Based on the field of education (FOE) at the detailed level (6-digit), which is one part of the Australian Standard Classification of Education (ASCED), ABS cat. no.1272.0, 2001</p> <p>The FOE is the subject matter of the unit of competency or module.</p> <p>Alphanumeric 6</p> <p>BBNDD Valid 6-digit detailed level FOE code</p>
Outcome identifier—national	<p>Identifies the result or outcome of a client's participation in a unit of competency or module</p> <p>Numeric</p>

Data item	Data description
	2 Assessable 20 Competency achieved/pass 30 Competency not achieved/fail 40 Withdrawn/discontinued 51 Recognition of prior learning granted 52 Recognition of prior learning not granted 60 Credit transfer/national recognition 70 Continuing enrolment Non-assessable 81 Non-assessable enrolment—satisfactorily completed 82 Non-assessable enrolment—withdrawn or not satisfactorily completed Other 90 Not yet available
Disability flag	Indicates whether clients consider themselves to have a disability, impairment or long-term condition Alphanumeric 1 Y Yes—the client has a disability, impairment or long-term condition N No—the client does not have a disability, impairment or long-term condition
Disability type identifier	Uniquely identifies the type(s) of disability, impairment or long-term condition that a client indicates Classification is intended to be consistent with the International Classification of Impairments, Disabilities and Handicaps as published by the World Health Organization (Geneva, 1980). Numeric 2 11 Hearing/deaf 12 Physical 13 Intellectual 14 Learning 15 Mental illness 16 Acquired brain impairment 17 Vision 18 Medical condition 19 Other 99 Not specified

Source: NCVET 2013c.

Appendix D: Questionnaires

Stakeholder questionnaire

Please include any supporting information as attachments if required.

Organisation

- 1. Name of organisation and program area(s), if applicable

Data collections

- 2. Which of the national data collections listed at Attachment A does your organisation/area use to inform policy or for research needs in the education and training sector?

- 3. Are there any other data collections that your organisation/area uses to inform policy and research in the education and training sector?

Include any other national or state/territory data collections.

Reporting requirements

- 4. Please provide information on any regular reporting requirements related to education and training that your organisation/area has responsibility for.

5. **Are there any specific reporting requirements that can not currently be met with existing data, either from the collections listed in Attachment A or from other collections?**

Please provide details, including reference to any specific indicators and data items.

Additional information needs

6. **Are there any identified policy issues for which there is currently inadequate information to inform, develop or evaluate policy?**

Please copy and complete the table below for all relevant policy issues.

a) Policy issue:

b) Information required:

c) Are there any data developments underway or planned to address this information gap?
Please provide details.

d) Suggested method (including data collection) to obtain this information:

e) Priority (high/medium/low):

7. **Are there any data items, not currently collected in the collections listed at Attachment A, which there is a need to collect for policy, monitoring, research or reporting purposes that are not covered above?**

Please copy and complete the table below for all relevant data items.

a) Data item:

b) Reason required:

c) Are there any data developments underway or planned to address this information gap?
Please provide details.

d) Suggested method (including data collection) for obtaining this information:

e) Priority (high/medium/low):

Data standards

8. Are there any data items for which it would be beneficial to have greater consistency across collections, and potentially develop a national data standard?

Please copy and complete the table below for all relevant data items.

a) Data item/element:

b) Data collection(s):

c) Rationale for, and benefit from, greater consistency:

d) Should consideration be given to the development of a national data standard for this data item/element:

e) Priority (high/medium/low):

Data linkage

9. Are there any identified data linkage projects using one or more of the data collections listed at Attachment A, that can not currently be undertaken due to data limitations?

Please copy and complete the table below for all relevant projects.

a) Project description:

b) Datasets for linkage (and year of data):

c) Limitation (include specific data items if relevant):

d) Priority (high/medium/low):

Data custodian questionnaire

Please include any supporting information as attachments if required.

Data collection

10. Name of data collection

11. Frequency of collection and year commenced (e.g. annual, 4-yearly)

12. Please provide some brief background on the data collection, including purpose, aims, scope and policy intent.

13. Please list the metadata that is available for this collection, e.g. data dictionary, data standards, data item definitions, data collection manual.

Please include hyperlinks if this information is available online.

14. Is there a data quality statement or quality declaration available for this data collection and/or individual data elements, or are there plans to develop these in the future?

Please provide details and include hyperlinks or attachments as required.

Reporting outputs

15. Please list the regular reporting outputs from this data collection, and provide a summary of any ad-hoc reporting outputs.

Please include hyperlinks to reporting outputs available online.

Data collection, reporting and review cycle

16. Please provide details and timeframes on the *data collection and reporting cycle* for this collection.

Note: Include task, month and time, e.g. Jurisdictions provide data (August, 4 weeks).

17. How often do you review changes for this data collection, e.g. review data items and data standards, and when will this process next occur?

18. Please provide details and timeframes on the process to review and implement changes for this collection.

Note: Include task, month and time, e.g. Review of data standards (Jan to Mar, 12 weeks); Endorsement from parent committee (August, 4 weeks).

Data improvements

19. In relation to the purpose, aims, scope and policy intent of the data collection (refer to Question 3), are there any restrictions (legislative or otherwise) on changes to data standards or the inclusion of additional data items in this collection?

Please provide details.

20. Please describe the process for making changes to this collection (such as adopting a new data standard or adding a new data item).

Note: Include details on governance structures that must approve/endorse changes (e.g. committees, subgroups, overseeing body) and data systems upgrades.

21. Please describe any recent (approximately the last 5 years) changes to this collection to improve data standards and the quality of the collection.

Please describe the change and rationale/benefit.

22. Are there currently any existing projects/plans for improvements to this data collection?

If yes → please provide details below

If no → go to question 13

Please provide details, including timeframes if known.

23. What is the rationale for, and expected benefits from, these data improvements?

For example, will it allow the data to be used for additional reporting purposes, for data linkage or to address identified policy issues or knowledge gaps?

24. What revisions would be required to data items and reporting requirements in order to improve the ability of data from this collection to *inform identified policy issues*?

Please copy and complete the table below for all relevant policy issues.

- a) Policy issue:
- b) Revisions to data items:
- c) Revisions to reporting requirements:

25. Are there any data items for which it would be beneficial to have greater consistency across collections, and potentially develop a national data standard?

Please copy and complete the table below for all relevant data items.

- a) Data item/element:
- b) Data collection(s):
- c) Rationale for, and benefit from, greater consistency:
- d) Should consideration be given to the development of a national data standard for this data item/element:
- e) Priority (high/medium/low):

26. Are there any other identified needs for improvements to this data collection that are not covered above? E.g. any other revisions to improve the data collection and reporting methods?

If yes → please copy and complete the table below for all identified changes.

If no → go to question 17

- a) Description of change required:
- b) Data elements involved:
- c) Expected benefits:
- d) Priority (high/medium/low):

Data linkage

27. Please provide details on any data linkage/integration projects using this data collection, either current, recently completed or expected.

Please copy and complete the table below for all relevant projects.

a) Project title:

b) Datasets for linkage (and years of data):

c) Project timeframe:

d) Status (completed/current/expected):

28. Please provide details on the data elements from this collection that have been, or could be used, to link or integrate with other data collections, e.g. name, address.

Please provide the name of the data element(s), data standard and the name of any data collections that have been linked on these data elements.

29. For current and expected projects, is any data development required before data linkage or integration can occur? Please provide details.

Appendix E: Stakeholder and data custodian consultation

This appendix summarises and discusses the results of the stakeholder and data custodian consultation undertaken in Phase 2 of the project via questionnaire. Specifically, it summarises the aspects related to the identified needs for consistency in data standards at the data item level and at the collection level.

Stakeholders and data custodians noted that while national consistency for core data items may be desirable in many cases, it is not always appropriate or needed – it depends on the purpose of the collection. There should therefore be a clear rationale for adopting a data standard to be consistent with an existing national data standard or with the standard of another collection, in terms of policy priority, benefit and cost.

Data item level

Indigenous status

Stakeholders and data custodians noted that there was a need for consistency in data item standards for core equity groups such as Indigenous students, particularly for schools, VET, and higher education collections. The data mapping indicated a fairly high level of consistency (fully consistent or mappable) to the national data standard for this data item across the collections, but further consideration could be given to ensuring that the response categories are fully consistent with the national standards, rather than mappable.

Remoteness

Disaggregation of student attendance data by remoteness was identified as a reporting requirement that cannot currently be met, but would help researchers and policymakers formulate and target strategies.

Currently, there is inconsistency in the remoteness information available across the collections (refer to Chapter 2). For some collections, it is collected based on the usual residence of the child/student and for others it is based on the school. The school-based collections use the MCEECDYA Schools Geographic Location Classification rather than the ABS remoteness classification. ACARA is investigating this issue.

For some data collections, information for remoteness could be derived based on available address information. It would be possible to add a derived remoteness variable to these collections to allow disaggregation and comparison with output from other collections.

It was also noted that consistency in standards between VET Provider, VET in Schools and Higher Education collections would be useful.

Socioeconomic status (SES)

Stakeholders noted that funding is increasingly using SES indicators, which increases the need for consistency across jurisdictions and sectors. Disaggregation of student enrolment and student attendance data by SES was identified as a reporting requirement that cannot currently be met, and this information would enable this to be reported and would help researchers and policymakers formulate and target strategies.

There are currently a number of different measures of SES used in education and training collections, and the former SCDC did work to assess the coherence of measures of SES within and across education and training sectors. This work highlighted the difficulties in developing and applying consistent SES measures across the education and training sectors, and suggested that, in the short term, work in this area should focus on improving the quality and consistency of information within sectors (noting that cross-sectoral consistency would still be a desirable outcome).

Stakeholders specifically noted that consistency in standards between VET Provider, VET in Schools and Higher Education collections would be useful.

For some data collections, area-based SES information could be derived based on available address information, according to multiple indexes of SES. It would be possible to add a derived SES variable to these collections to allow disaggregation and comparison with output from other collections.

Disability status

The need for the collection of consistent information across the data collections on children and students with disability or special needs was identified by multiple stakeholders. This information could potentially address a number of policy and research needs in relation to the participation of children/students with disability in education and training.

Some national standards exist in relation to disability (refer to Chapter 2), and the AIHW has developed a standardised disability flag module that is currently used in the community services sector, that is available in METeOR, AIHW's Metadata Online Registry (AIHW 2013). Elements of this module could be considered for suitability. However, whether these national standards are appropriate to use across all education and training sectors and for children/students at different developmental stages needs further consideration. Whether this information could also be reported consistently might depend on who is making the assessment (for example, teacher, parent, self-reported and/or whether it is based on medical diagnosis).

Further work would need to be done to determine the appropriateness and feasibility of the development of consistent standards for disability in education and training data collections, and whether a single standard would be possible across the sectors, or whether linkage with other collections would be viable to obtain this information.

Work to develop and collect nationally consistent data on school students with disability is being done by the Students with Disability Working Group and the Department of Education and Training. The National Student with Disability Collection collects standardised data; however, it is separate from the NSSC with no consistency of recording data across jurisdictions.

English proficiency

Standardisation of data items related to English proficiency nationally – for example, language background other than English, language other than English, English as an additional language or dialect and English as a second language (ESL) – was identified as important for student performance data and to determine progress of these students in education and training.

Four of the education and training data collections in scope of the national data standards strategy collect information on proficiency in English in some form.

Migrant status, Australian citizenship status and year of arrival in Australia

A number of data items related to migrant and citizenship status were identified by stakeholders as important for improving policy development and program monitoring and reporting, particularly in relation to VET and higher education. These items were generally identified as medium priority for collection.

Year of arrival in Australia was the only data item that was identified as in scope for the data mapping exercise, and was collected by the Census and in the Higher Education Statistics Collection only (both were fully consistent with the national standard).

Visa/migrant status was not within scope of the data items to be mapped, and due to ongoing changes to these categories it was considered to be difficult to obtain consistent information between collections and over time.

Other data items

Other data items identified by stakeholders for collection and standardisation that were deemed to be relevant to the core data item categories were family composition, sibling order, refugee status, prior education, child protection status and health care card status. These items were only identified once by stakeholders.

It was also noted that there is an increasing demand for reporting on vulnerable children, but no definition is in place for measurement of this concept.

Collection level

VET and higher education

A number of stakeholders commented on the benefits of greater consistency between the VET Provider, VET in Schools and Higher Education data collections, in relation to specific data items and the collections more broadly.

The VET collections operate under the same standard (AVETMISS), which evolves in response to agreed policy priorities. The Higher Education collection uses a different standard and is not always consistent with ABS standards.

There are a number of providers operating in both VET and higher education that have to collect and report their data differently for essentially the same items, and improved consistency would ultimately help these dual-sector providers, although cost and timing are barriers.

National Schools Attendance Collection

Nationally comparable data are not currently available from this collection. However, ACARA has developed the National Standards for Student Attendance Data Reporting. These were endorsed by SCSEEC and will be implemented for reporting in 2015.

