NAP

CIVICS AND CITIZENSHIP 2019

TECHNICAL REPORT

acara AUSTRALIAN CURRICULUM, ASSESSMENT AND REPORTING AUTHORITY

Australian CURRICULUM

NAP NATIONAL ASSESSMENT PROGRAM

My School®

The technical report was written by Kate O'Malley, Renee Kwong, Julian Fraillon, Jorge Fallas, Dr Martin Murphy, Dr Judy Nixon, Eunjung Lee, Greg Macaskill and Vernon Mogul from the Australian Council for Educational Research (ACER).

The NAP–CC project staff from ACARA included Michelle Robins, ACARA's NAP Sample Project Manager; Eveline Gebhardt, ACARA's Lead Psychometrician; Mark McAndrew, ACARA's Civics and Citizenship Curriculum Specialist; and Matthew Stokoe, ACARA's NAP Sample Project Officer.

Julian Fraillon (ACER) directed the ACER work on NAP–CC. Dr Judy Nixon (ACER) led the test development and marking operations teams, Kate O'Malley (ACER) led the field operations and assessment administration team, Renee Kwong (ACER) led the scaling and data analysis team and Dr Martin Murphy (ACER) led the sampling team.

Table of Contents

List of Tables

List of Figures

Chapter 1 INTRODUCTION

The National Assessment Program (NAP) began as an initiative of ministers of education in Australia to monitor outcomes of schooling specified in the 1999 Adelaide Declaration on National Goals for Schooling in the 21st Century (Adelaide Declaration). The NAP was established to measure student achievement and to report this against key performance measures (KPMs) in relation to the national goals. It was agreed that nationally comparable data across jurisdictions would be collected in the domains of literacy, numeracy, science literacy, information and communication technology (ICT) literacy, and civics and citizenship.

Literacy and numeracy achievements are measured and reported via the National Assessment Program – Literacy and Numeracy (NAPLAN). Achievement in science literacy, ICT literacy, and civics and citizenship are assessed under the NAP sample assessments program. These assessments are developed and managed by the Australian Curriculum, Assessment and Reporting Authority (ACARA) under the auspices of the Education Council.

In 2008, the Adelaide Declaration was superseded by the Melbourne Declaration on Educational Goals for Young Australians (Melbourne Declaration). In 2019, the Melbourne Declaration was superseded by the Alice Springs (Mparntwe) Education Declaration. Throughout this time the work of the NAP has continued.

The first collection of data from students in the National Assessment Program – Civics and Citizenship (NAP–CC) was in 2004; subsequent cycles of assessment have been conducted in 2007, 2010, 2013, 2016 and 2019.

This report describes the procedures and processes involved in the conduct of NAP–CC 2019.

To access the NAP–CC public report and technical report documents visit www.nap.edu.au > ‘Results and reports’ section > ‘National reports’ page.

WHAT IS ASSESSED IN NAP–CC?

The context in which civics and citizenship is assessed in Australia has evolved since the beginning of the NAP–CC program. Throughout this period, a commonly agreed theme has been that civics and citizenship education aims to enable students to become active and informed citizens. From its inception, NAP–CC has consequently collected data on students’ knowledge and understanding of civics and citizenship content as well as the attitudes, values and behaviours that relate to participatory citizenship.

The NAP–CC Assessment Framework which guides the development of the assessment was revised in 2018 to align with the content knowledge and skills of various sections of the Australian Curriculum. The NAP–CC Assessment Framework includes five components based on the Australian Curriculum in

Humanities and Social Sciences and Civics and Citizenship. Steps were taken to ensure that trends over time could continue to be reported.

As part of the development of NAP–CC 2019, a new set of assessment items was developed at each year level with a focus on the NAP–CC History sub-strand of the assessment framework.

The student survey collected data relevant to the affective domain of the assessment framework. Further details of the assessment framework are presented in Chapter 2 of this report.

CIVICS AND CITIZENSHIP STUDENT ASSESSMENT AND SURVEY

Assessment delivery

The NAP–CC 2019 assessment was delivered exclusively via the national online assessment platform. All student cognitive and survey data were captured using this online method, and participating students used either their own devices or school-supplied devices that were connected to the internet to complete the assessment.

The online platform used for NAP–CC was the same as that used in NAPLAN Online. Given the widespread compatibility of schools' IT systems with the online platform, offline delivery methods such as school-server solutions or USB delivery methods were not used to administer the assessment in 2019.

The assessment comprised a single test session of 60 minutes for Year 6 students and 75 minutes for Year 10 students. The entire assessment administration time was no more than two hours in total. This two-hour period included time for settling the students into the test room, logging students into the devices and then into the assessment platform, reading the test instructions to students, administering the test itself and then conducting the student survey.

Before starting the assessment component, students completed a set of three practice questions. These practice questions introduced students to the navigation features of the online testing environment as well as to the different item types and formats used in the assessment.

Civics and citizenship assessment instrument

The NAP–CC student assessment instrument comprised a total of 179 items. This pool was divided into items that were delivered to Year 6 students only, to Year 10 students only and to both year levels. In total, 91 items were available to be completed by Year 6 students and 126 items were available to be completed by Year 10 students. At each year level not all items were completed by any single student. The full set of assessable content was distributed across a number of rotated tests for each year level. The Year 6 test booklets contained 39 items and the Year 10 test booklets contained 42 items each.

Item response types

The items developed for the NAP–CC 2019 assessment instrument belonged to one of four response categories:

- standard multiple choice, for which students were asked to select the best answer from a list of typically four distinct options
- multiple choices response, for which students were asked to select all possible answers from a list of four or more distinct options
- short constructed response, which required students to provide typed responses from one word through to a maximum of three sentences
- interactive match, which required students to provide their response to an item by using 'drag and drop' or hotspot functions.

Civics and citizenship survey

At the conclusion of the civics and citizenship assessment, all students completed a survey. The Year 6 survey contained 78 items, while the Year 10 survey contained 97 items. As in previous cycles, the Year 10 survey comprised the Year 6 survey with additional items that were exclusive to Year 10.

Unlike the actual NAP–CC assessment, the student survey was not timed and students were able to take as long as they needed to complete it.

The student survey collected information relating to students':

- perceptions of the importance of citizenship behaviours
- trust in civic institutions and processes
- attitudes towards Australian Indigenous cultures
- attitudes towards Australian diversity
- perceptions of problems affecting Australia
- civics and citizenship participation at school and in the community
- intentions to engage in civic action.

STUDENT BACKGROUND DATA

Data regarding individual student background characteristics were collected and matched to students' cognitive and survey responses for analysis and reporting purposes. Where data were held centrally, it was supplied to ACER by the relevant educational authority. Where central data collection was not possible, this information was supplied directly by the schools themselves.

SAMPLE

The NAP–CC 2019 assessment was administered to a representative sample of students across Australia, at Year 6 and Year 10. For this purpose, a two-stage sampling design was implemented, in line with the methodology used in previous NAP–CC cycles as well as all other assessments (NAP–ICT Literacy and NAP–Science Literacy) within the NAP sample assessments program. The sample for Year 6 and the sample for Year 10 were drawn independently of each other. Full details of the sampling procedures are provided in Chapter 3.

The assessment instrument was administered in Term 4 of 2019. Data were provided by 5,611 Year 6 students in 332 schools and 4,510 Year 10 students in 295 schools.

Results for Tasmanian Year 10 students should be interpreted with caution in this report. Issues with test administration may have reduced the representativeness of participating schools and may have caused a negative impact on student engagement and performance due to timing of the testing near the end of the school year. Tasmanian government schools were given late notification of requirements to participate, resulting in non-participation by 10 of the 26 sampled schools. The participation by the other 16 schools between 25/11/2019–6/12/2019 was beyond the scheduled testing window i.e. 21/10/2019–1/11/2019.

Participation of Tasmanian government high schools later in the year than originally planned coincided with competing priorities of students and school staff, such as end of year exams, reporting, and planned excursions. This may have negatively impacted engagement with this assessment by some students, which is difficult to quantify, but may be evident in a higher proportion of students with very low achievement results.

In the Northern Territory, participation rates were lower than previous years for Year 6 and Year 10. A likely reason for this is a cluster of schools that were sampled but were found not to have sufficient bandwidth to administer the test online. Replacement schools were in the same situation.

Non-participation issues were reduced by adjusting weights within jurisdictions and within sector (see Chapter 3 for further details). However, these adjustments were not able to control for socio-economic differences between participating and non-participating schools within that sector.

At the national level, the impact of the sample shortfall was negligible. The national participation rates were acceptable at both year levels. The national estimates are comparable with those of all previous cycles.