National Schools Statistics Collection

The value of consistent definitions and standards across government and non-government schools was noted, in order to provide comparable data for analysis.

NAPLAN and National Schools Statistics Collection

It was noted that currently ACARA and the ABS have inconsistent response values for a number of data items, which creates issues for data providers to meet national reporting requirements in relation to both collections.

Appendix F: Summary of policy priorities

Table F1: Summary of current policy priorities with relevance to education and training as at October 2014

Policy document	Summary of priorities	Source
Early childhood		
National Partnership Agreement for Universal Access to Early Childhood Education	Maintaining Universal Access to quality early childhood education program(s) for all children in the year before full-time school for 600 hours per year, delivered by a degree qualified early childhood teacher who meets the National Quality Framework requirements with a focus on participation by vulnerable and disadvantaged children. Includes the following outcomes: <ul style="list-style-type: none"> vulnerable and disadvantaged children have access to and participate in an affordable, quality early childhood education program Indigenous children have access to and participate in an affordable, quality early childhood education program the achievement of the Closing the Gap target is maintained, to ensure access to early childhood education for all Indigenous 4-year-olds in remote communities 	COAG 2013b
National Partnership Agreement on the National Quality Agenda for Early Childhood Education and Care	Jointly governed unified National Quality Framework for early childhood education and care and outside school hours care to improve outcomes for all children especially for those children from disadvantaged or at risk backgrounds Build a highly skilled workforce National Quality Standard provides clarity about expectations for quality early childhood education and care National quality rating system to provide accountability	COAG 2009d
Investing in the early years—a national early childhood development strategy	Enhance collaboration across schools and early childhood services to improve transition to school from early childhood settings Consider AEDC results in community, state and national planning processes	COAG 2009a
Indigenous		
National Indigenous Reform Agreement (NIRA) (Closing the Gap)	Halve the gap for Indigenous children in reading, writing and numeracy within a decade (by 2018) Halve the gap for Indigenous people aged 20–24 in Year 12 or equivalent attainment rates by 2020 Halve the gap in employment outcomes between Indigenous and non-Indigenous Australians within a decade (by 2018) Close the gap between Indigenous and non-Indigenous school attendance within 5 years	Australian Government 2014; COAG 2012e
Aboriginal and Torres Strait Islander Education Action Plan	Reinforces the targets and outcomes of the NIRA (above) along with additional objectives: <ul style="list-style-type: none"> increase in the number of Aboriginal and Torres Strait Islander teachers, principals and education workers Aboriginal and Torres Strait Islander students make a successful transition from school to work and further study 	MCEECDYA 2011
Health and disability		
National Partnership Agreement on Preventive Health	Reforms Australia's efforts in preventing the lifestyle risks that cause chronic disease. Several initiatives, specifically Healthy Communities and Healthy Children	Australian Government Department of Health 2013; COAG 2009b

(continued)

Table F1 (continued): Summary of current policy priorities with relevance to education and training

Policy document	Summary of priorities	Source
Health and disability (continued)		
National Disability Agreement	Build the evidence base for disability policies and strategies Develop employment opportunities for people with disability with a target of a 5 percentage point increase in the proportion of people with disability participating in the labour force between 2009 and 2018	Australian Government Department of Social Services 2013; COAG 2012c
National Partnership Agreement for More Support for Students with Disabilities	Funding to support schools and teachers to provide services for students with disability and learning difficulties including services by health or allied health specialists in schools, curriculum differentiation, and assistive technology in the classroom	COAG 2014
National Disability Strategy	The learning and skills policy area in the strategy has the greatest relevance to education and training, and includes early childhood education and care, schools, further education, vocational education, transitions from education to employment and lifelong learning. The commitments and areas for future action that are most relevant to an education and training data standards strategy are the nationally consistent collection of data on school students with disability and the reporting of educational outcomes for students with disability. The importance of consistent data on disability is emphasised in the strategy. Draft trend indicators for learning and skills include: <ul style="list-style-type: none"> • proportion of people with disability in mainstream schools • proportion of people aged 19–25 with disability who have attained at least Year 12 or equivalent qualification • proportion of people with disability with post-school qualifications. 	COAG 2011; SCSEEC 2014
Intergovernmental Agreement for the National Disability Insurance Scheme (NDIS) launch	Provides the foundation for governments to work together to develop and implement the first stage of the NDIS to ensure that people with disability and their carers receive access to the required support. The full scheme will be rolled out progressively across the states and territories from 2016. The NDIS is dedicated to the vision of a community that values people with disability.	COAG 2012a
Education and training		
Melbourne Declaration on Educational Goals for Young Australians	The goals are for Australian schooling to promote equity and excellence, and that all young Australians become successful learners, confident and creative individuals, and active and informed citizens. Ensure that socioeconomic disadvantage ceases to be a significant determinant of educational outcomes Reduce the effect of other sources of disadvantage, such as disability, homelessness, refugee status and remoteness	MCEETYA 2008

(continued)

Table F1 (continued): Summary of current policy priorities with relevance to education and training

Policy document	Summary of priorities	Source
Education and training (continued)		
National Education Agreement	<p>Objective of the agreement is that all Australian school students acquire the knowledge and skills to participate effectively in society and employment in a globalised economy</p> <p>Contributes to 5 outcomes:</p> <ul style="list-style-type: none"> • all children are engaged in and benefiting from schooling • meeting and improving basic literacy and numeracy standards • schooling promotes the social inclusion and reduces the education disadvantage of children, especially Indigenous children • Australian students excel by international standards • young people make a successful transition from school to work and further study. <p>Includes 4 targets critical to achieving the outcomes and objective:</p> <ul style="list-style-type: none"> • lift the Year 12 or equivalent or Certificate II attainment rate to 90% by 2015 • lift the Year 12 or equivalent or Certificate III attainment rate to 90% by 2020 • halve the gap for Indigenous students in reading, writing and numeracy by 2018 • at least halve the gap for Indigenous students in Year 12 or equivalent attainment rates by 2020. 	COAG 2012d
National Education Reform Agreement	<p>Builds on the Melbourne Declaration on Educational Goals for Young Australians (2008) and the NEA (2008) and provides the basis for working towards the national goals set out in the National Plan for School Improvement</p> <p>Incorporates existing COAG targets from the NEA as well as new targets to achieve the outcomes of the NEA and the objective that Australian schooling provides a high quality and equitable education for all students</p> <p>Includes a commitment to social inclusiveness and addressing disadvantage</p>	COAG 2013a
Australian Qualifications Framework (AQF)	Encourages lifelong learning	Australian Qualifications Framework Council 2013
Higher Education Participation Programme	<p>Ensuring that Australians from disadvantaged backgrounds who have the ability to study at university get the opportunity to do so and succeed</p> <p>Funds universities to increase participation of domestic students from low socioeconomic status backgrounds in undergraduate qualifications and support the retention and success of those students</p> <p>Funds activities in partnership between universities and primary and secondary schools, VET providers and state and territory governments to raise aspiration and build capacity of people from low socioeconomic backgrounds to participate in higher education</p>	Australian Government Department of Education 2014b

(continued)

Table F1 (continued): Summary of current policy priorities with relevance to education and training

Policy document	Summary of priorities	Source
Workforce		
National Agreement for Skills and Workforce Development	<p>The objective of this agreement is to achieve a VET system that delivers a more productive and highly skilled workforce, enabling all working-age Australians to participate effectively in the labour market and contribute to Australia's economic future. Measures include:</p> <ul style="list-style-type: none"> • proportion of working-age population with higher level qualifications (Certificate III and above) • proportion of working-age population with or working towards a non-school AQF qualification. 	COAG 2012b
National Partnership Agreement on School Pathways Programs	To increase the pool of young people with the prerequisite skills and entry capabilities to move from school into further education, apprenticeships, internships, scholarships and part-time work/study combinations in the occupational areas that supports the maintenance and expansion of the defence industry sector. Applies to South Australia and Western Australia only.	COAG 2009c
Industry Skills Fund	<p>Complements Trade Support Loan Scheme to offer loans up to \$20,000 over the life of an apprenticeship to increase apprenticeship completion rates</p> <p>Deliver 200,000 targeted training places and training support services over 4 years</p>	Macfarlane 2014
National Early Childhood Education and Care workforce initiatives	Initiatives including extra university places for early childhood teachers, the removal of TAFE fees for child care students and paying 50% of the HECS-HELP fees of early childhood teachers who work in areas of high disadvantage	Australian Government Department of Education 2014a

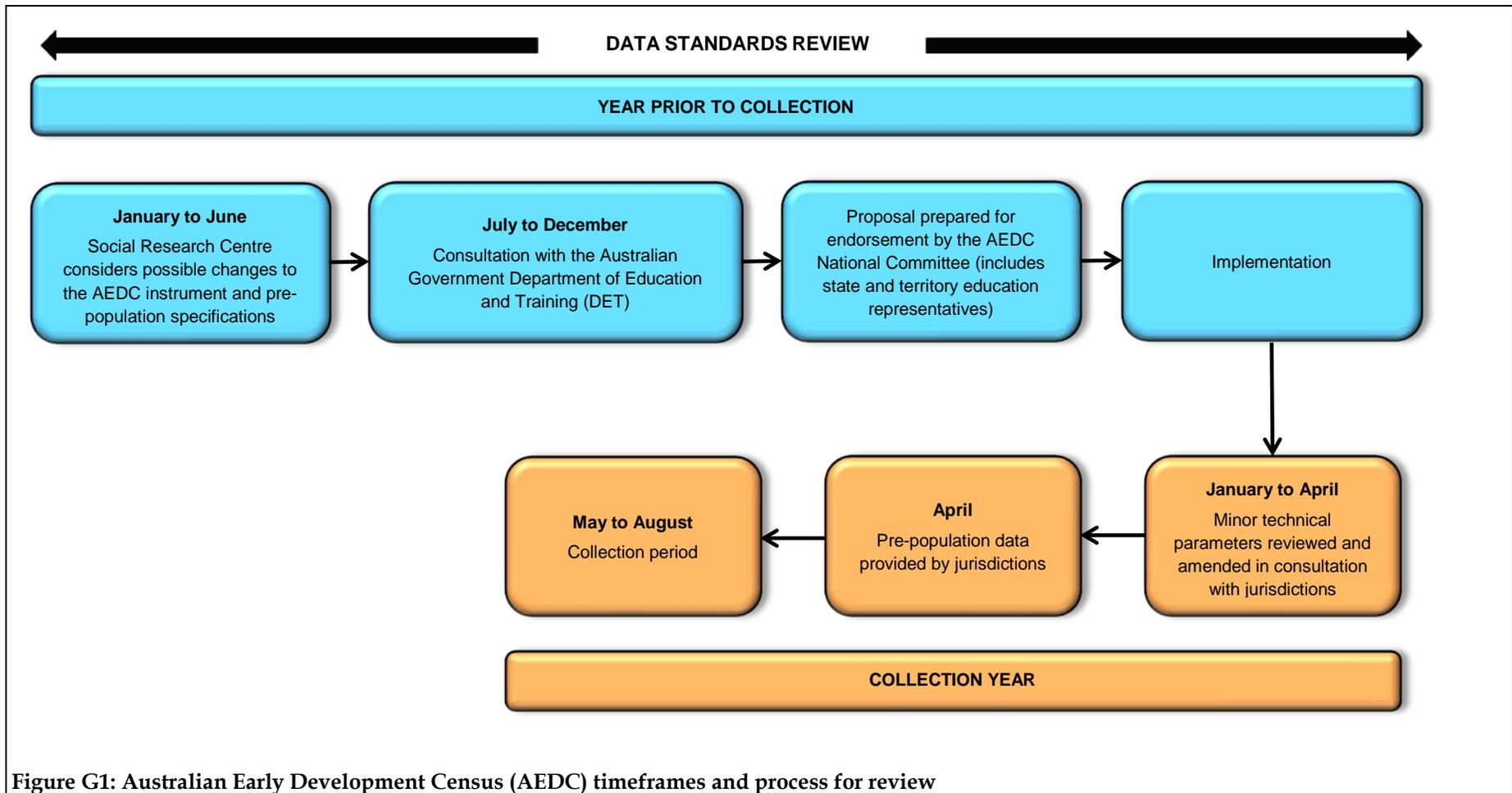
Appendix G: Data collection review timeframes and processes

Information was obtained from data custodians regarding data collection cycles, including the frequency with which data items and standards are reviewed, and the process for implementing changes to data standards.

This information can be used to identify opportunities for changes to be made in relation to data standards for these collections based on existing data collection cycles/data systems upgrades.

The information on processes and timeframes in figures G1 to G12 is based on the most recent cycle of the data collection and is subject to change for future cycles.

The information that is available for each data collection varies and is dependent on the input from data custodians.