REPORTING THE NAP–CC 2019 RESULTS

The results of the assessment are reported in the National Assessment Program – Civics and Citizenship 2019 Years 6 and 10 Report (ACARA, 2020).

A reporting scale for NAP–CC student achievement was established, using methods based on the one-parameter item response theory model (the Rasch model). In 2004, the Year 6 cohort was defined as having a mean scale score of 400 and a standard deviation of 100 scale score units. The Year 10 mean and standard deviation in 2004 were determined by the performance of Year 10 relative to the Year 6 parameters.

Using common item equating procedures (for secure trend items administered in more than one testing cycle) the results for NAP–CC 2019 were reported on the scale established in 2004. Consequently, the results from NAP–CC 2019 are directly comparable with those from all five previous cycles of NAP–CC (2016, 2013, 2010, 2007 and 2004).

It was also possible to describe students' achievement according to six proficiency levels. Summary descriptions for levels 1 to 5 of the NAP–CC scale were established in the first cycle of NAP–CC in 2004. A description for 'below level 1' achievement was developed in 2007 when more test material was available to support this description. Each level description provides a synthesised overview of the civics and citizenship and history knowledge and understanding that a student working within the level is able to demonstrate. The proficiency level descriptors were updated in 2013 to reflect the larger pool of items that had been developed over the cycles since 2004.

In 2019, the scale descriptors were further revised to reflect the inclusion of items from the NAP–CC History sub-strand of the revised NAP–CC Assessment Framework.

In addition to deriving the NAP–CC scale, proficient standards were established in 2004 for Year 6 and Year 10. The proficient standards "represent a 'challenging but reasonable' expectation of student achievement at a year level with students needing to demonstrate more than elementary skills expected at that year level" (ACARA 2019a, p. 5).

The proficient standard for Year 6 and the proficient standard for Year 10 were established in 2004 on the NAP–CC scale. The proficient standard for Year 6 is 405 scale points, which is the boundary between levels 1 and 2 on the NAP–CC scale. The proficient standard for Year 10 is 535 scale points, which is the boundary between levels 2 and 3 on the scale. Year 6 students performing at level 2 and above, and Year 10 students performing at level 3 and above have consequently met or exceeded their relevant proficient standard. Further details of the proficient standards are provided in Chapter 6.

Student achievement for Year 6 and for Year 10 was reported at the national level and by the following population subgroup categories: jurisdiction, gender, Indigenous status, language spoken at home, school geographic location, and parental occupation and education. Appendix A7 of this technical report includes student achievement reported by percentiles nationally and by jurisdiction across all cycles of NAP–CC.

Results of the student survey for Year 6 and for Year 10 were reported at the national level and by gender only. Where relevant, measures of the association between information from the student survey and student achievement were also reported.

STRUCTURE OF THE TECHNICAL REPORT

This report describes the technical aspects of the NAP–CC 2019 sample assessment and summarises the main activities involved in the data collection, the data collection instruments and the analysis and reporting of the data.

Chapter 2 describes the assessment framework and the process of item development and construction of the instruments.

Chapter 3 reviews the sample design and describes the sampling process. It also describes the weighting procedures that were implemented to derive population estimates and the calculation of response rates.

Chapter 4 describes the data collection and data management procedures used in NAP–CC 2019. This includes the various methods of data capture that were employed before, during and after the administration of the assessment, as well as the procedures applied in the transfer, tracking, verification and transformation of the data collected.

Chapter 5 describes the scaling model and procedures, item calibration, the creation of plausible values and the standardisation of student scores.

Chapter 6 outlines the NAP–CC achievement levels and proficiency standards.

Chapter 7 discusses the reporting of student results, including the procedures used to estimate sampling and measurement variance, and the multivariate analyses conducted with data from NAP–CC 2019.

Chapter 2 ASSESSMENT FRAMEWORK AND INSTRUMENT DESIGN

INTRODUCTION

As part of the preparation for the NAP–CC assessment in 2019, the NAP–CC assessment domain was revised with reference to the Australian Curriculum: HASS (Foundation to Year 6) and Australian Curriculum: History (Years 7 to 10). Some items relating to both the civics and citizenship and the history curriculum were aligned where appropriate for the 2019 assessment.

The revised NAP–CC Assessment Framework for the 2019 cycle, released in December 2018 (ACARA, 2018) summarises the civics and citizenship and history content to be assessed, and the cognitive processes that are extant when students complete the NAP–CC assessment items. The revised framework also contains a review of the affective domain of civics and citizenship which is assessed through the student survey.

As in previous cycles, the items for the NAP–CC 2019 assessment cycle were developed in units. Each unit comprised one or more assessment items that were developed around a single theme or stimulus. In its simplest form a unit was a single, self-contained item, and, in its most complex form, it was a piece of stimulus material (text and/or graphic images) with a set of assessment items related to it. Each assessment item was referenced to an Australian curriculum code and to a single cognitive process from the revised NAP–CC Assessment Framework.

Item-response types in the 2019 NAP–CC assessment were more varied in this cycle than in previous cycles and included multiple-choice, multiple-choices, interactive (drag-and-drop) and constructed response. The scores allocated to items varied: multiple–choice items had a maximum score of one point for correct responses and zero points for incorrect ones. For items where there were multiple choices, the maximum possible scores were either two or three. For constructed response items students could receive between zero and three points. The assessment was conducted using a total of 179 items.

THE REVISED NAP–CC ASSESSMENT FRAMEWORK AND THE AUSTRALIAN CURRICULUM

To complement the development of the Australian Curriculum, ACARA specified that NAP studies should be directly aligned with the Australian Curriculum and in 2018 the NAP–CC Assessment Framework was reviewed and adjusted to reflect this focus.

The alignment was achieved through the modification of the existing assessment framework and its coverage of the knowledge, processes and skills relevant for assessing Civics and Citizenship through the HASS F–6/7 curriculum (where Civics and Citizenship is a sub-strand) and the Civics and Citizenship and History

7–10 curriculum. It was the intention of the revision of the framework to provide some information on the intersection between history and civics and citizenship in the Australian Curriculum.

The 2010 NAP–CC Assessment Framework was used as the reference for the expanded 2018 review. The 2010 NAP–CC Assessment Framework was a comprehensive coverage of the assessment domain, based on the statements of goals of the Melbourne Declaration (2008). The 2010 framework provided guidance for the 2010, 2013 and 2016 NAP–CC assessments. Fundamentally, the 2010 framework also provides the underpinning aspects of the revised 2018 framework.

The following points are contained in the description of the content domain in the revised 2018 framework:

- Political and legal systems
- The nature of citizenship, diversity and identity in contemporary society
- The federal system of government, derived from the Westminster system
- The liberal democratic values such as freedom, equality and the rule of law
- How the people, as citizens, choose their governments
- How the system safeguards democracy by vesting people with civic rights and responsibilities
- How laws and the legal system protect people's rights
- How individuals and groups can influence civic life
- How Australia is a secular nation with a multicultural, multifaith society and a Christian heritage
- The broader values such as respect, civility, equity, justice and responsibility
- The experiences and contributions of Aboriginal and Torres Strait Islander Peoples and their identities within contemporary Australia
- Australia's position and international obligations and the role of citizens today, both within Australia and in an interconnected world.

The revised framework also referenced the cognitive skills outlined in the 2010 framework.

- The skills of inquiry, values and dispositions that enable them to be active and informed citizens
- The skills to question, understand and contribute to the world in which they live
- The skills to recognise and appreciate diverse perspectives, empathy, collaboration, negotiation, self–awareness and intercultural understanding.

The 2019 NAP–CC assessment took into account the jurisdictional differences between the states and the territories in their application of the Australian Curriculum.

ITEM DEVELOPMENT

The new items for the 2019 NAP–CC assessment were developed by a team of ACER's expert test developers, many of whom had experience developing NAP–CC items in previous cycles. The test developers first sourced and developed relevant, engaging and focused civics and citizenship (and history) stimulus materials that addressed the revised framework and the relevant Australian Curriculum codes.

As noted previously, the items were developed as units. A unit consisted of one or more assessment items directly relating to a single theme or stimulus. In its simplest form, a unit is a single self-contained item. In its most complex form, a unit is a piece of stimulus material with a set of assessment items directly related to it.

FIELD TRIAL

Prior to the conduct of the main study, the new 2019 NAP–CC assessment items were trialled on a sample of students in order to attest to the items' reliability as measures of the constructs being assessed. The field trial was conducted in June 2019 in a sample of 118 schools drawn from all educational sectors (government, Catholic and independent) in New South Wales, Victoria and Queensland to ensure a minimum of 200 responses per item.