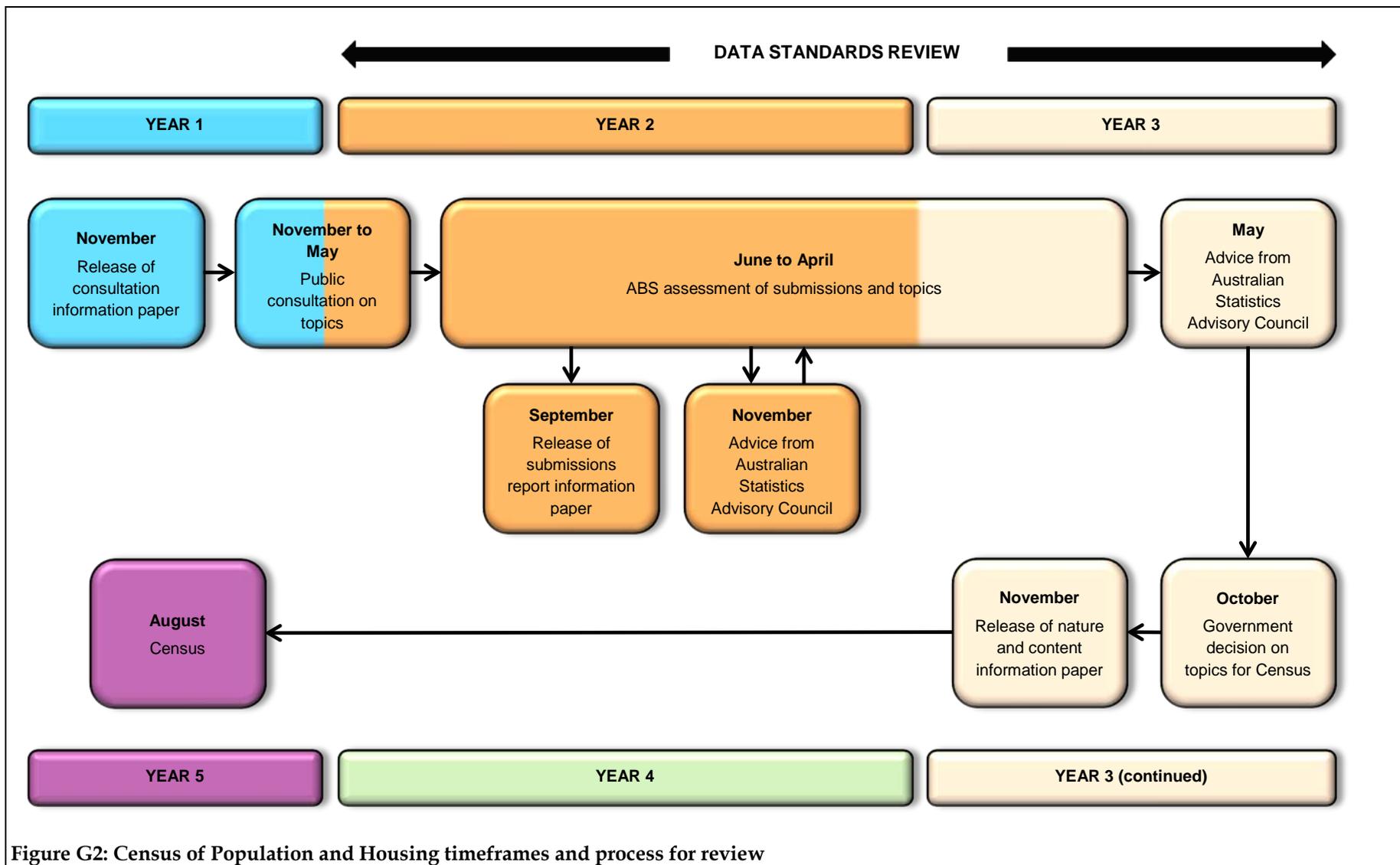
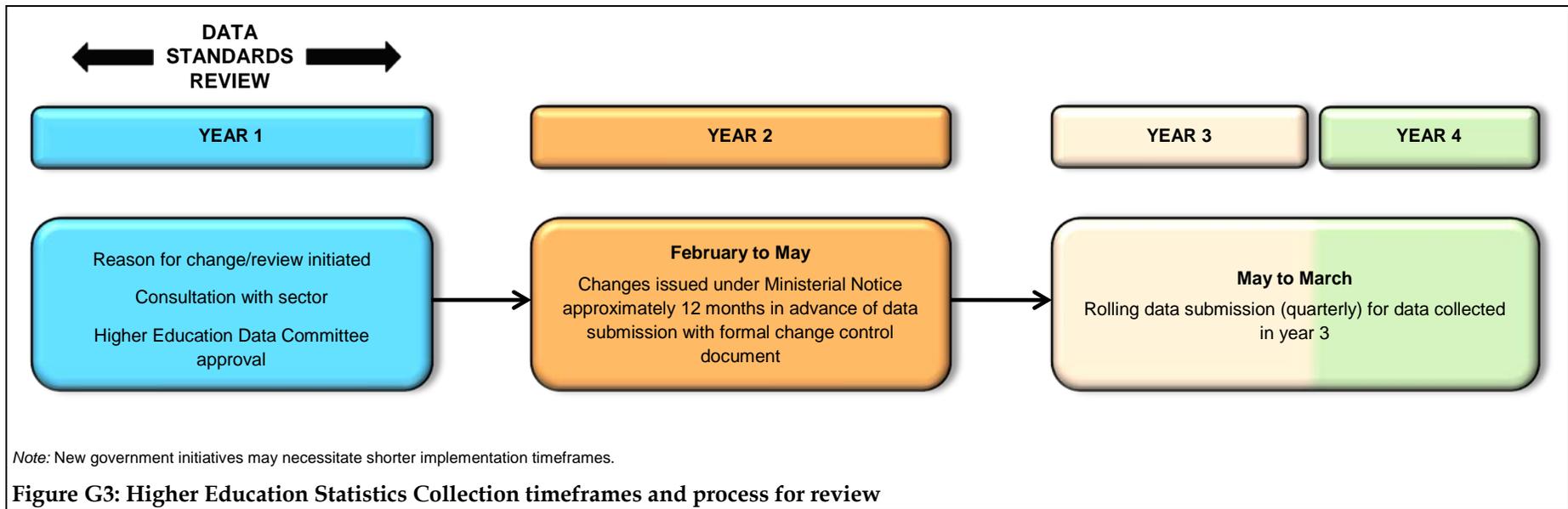
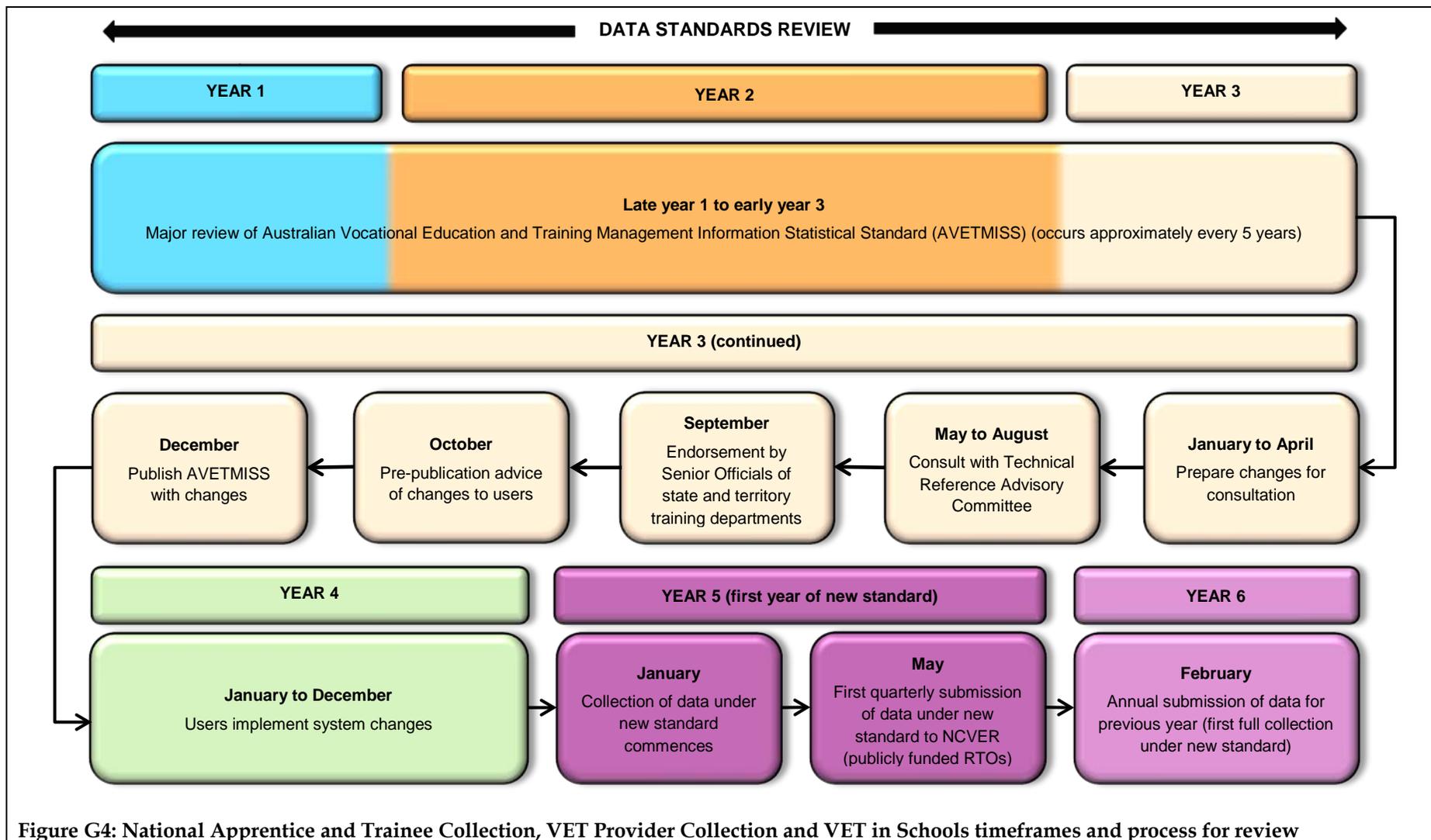
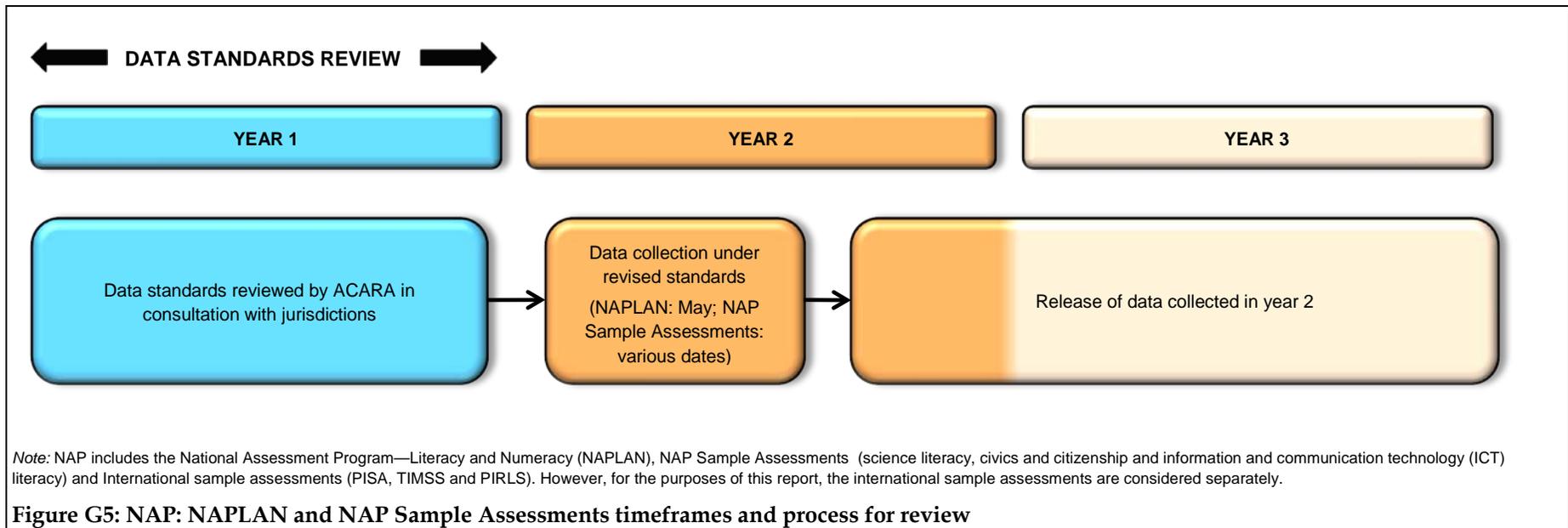