In each school, the field trial assessment involved one randomly selected intact class from either Year 6 or Year 10. Each student completed a test of civic and citizenship (and related historical) knowledge, followed by a survey about students' experience of, and engagement in, civic issues.

In total, 219 items were used in the field trial, 69 of which were secure items from previous assessment cycles. The items were presented in balanced cluster rotations, with three clusters per test booklet. This is further described in the next section.

MAIN STUDY ASSESSMENT INSTRUMENT

For the main study assessment, schools from all educational sectors and from all states and territories across Australia participated. For the Year 6 assessment, data were gathered from 5,611 students from 332 schools. For the Year 10 assessment, data were collected from 4,510 students from 295 schools.

The main assessment was conducted using seven different test forms at Year 6 and nine different test forms at Year 10. Each test form contained approximately 39 items at Year 6 and 42 items at Year 10 and were compiled using new items and secure items from previous cycles, for equating purposes. The assessment was conducted using a total of 179 items, with 30 secure items at Year 6 and 40 secure items at Year 10.

The secure items included in the 2019 NAP–CC assessment are a subset of items from previous assessments that have not been released to the public. These items enabled the equating of the 2019 scale, via the 2016 scale, onto the historical scale created in the first NAP–CC assessment cycle in 2004. Transformations are conducted of the logit scale to a scale with a mean of 400 and a standard deviation of 100, as per the historical 2004 scale. This has enabled an examination of student performance over time through each of the NAP–CC assessment cycles.

Table 2.1 shows the number of main study test items by year level corresponding to the content knowledge codes in the Australian Curriculum: Civics and Citizenship.

Table 2.1: Main study test items corresponding to knowledge codes in the Australian Curriculum: Civics and Citizenship

Year 6		Year 10	
Australian Curriculum Knowledge Code	Number of items	Australian Curriculum Knowledge Code	Number of items
ACHASSI098	1	ACDSEH087	1
ACHASSI127	1	ACDSEH090	1
ACHASSK062	1	ACDSEH091	4
ACHASSK065	1	ACDSEH095	1
ACHASSK083	1	ACDSEH096	2
ACHASSK085	1	ACDSEH108	1
ACHASSK093	1	ACDSEH109	9
ACHASSK115	5	ACDSEH144	2
ACHASSK116	1	ACHASSI098	1
ACHASSK117	1	ACHASSI127	1
ACHASSK118	6	ACHASSK064	1
ACHASSK134	1	ACHASSK134	1
ACHASSK135	7	ACHASSK135	6
ACHASSK143	3	ACHASSK143	1
ACHASSK144	3	ACHASSK147	6
ACHASSK145	3	ACHASSK197	1
ACHASSK146	1	ACHCK014	1
ACHASSK147	13	ACHCK023	1
ACHASSK148	1	ACHCK024	1
ACHASSK197	2	ACHCK027	2
ACHCK012	1	ACHCK035	2
ACHCK014	1	ACHCK036	4
ACHCK022	2	ACHCK039	2
ACHCK023	1	ACHCK048	4
ACHCK024	1	ACHCK049	2
ACHCK025	1	ACHCK052	1
ACHCK027	2	ACHCK053	4
ACHCK035	1	ACHCK061	1
ACHCK036	5	ACHCK062	7
ACHCK039	2	ACHCK063	2
ACHCK049	1	ACHCK066	7
ACHCK052	3	ACHCK075	6
ACHCK053	3	ACHCK076	6
ACHCK061	2	ACHCK077	1
ACHCK062	6	ACHCK078	4
ACHCK066	1	ACHCK080	4
ACHCK075	1	ACHCK081	2
ACHCK077	1	ACHCK090	1
ACHCK078	2	ACHCK091	5
		ACHCK092	3
		ACHCK093	2
		ACHCK094	6
		ACHCK103	2
		ACOKFH015	1
		ACOKFH019	1
		ACSEH095	2
Total	**91**	**Total**	**126**

Note: Common items used at each of Year 6 and Year 10 have been included in both year levels in the above table.

Table 2.2 shows the number of items by year level and item format that were used in the main study item pool.

Table 2.2: Main study test item pool by item format

Item Format	Year Level			Total
	6 only	6 & 10	10 only	
Composite	10	2	6	18
Extended text	9	9	35	53
Interactive match		2	3	5
Multiple choice	31	24	40	95
Multiple choices	3	1	4	8
Total	**53**	**38**	**88**	**179**

Test forms comprised three test clusters. As well as balancing the order and combinations of clusters across the test forms, each individual cluster was matched for reading load (length and difficulty), item format and use of graphic images. A small number of the secure items was also distributed across the clusters. By matching each individual cluster for these characteristics, each test form could then be considered as matched and equivalent in terms of its characteristics.