Figure G2: Census of Population and Housing timeframes and process for review







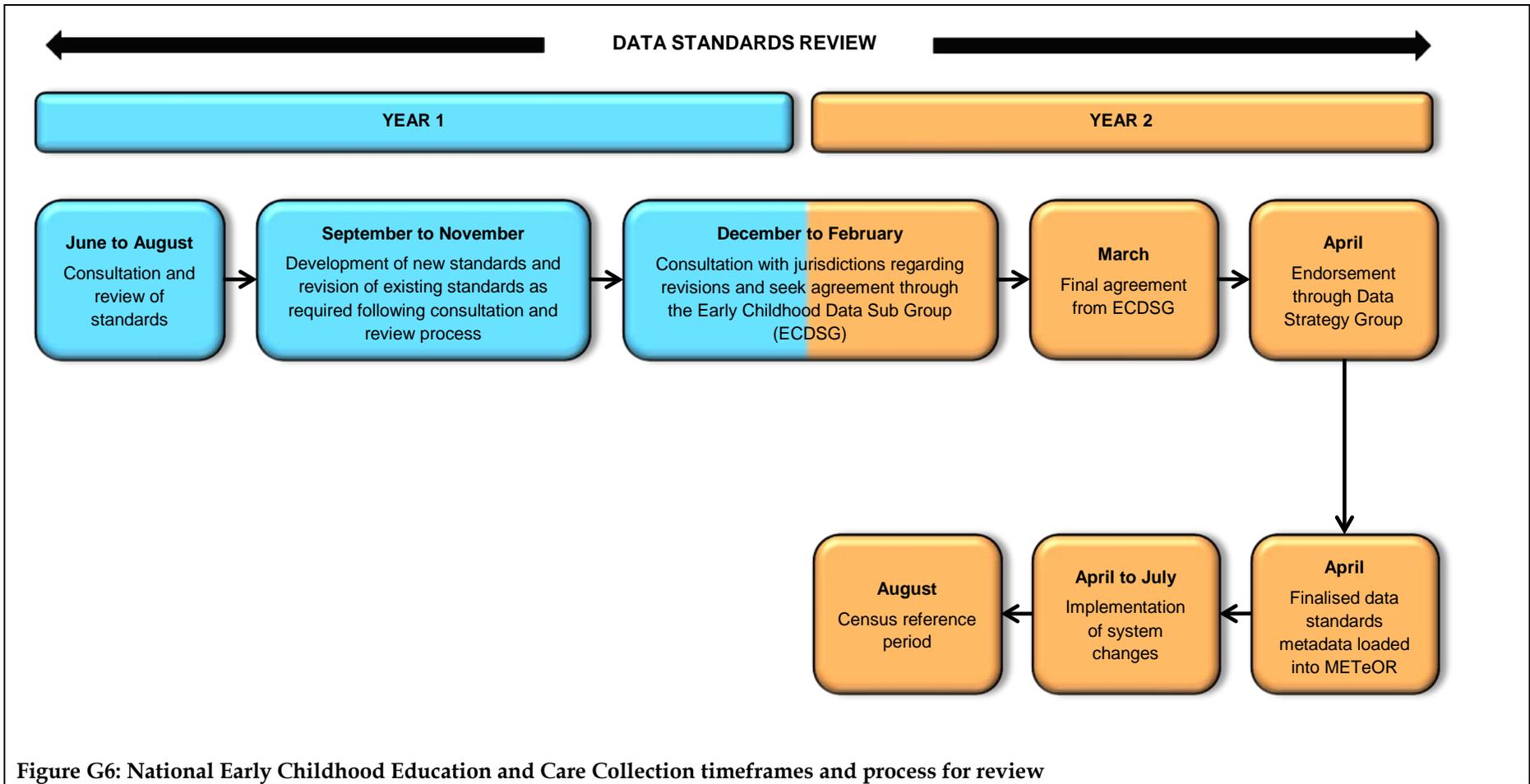
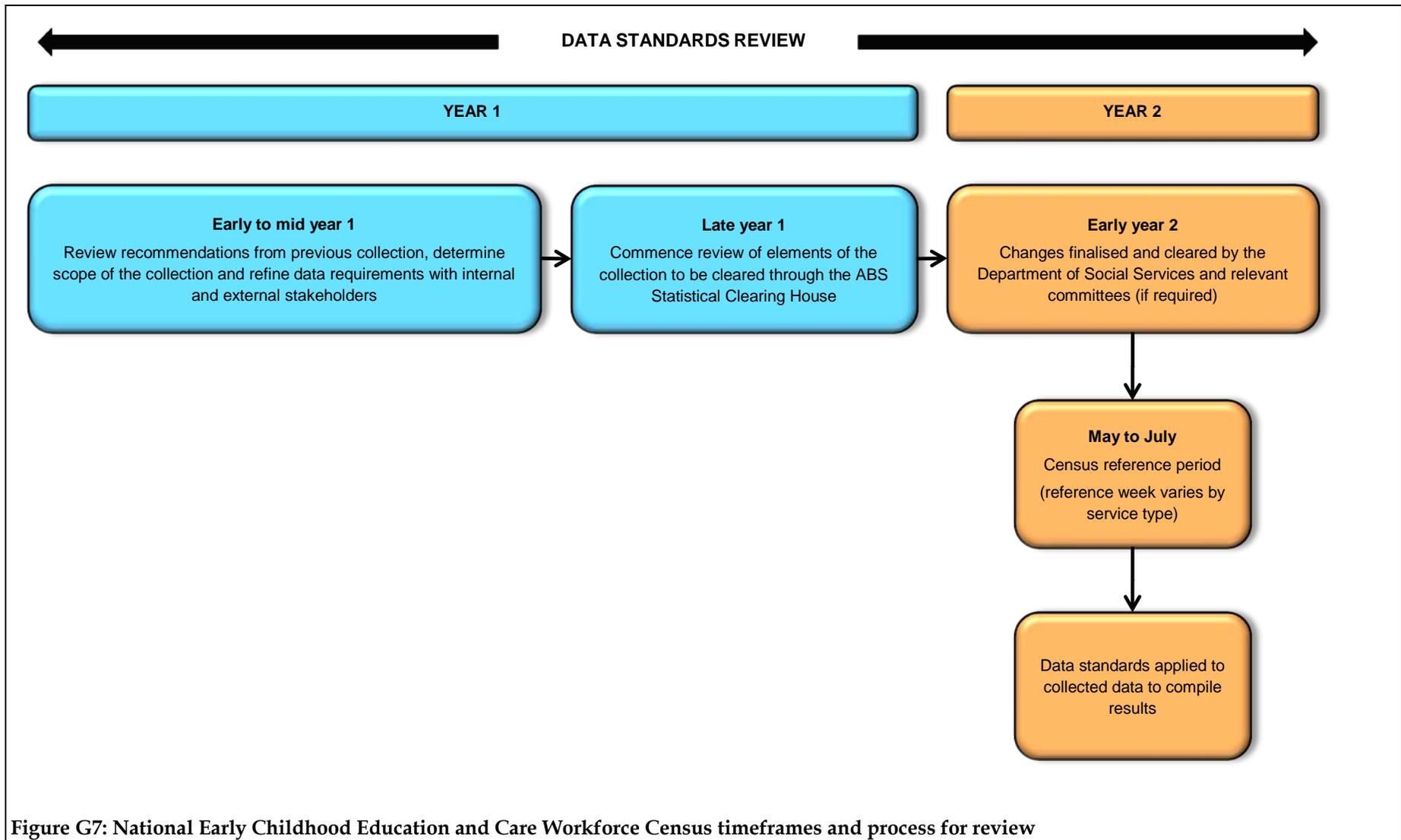
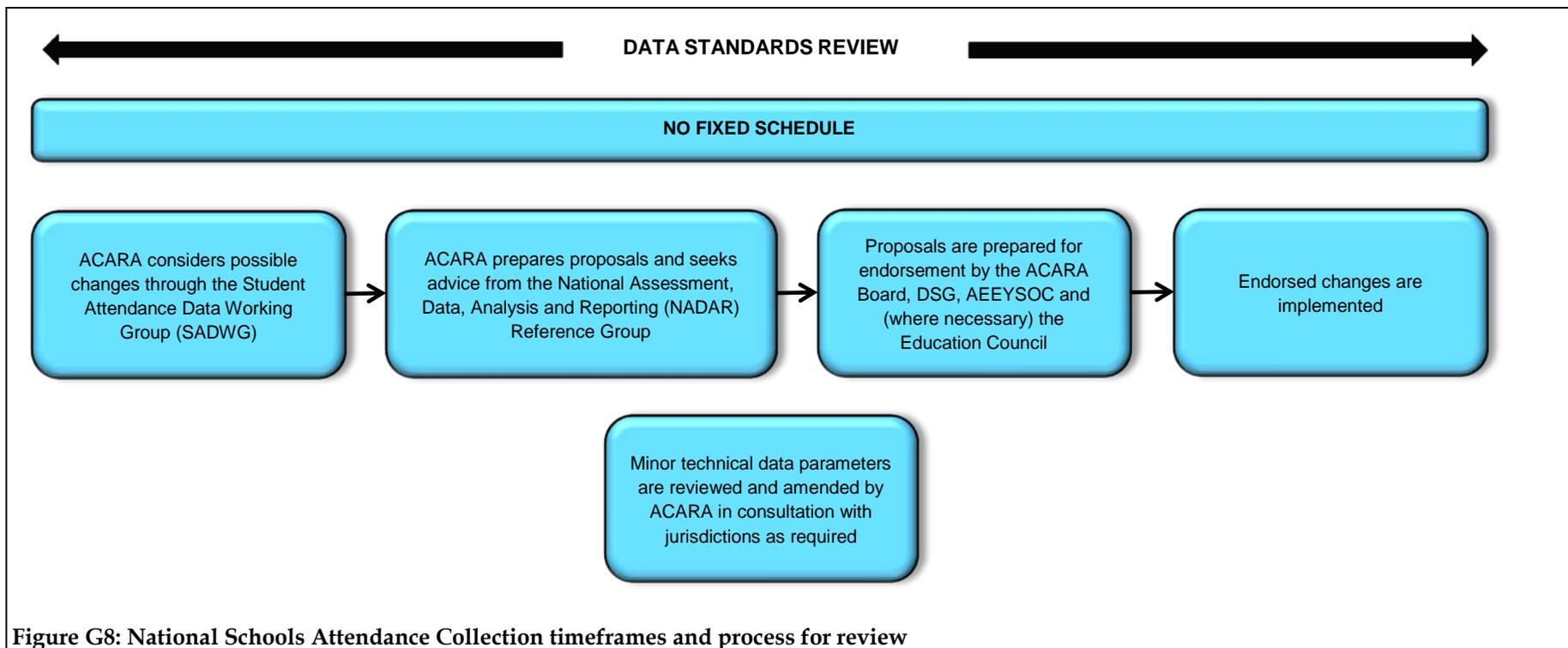
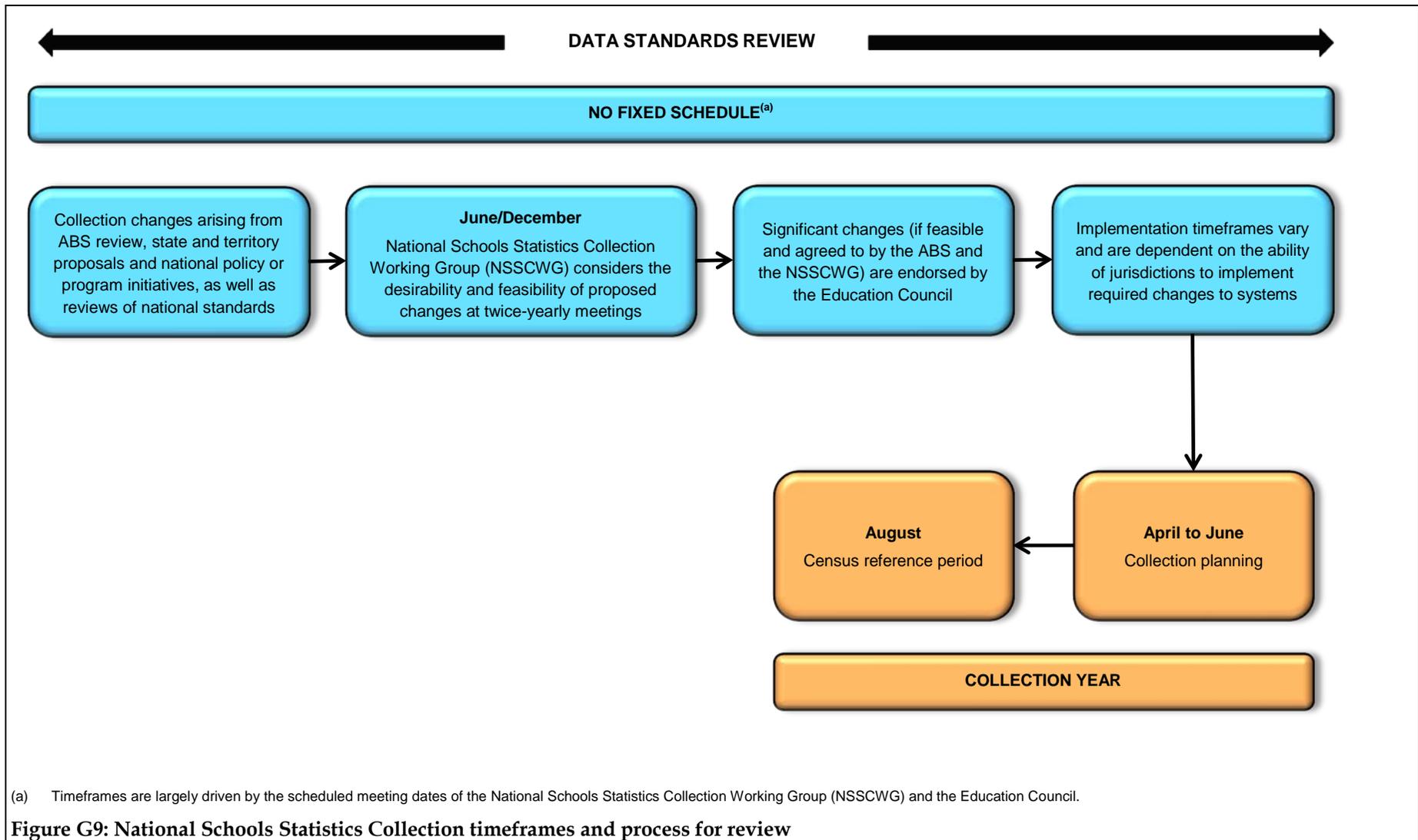
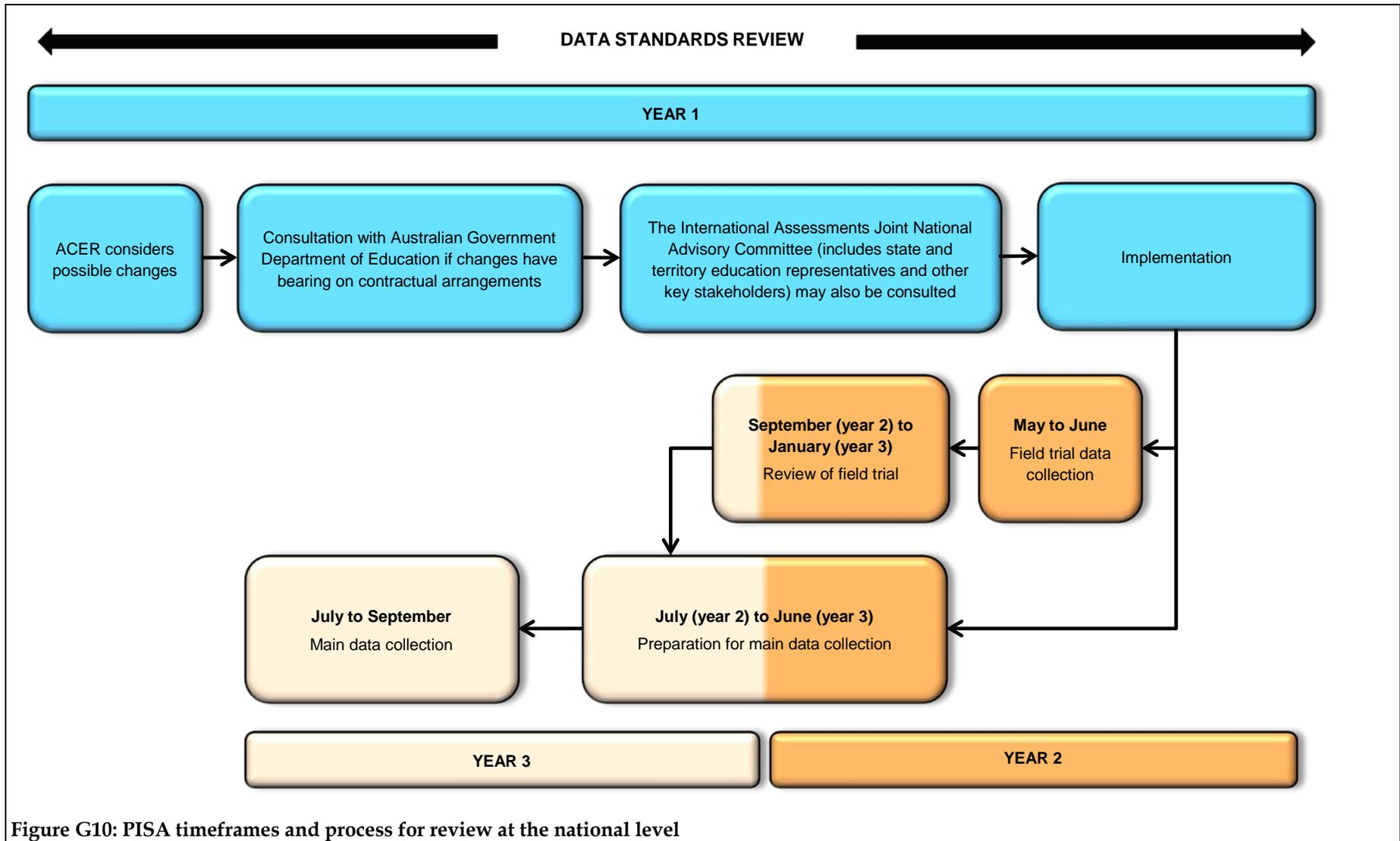


Figure G6: National Early Childhood Education and Care Collection timeframes and process for review









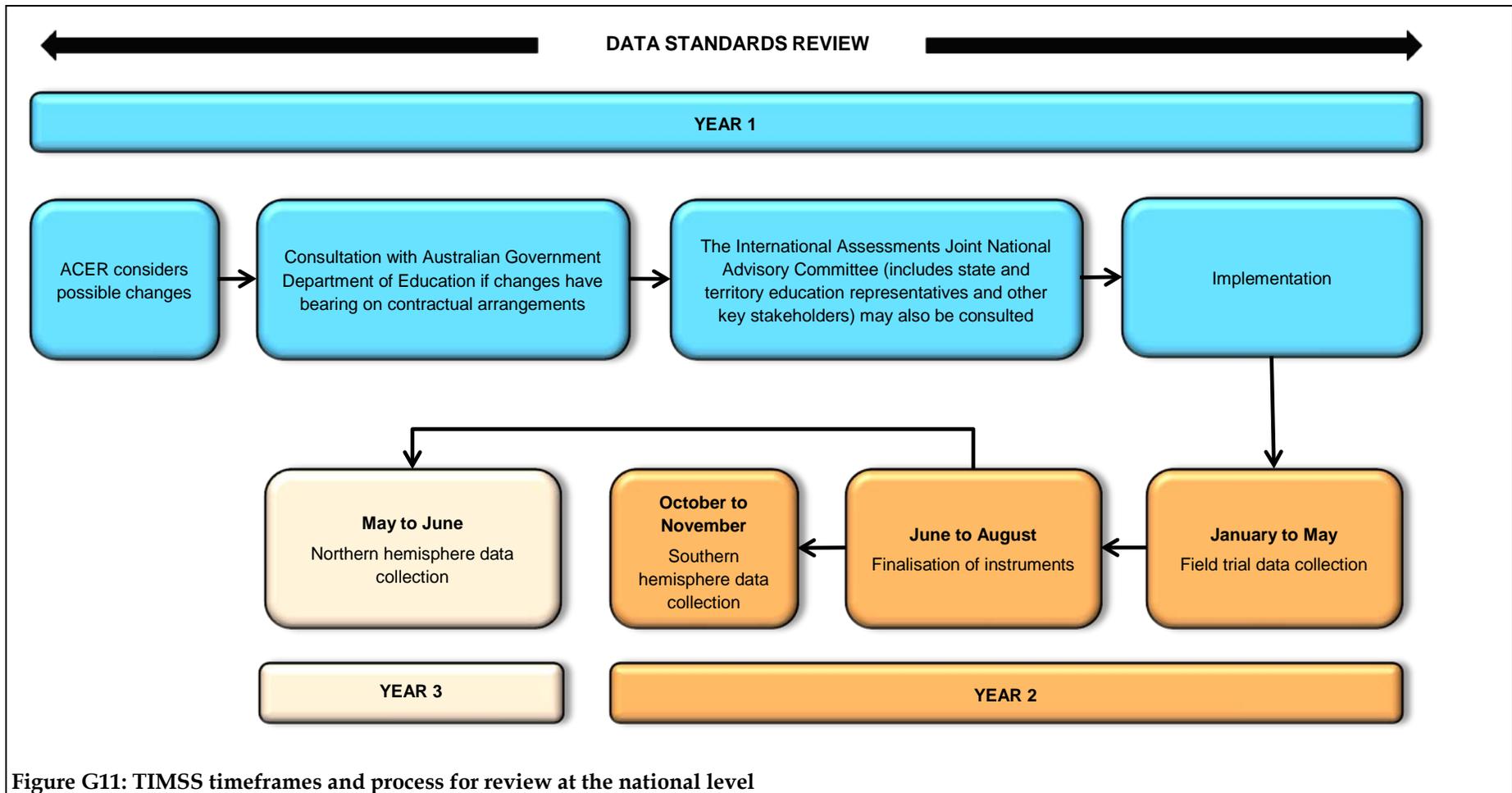
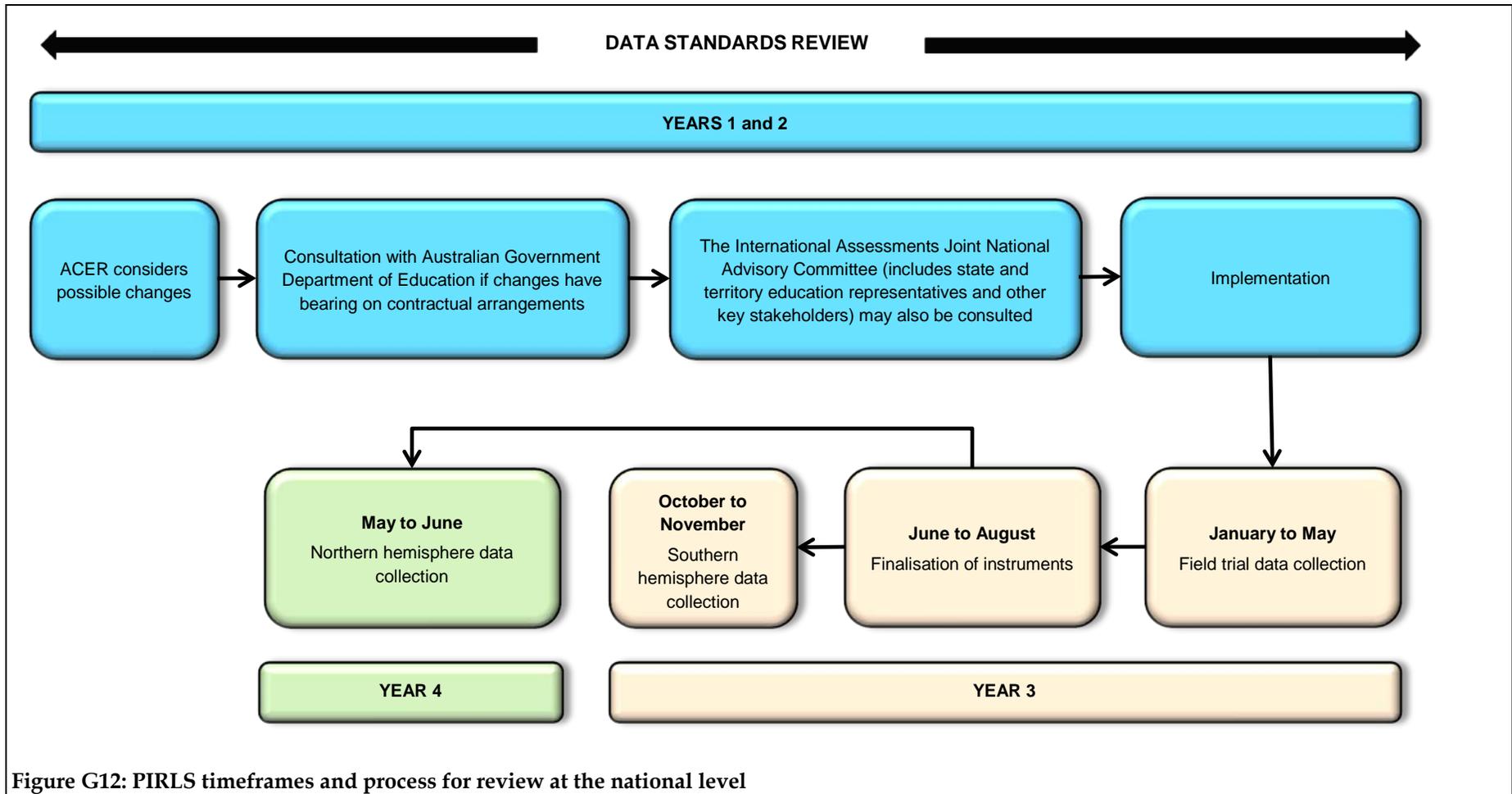


Figure G11: TIMSS timeframes and process for review at the national level



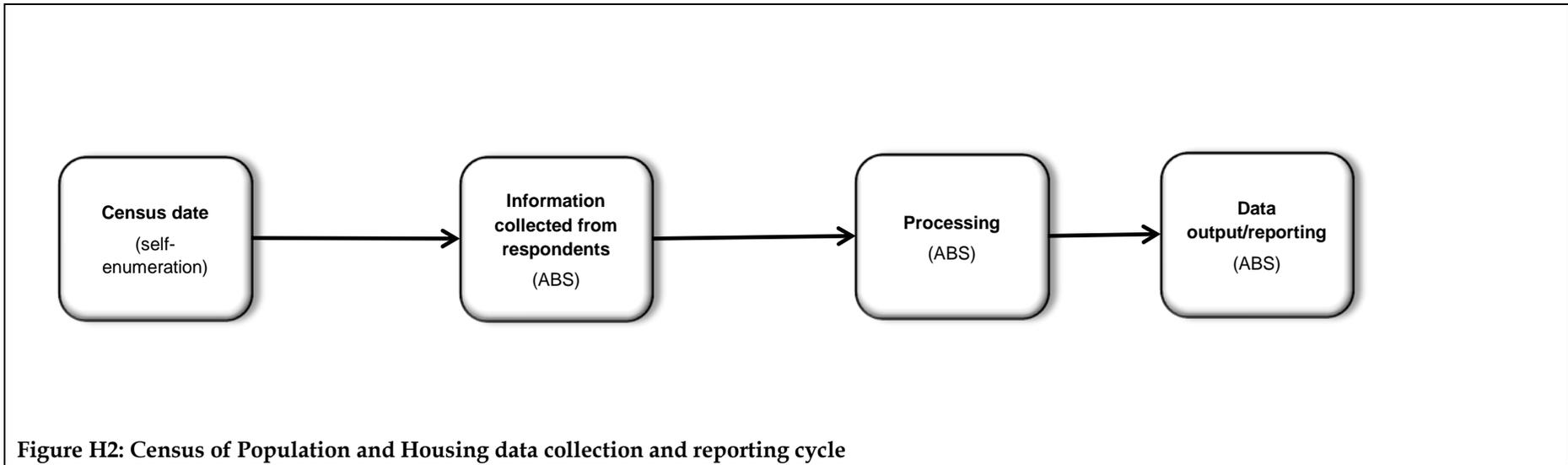
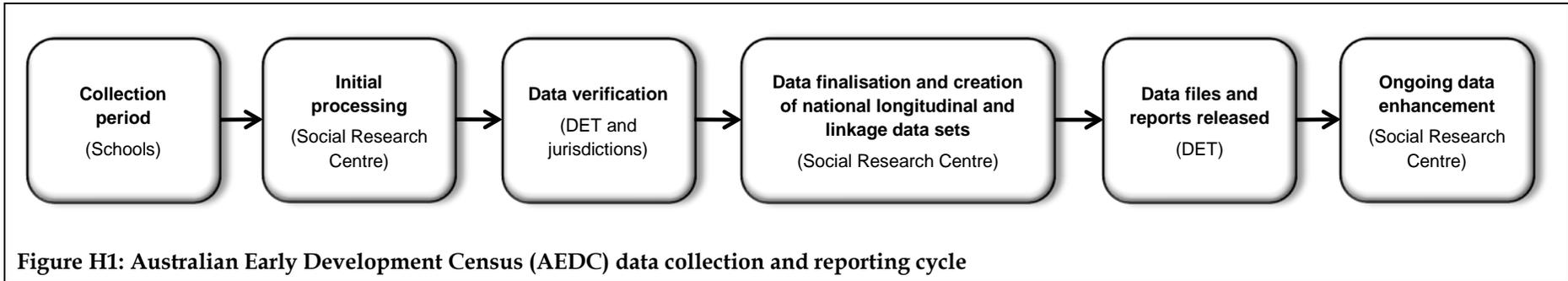
Appendix H: Data flows

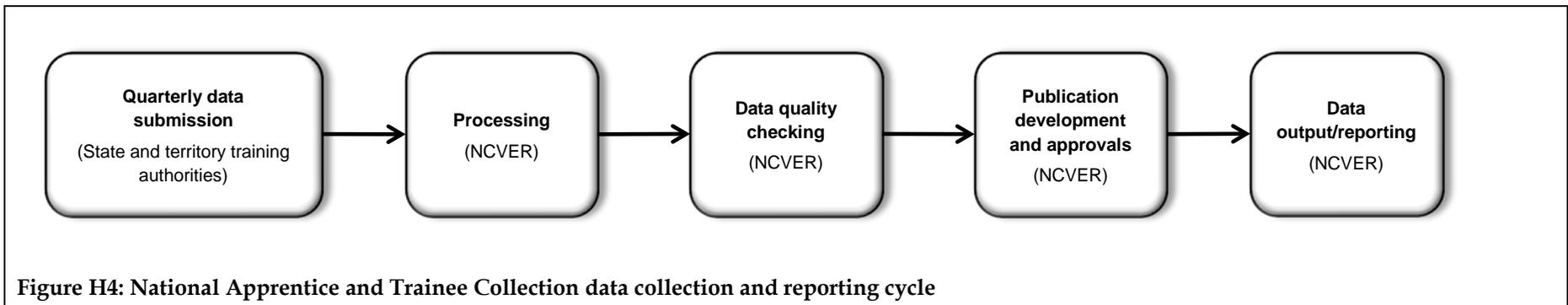
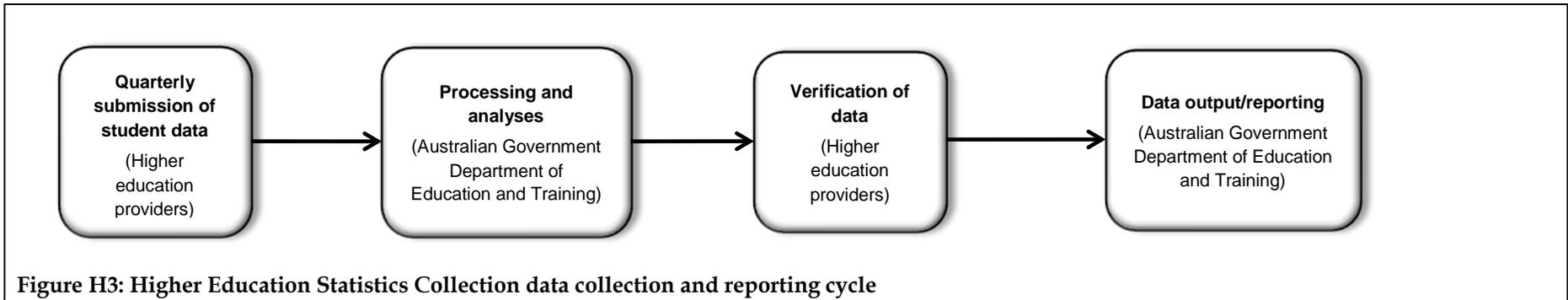
Information was obtained from data custodians regarding the data collection and reporting cycles for the education and training data collections.

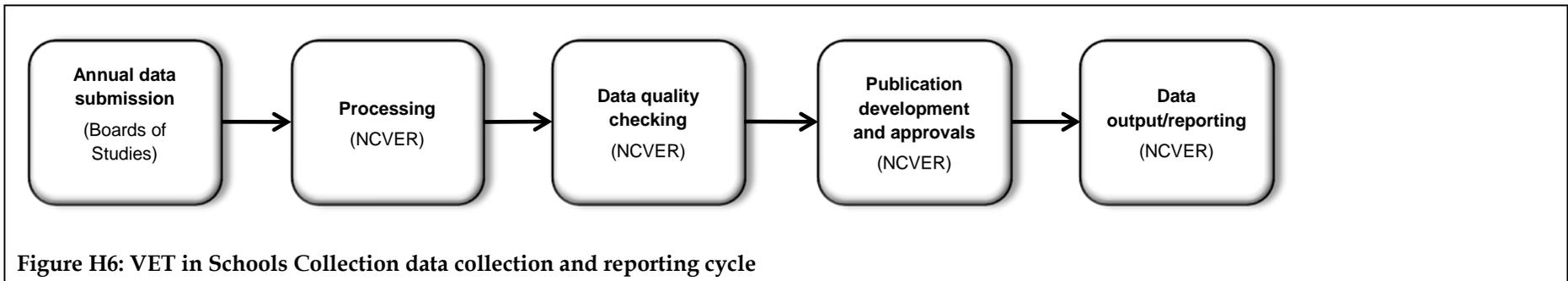
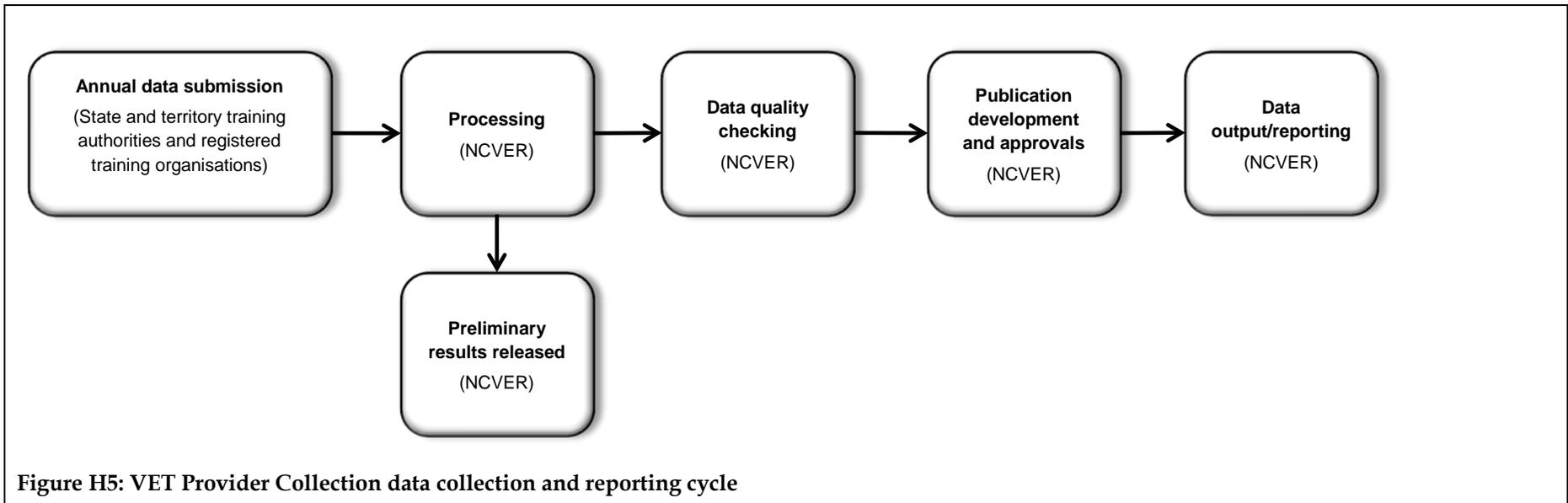
This information can be used to show the flow of data from the collection stage through to reporting outputs and highlights the stakeholders involved in this process. A common flow of data includes: data collection, data submission, data processing, data verification, analysis, data output/reporting development, approvals and release.

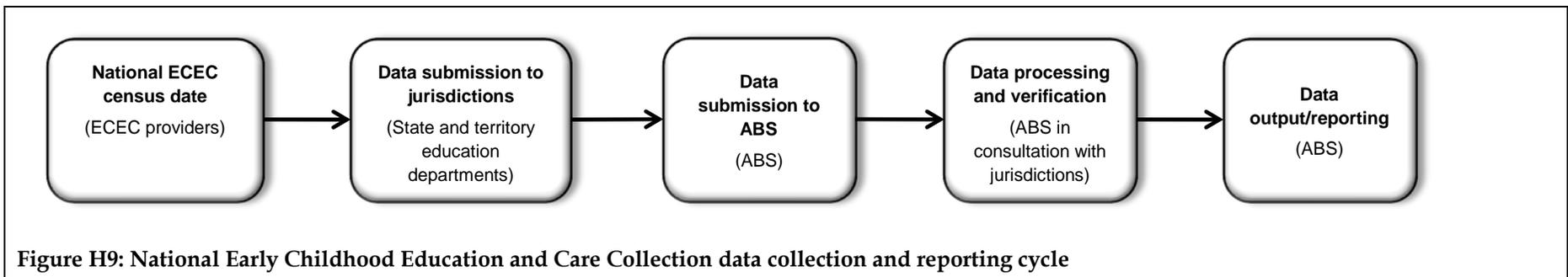
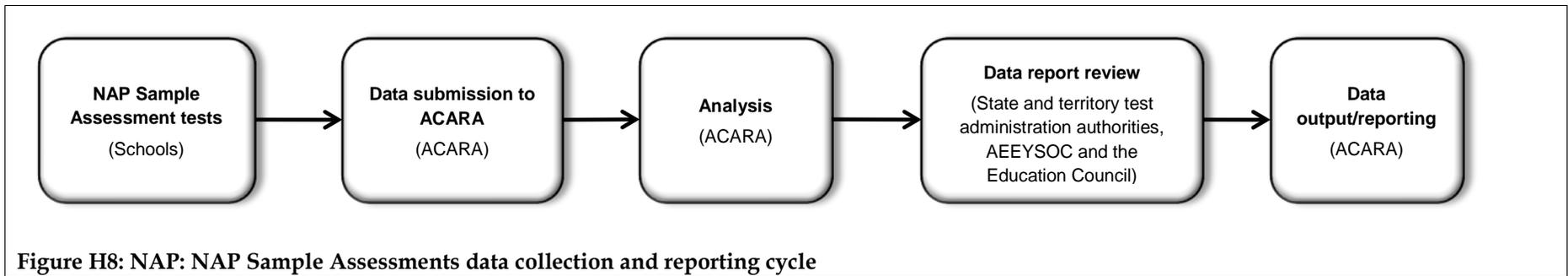
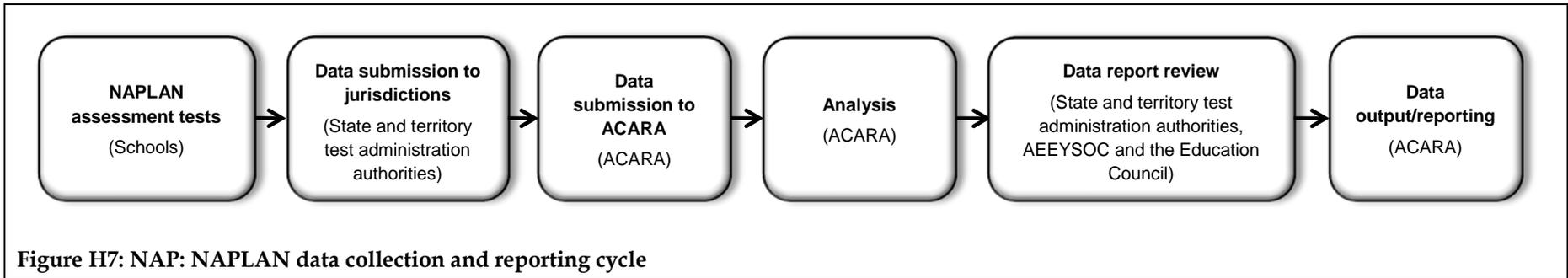
The information on processes and timeframes in figures H1 to H13 is based on the most recent cycle of the data collection and is subject to change for future cycles.

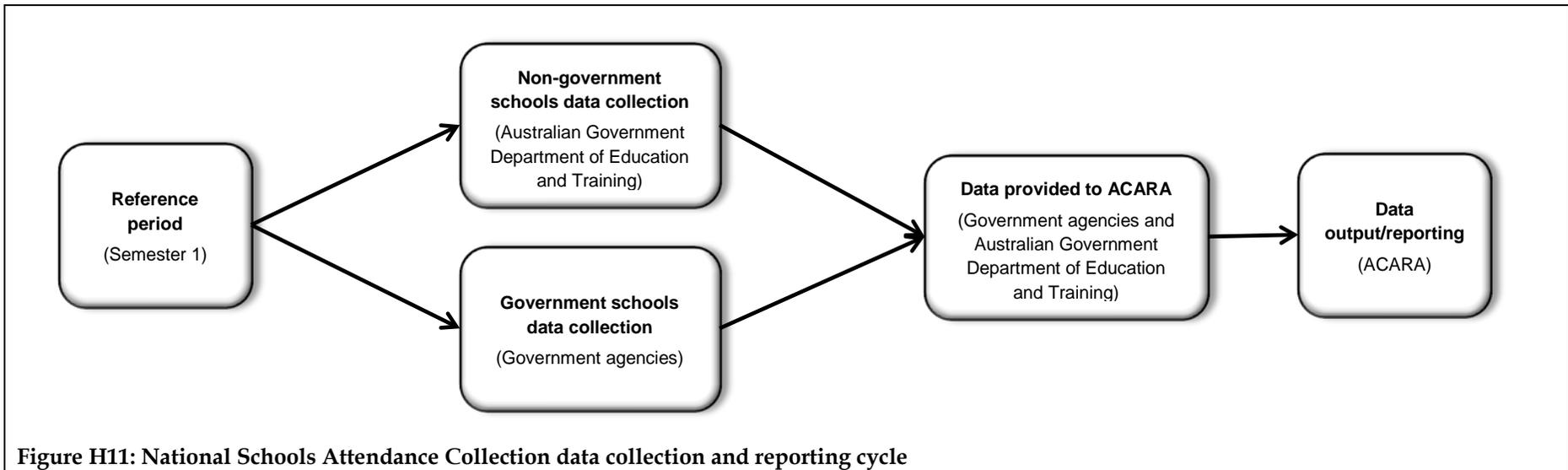
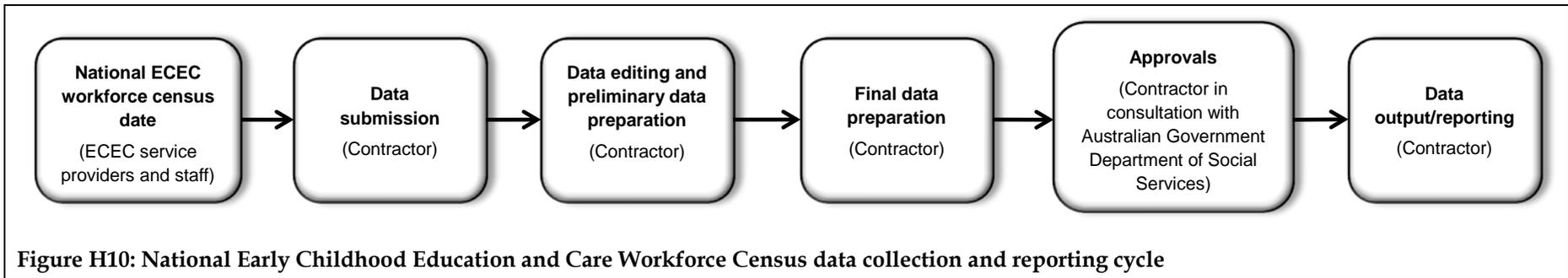
The information that is available for each data collection varies and is dependent on the input from data custodians.

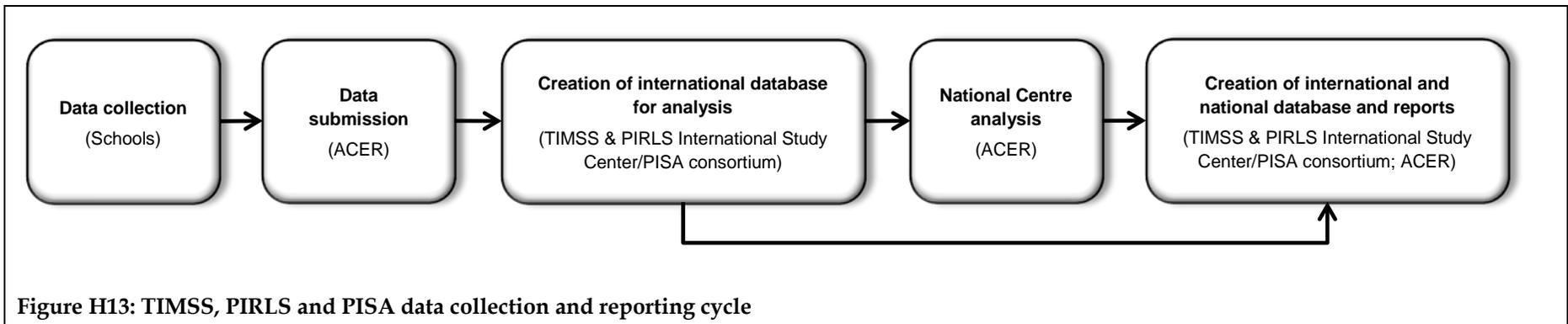
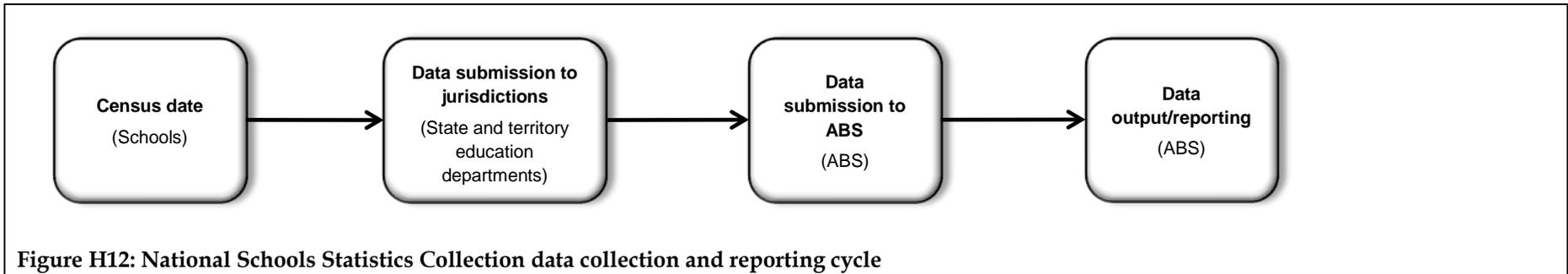












References

- ABS (Australian Bureau of Statistics) 2011. Census dictionary 2011. ABS Cat. No. 2901.0. Canberra: ABS. Viewed 19 December 2013, <<http://www.abs.gov.au/ausstats/abs@.nsf/mf/2901.0>>.
- ABS 2013a. Australians' journeys through life: stories from the Australian Census Longitudinal Dataset, 2006–2011. ABS Cat. No. 2801.0. Canberra: ABS. Viewed 24 February 2014, <<http://www.abs.gov.au/ausstats/abs@.nsf/mf/2081.0>>.
- ABS 2013b. National Schools Statistics Collection (NSSC) data collection manual 2013. Canberra: ABS.
- ABS 2013c. Research paper: Assessing the quality of linking school enrolment records to 2011 Census data, Apr 2013. ABS Cat. No. 1351.0.55.041. Canberra: ABS.
- ABS 2013d. Research paper: Assessing the quality of linking school enrolment records to 2011 Census data: deterministic linkage methods, Dec 2013. ABS cat. no. 1351.0.55.045. Canberra: ABS.
- ABS 2014a. Educational outcomes, experimental estimates, Tasmania, 2006–2013, July 2014. ABS cat. no. 4261.6. Canberra: ABS. Viewed 26 March 2015, <<http://www.abs.gov.au/ausstats/abs@.nsf/mf/4261.6>>.
- ABS 2014b. Educational outcomes, experimental estimates, Queensland, 2011, December 2014. ABS cat. no. 4261.3. Canberra: ABS. Viewed 26 March 2015, <<http://www.abs.gov.au/ausstats/abs@.nsf/mf/4261.3>>.
- ABS 2014c. Outcomes from Vocational Education and Training in Schools, experimental estimates, Australia, 2006–2011. ABS cat. no. 4260.0. Canberra: ABS.
- ABS 2014d. Research paper: assessing the effect of geographic information on data linkage quality using Vocational Education and Training in Schools data, October 2014. ABS cat. no. 4269.0.55.001. Canberra: ABS.
- ACARA (Australian Curriculum, Assessment and Reporting Authority) 2013. Data catalogue first edition 2013. Sydney: ACARA. Viewed 14 January 2014, <http://www.acara.edu.au/verve/_resources/Data_Catalogue.pdf>.
- ACER (Australian Council for Educational Research) 2012. Codebook for OECD Programme for International Student Assessment 2012. Viewed 28 January 2014, <<http://pisa2012.acer.edu.au/downloads.php>>.
- ACER 2014. Data and reports: Australian TIMSS & PIRLS SPSS data files. Melbourne: ACER. Viewed 28 January 2014, <<http://www.acer.edu.au/timss/data-and-reports>>.
- AIHW (Australian Institute of Health and Welfare) 2014a. Developing the National Early Childhood Development Researchable Data Set. Cat. no. PHE 179. Canberra: AIHW.
- AIHW 2014b. Early childhood education and care: data set specification 2014. Canberra: AIHW. Viewed 28 May 2014, <<http://meteor.aihw.gov.au/content/index.phtml/itemId/555371>>.
- Australian Government 2014. Closing the Gap: Prime Minister's report. Canberra: Australian Government. Viewed 7 May 2014, <http://www.dpmmc.gov.au/publications/docs/closing_the_gap_2014.pdf>.
- Australian Government Department of Education 2013. HEIMS Help, Higher education student elements by name, 2013. Viewed 3 February 2014,

<http://heimshelp.education.gov.au/sites/heimshelp/2013_data_requirements/pages/highereducationstudent_elementsbyname>.

Australian Government Department of Education 2014a. Early childhood workforce initiatives. Canberra: Australian Government Department of Education. Viewed 4 June 2014, <<https://education.gov.au/early-childhood-workforce-initiatives>>.

Australian Government Department of Education 2014b. Higher Education Participation Programme. Canberra: Australian Government Department of Education. Viewed 3 June 2014, <<https://education.gov.au/higher-education-participation-programme-hepp>>.

Australian Government Department of Education (unpublished). Early childhood education and care workforce census: data dictionary 2013.

Australian Government Department of Health 2013. National Partnership Agreement on Preventive Health. Canberra: Australian Government Department of Health. Viewed 2 June 2014, <<http://www.health.gov.au/internet/main/publishing.nsf/Content/phd-prevention-np>>.

Australian Government Department of Social Services (DSS) 2013. National Disability Agreement. Canberra: DSS. Viewed 2 June 2014, <<http://www.dss.gov.au/our-responsibilities/disability-and-carers/program-services/government-international/national-disability-agreement>>.

Australian Qualifications Framework Council 2013. Australian Qualifications Framework. Viewed 3 June 2014, <<http://www.aqf.edu.au/wp-content/uploads/2013/05/AQF-2nd-Edition-January-2013.pdf>>.

COAG (Council of Australian Governments) 2009a. Investing in the early years – a national early childhood development strategy. Canberra: COAG. Viewed 7 May 2014, <https://www.coag.gov.au/sites/default/files/national_ECD_strategy.pdf>.

COAG 2009b. National Partnership Agreement on Preventive Health. Canberra: COAG. Viewed 2 June 2014, <http://www.federalfinancialrelations.gov.au/content/npa/health_preventive/national_partnership.pdf>.

COAG 2009c. National Partnership Agreement on School Pathways Programs. Canberra: COAG. Viewed 4 June 2014, <http://www.federalfinancialrelations.gov.au/content/npa/education/school_pathways_program/national_partnership.pdf>.

COAG 2009d. National Partnership Agreement on the National Quality Agenda for Early Childhood Education and Care. Canberra: COAG. Viewed 20 May 2014, <http://files.acecqa.gov.au/files/NQF/nap_national_quality_agenda_early_childhood_education_care_signature.pdf>.

COAG 2011. National disability strategy 2010–2020. Canberra: COAG. Viewed <https://www.coag.gov.au/sites/default/files/national_disability_strategy_2010-2020.pdf>.

COAG 2012a. Intergovernmental Agreement for the National Disability Insurance Scheme (NDIS) launch. Canberra: COAG. Viewed 26 August 2014, <<http://www.ndis.gov.au/document/180>>.

COAG 2012b. National Agreement for Skills and Workforce Development. Canberra: COAG. Viewed 4 June 2014, <http://www.federalfinancialrelations.gov.au/content/npa/skills/skills-reform/national_agreement.pdf>.

COAG 2012c. National Disability Agreement. Canberra: COAG. Viewed 2 June 2014, <http://www.federalfinancialrelations.gov.au/content/national_agreements.aspx>.

COAG 2012d. National Education Agreement. Canberra: COAG. Viewed 3 June 2014, <<http://www.federalfinancialrelations.gov.au/content/npa/education/national-agreement.pdf>>.

COAG 2012e. National Indigenous Reform Agreement (Closing the Gap). Canberra: COAG. Viewed 7 May 2014, <http://www.federalfinancialrelations.gov.au/content/npa/health_indigenous/indigenou s-reform/national-agreement_sept_12.pdf>.

COAG 2013a. National Education Reform Agreement. Canberra: COAG. Viewed 13 August 2015, <http://www.federalfinancialrelations.gov.au/content/other_related_agreements/current/National_Education_Reform_Agreement_2013.pdf>.

COAG 2013b. National Partnership Agreement for Universal Access to Early Childhood Education. Canberra: COAG. Viewed 2 June 2014, <http://www.federalfinancialrelations.gov.au/content/npa/education/early_childhood/n ational_partnership.pdf>.

COAG 2014. National Partnership Agreement for More Support for Students with Disabilities. Canberra: COAG. Viewed 2 June 2014, <http://www.federalfinancialrelations.gov.au/content/npa/education/more_support_for student_with_disabilites/national-agreement.pdf>.

Gemici S & Nguyen N 2013. Data linkage and statistical matching: options for the Longitudinal Surveys of Australian Youth. NCVER Technical paper no. 77. Adelaide: NCVER.

IEA (International Association for the Evaluation of Educational Achievement) 2013a. PIRLS 2011 international database. Boston: TIMSS & PIRLS International Study Center and IEA. Viewed 28 January 2014, <<http://timss.bc.edu/pirls2011/international-database.html>>.

IEA 2013b. TIMSS 2011 international database. Boston: IEA. Viewed 28 January 2014, <<http://timssandpirls.bc.edu/timss2011/international-database.html>>.

Macfarlane I 2014. Skilling Australia to grow industry. Viewed 4 June 2014, <<http://minister.innovation.gov.au/ministers/macfarlane/media-releases/skilling-australia-grow-industry>>.

MCEECDYA (Ministerial Council for Education, Early Childhood Development and Youth Affairs) 2011. Aboriginal and Torres Strait Islander Education Action Plan 2010–2014. Melbourne: MCEECDYA and Education Services Australia. Viewed 2 June 2014, <http://scseec.edu.au/site/DefaultSite/filesystem/documents/ATSI%20documents/ATSIE AP_web_version_final.pdf>.

MCEETYA (Ministerial Council on Education, Employment, Training and Youth Affairs) 2008. Melbourne Declaration on Educational Goals for Young Australians. Viewed 3 June 2014, <http://www.curriculum.edu.au/verve/_resources/National_Declaration_on_the_Educati onal_Goals_for_Young_Australians.pdf>.

National Statistical Service 2015. Statistical data integration involving Commonwealth data. Canberra: ABS. Viewed 8 April 2015, <<http://www.nss.gov.au/nss/home.NSF/pages/Data+Integration+Landing%20Page?OpenDocument>>.

NCVER (National Centre for Vocational Education Research) 2013a. AVETMISS data element definitions: edition 2.2. April 2013. Adelaide: NCVER. Viewed 11 December 2013, <<http://www.ncver.edu.au/wps/wcm/connect/e659553e-7e44-401a-b8f5-b3a69eb4cf8f/AVETMISS7-2-2-data-elements-2579.pdf?MOD=AJPERES>>.

NCVER 2013b. AVETMISS VET Provider collection specifications: release 7.0. April 2013. Adelaide: NCVER. Viewed 11 December 2013, <<http://www.ncver.edu.au/wps/wcm/connect/dd52c92e-46c2-4217-9a21-6135c405f482/Avetmiss-r7-vet-provider-specifications-2578.pdf?MOD=AJPERES&CACHEID=dd52c92e-46c2-4217-9a21-6135c405f482>>.

NCVER 2013c. VET in Schools 2013: terms and definitions. Adelaide: NCVER. Viewed 11 December 2013, <<http://www.ncver.edu.au/wps/wcm/connect/8f98f237-e5a7-41fc-845a-f03597b952c7/Terms+and+definitions+-+VET+in+schools.PDF?MOD=AJPERES&CACHEID=8f98f237-e5a7-41fc-845a-f03597b952c7>>.

Productivity Commission 2015. Childcare and early childhood learning. Productivity Commission inquiry report. Canberra: Productivity Commission. Viewed 5 May 2015, <<http://www.pc.gov.au/inquiries/completed/childcare/report>>.

SCSEEC (Standing Council on School Education and Early Childhood) 2014. Nationally consistent collection of data: school students with disability. Viewed 2 May 2014, <<http://www.schooldisabilitydatapl.edu.au/>>.

Social Research Centre 2013. Australian Early Development Index 2012 data dictionary. Melbourne: Social Research Centre.

List of tables

Table 2.1: National education and training data collections (15) in scope of the national data standards strategy	6
Table 2.2: Detailed list of data items	7
Table 3.1: Summary of data item consistency with national data standards, by 11 data collections	11
Table 3.2: Data items requiring some or considerable effort to meet national data standards, by data collection.....	13
Table 4.1: Data standards developments for the 15 national education and training data collections, as at June 2014	21
Table 4.2: Frequency of data collection and review of data standards.....	25
Table 4.3: Estimated data standards review periods (shaded cells), by data collection, 2014 to 2021	27
Table A1: Organisations consulted during the development of the national data standards strategy.....	33
Table B1: Groups (5) of consistency	39
Table B2: Summary of mapping of education and training data collections with national standards (number and per cent of data items selected for mapping)	43
Table B3: Detailed mapping of education and training data collections with national standards	48
Table B4: Summary of data standard consistency across collections	65
Table B5: Socio-demographic data items that are derivable or mappable, by data collection	68
Table B6: Socio-demographic data items that are relevant but do not exist in the collection.....	69
Table B7: Socio-demographic data items inconsistent with standards	70
Table B8: Socio-demographic data items not applicable by collection	71
Table B9: Education data items that are derivable or mappable by collection	74
Table B10: Education data items that are relevant but do not exist in the collection.....	75
Table B11: Education data items inconsistent with standards, by collection.....	76
Table B12: Education data items not applicable, by collection.....	77
Table B13: Disability data items that are derivable or mappable, by collection	80
Table B14: Disability data items that are relevant but do not exist in the collection.....	80
Table B15: Disability data items inconsistent with standards, by collection.....	81
Table B16: Family composition data items that are derivable or mappable	83
Table B17: Family composition data items that are relevant but do not exist in the collection.....	83
Table B18: Statistical linkage data items that are derivable or mappable, by collection	86
Table B19: Statistical linkage data items that are relevant but do not exist in the collection.....	86
Table B20: Statistical linkage data items inconsistent with standards, by collection	87
Table B21: Statistical linkage data items not applicable, by collection.....	87
Table F1: Summary of current policy priorities with relevance to education and training as at October 2014	162

List of figures

Figure 1.1: Development of the national data standards strategy: project overview	2
Figure 3.1: Consistency of data items with national data standards, by 11 data collections	12
Figure B1: Process to determine consistency of an item to a national standard and relevance to a data collection.....	37
Figure B2: Consistency of all data items, by data collection.....	64
Figure B3: Groups of consistency with standards for socio-demographic data items.....	66
Figure B4: Groups of consistency with standards for education data items.....	72
Figure B5: Groups of consistency with standards for disability data items.....	78
Figure B6: Groups of consistency with standard for family composition data item.....	82
Figure B7: Groups of consistency with standards for statistical linkage data items.....	84
Figure G1: Australian Early Development Census (AEDC) timeframes and process for review.....	167
Figure G2: Census of Population and Housing timeframes and process for review.....	168
Figure G3: Higher Education Statistics Collection timeframes and process for review.....	169
Figure G4: National Apprentice and Trainee Collection, VET Provider Collection and VET in Schools timeframes and process for review	170
Figure G5: NAP: NAPLAN and NAP Sample Assessments timeframes and process for review	171
Figure G6: National Early Childhood Education and Care Collection timeframes and process for review	172
Figure G7: National Early Childhood Education and Care Workforce Census timeframes and process for review.....	173
Figure G8: National Schools Attendance Collection timeframes and process for review	174
Figure G9: National Schools Statistics Collection timeframes and process for review	175
Figure G10: PISA timeframes and process for review at the national level.....	176
Figure G11: TIMSS timeframes and process for review at the national level.....	177
Figure G12: PIRLS timeframes and process for review at the national level.....	178
Figure H1: Australian Early Development Census (AEDC) data collection and reporting cycle.....	180
Figure H2: Census of Population and Housing data collection and reporting cycle.....	180
Figure H3: Higher Education Statistics Collection data collection and reporting cycle.....	181
Figure H4: National Apprentice and Trainee Collection data collection and reporting cycle	181
Figure H5: VET Provider Collection data collection and reporting cycle	182
Figure H6: VET in Schools Collection data collection and reporting cycle	182
Figure H7: NAP: NAPLAN data collection and reporting cycle	183
Figure H8: NAP: NAP Sample Assessments data collection and reporting cycle	183
Figure H9: National Early Childhood Education and Care Collection data collection and reporting cycle.....	183

Figure H10: National Early Childhood Education and Care Workforce Census data collection and reporting cycle.....184

Figure H11: National Schools Attendance Collection data collection and reporting cycle.....184

Figure H12: National Schools Statistics Collection data collection and reporting cycle185

Figure H13: TIMSS, PIRLS and PISA data collection and reporting cycle185

The Australian Institute of Health and Welfare developed a national data standards strategy and implementation plan to enhance the comparability, quality and coherence of information across the Australian education and training sectors, including early childhood education, school education, vocational education and training (VET) and higher education. This project report summarises the activities and process undertaken over 12 months from October 2013 to develop the strategy and implementation plan.