The test form designs for Years 6 and 10 are shown in Table 2.3 and Table 2.4.

Table 2.3: Test form design for Year 6 main study assessment

Test form	Position 1	Position 2	Position 3	Position 4	Position 5
06B1	Practice Questions	06C1	06C2	06C3	6 Survey
06B2	Practice Questions	06C2	06C3	06C4	6 Survey
06B3	Practice Questions	06C3	06C4	06C5	6 Survey
06B4	Practice Questions	06C4	06C5	06C6	6 Survey
06B5	Practice Questions	06C5	06C6	06C7	6 Survey
06B6	Practice Questions	06C6	06C7	06C1	6 Survey
06B7	Practice Questions	06C7	06C1	06C2	6 Survey

Table 2.4: Test form design for Year 10 main study assessment

Test form	Position 1	Position 2	Position 3	Position 4	Position 5
10B1	Practice Questions	10C1	10C2	10C4	10 Survey
10B2	Practice Questions	10C2	10C3	10C5	10 Survey
10B3	Practice Questions	10C3	10C4	10C6	10 Survey
10B4	Practice Questions	10C4	10C5	10C7	10 Survey
10B5	Practice Questions	10C5	10C6	10C8	10 Survey
10B6	Practice Questions	10C6	10C7	10C9	10 Survey
10B7	Practice Questions	10C7	10C8	10C1	10 Survey
10B8	Practice Questions	10C8	10C9	10C2	10 Survey
10B9	Practice Questions	10C9	10C1	10C3	10 Survey

SCORING GUIDES

Items requiring students to enter an extended text response are referred to as constructed response items. Some of these items had scoring guides that allowed for dichotomous scoring (sufficient/insufficient responses). Other constructed responses had scoring guides with partial credit (polytomous) scoring. In these items, the different categories of student responses could be scored according to the degree of knowledge, skill or understanding the students demonstrated.

Scoring guides for all constructed response items were developed in draft form in parallel with the item development. After the field trial they were edited and added to as needed by ACER test developers. Consultation also took place with the experts and stakeholders at ACARA.

Some of the closed and short constructed response items had a score value of zero (incorrect) or one (correct). The relative simplicity of this dichotomous scoring does not necessarily reflect the differing levels of complexity exhibited in the student responses.

The scoring guides for such items were therefore developed to define and describe these different levels of complexity in a meaningful way. Empirical data from the field trial were used to confirm whether the described semantic distinctions were indicative of actual differences in student achievement.

In those cases where hierarchical differences described in the scoring guides were not evident in the field trial data, these differences were removed from the scoring guides. Typically this would involve providing the same credit for responses that has previously been allocated to different levels of credit (referred to as *collapsing categories*). In this way, the scoring guides for the main study reflected more accurately the levels of achievement by students on these short constructed responses scored by the markers.

Accuracy of marker scoring was also ensured through the text of each item's scoring guide. Each score point allocation in an item's scoring guide was accompanied by text which described and characterised the kind of responses which would attract each score. These score points were then illustrated with actual student responses. The descriptive text, along with the sample student responses for each score point for each item, constituted the complete scoring guide.

Figure 2.1 displays an item from the 2019 main study, along with the full score guide for this item.

The key features of the score guide for this item are:

- The summary description of the substantive properties of the response of each score level
- The detailed description of the properties of the responses of each level and

- Sample student responses that illustrate the properties of the responses at each level.

Figure 2.1: Example of a constructed response item and its scoring guide

<table>
<tr><td>Item stimulus</td><td>
Bejah Dervish (BD_61)

Marie Williams lives in South Australia. Her grandfather and father were skilled camel drivers.
Marie's grandfather, Bejah Dervish, came to Australia from what is now Pakistan in about 1890.
Between the 1840s and the 1930s, camel drivers from countries including Afghanistan, Pakistan and India used camels to deliver goods to remote areas in Australia.

Marie Williams

Bejah Dervish with some of his camels

What is one positive contribution that immigrants such as Marie Williams' family have made to Australia?
</td></tr>
<tr><td>Scoring guide</td><td>
Code 1: Identifies a positive contribution of immigrants to Australia
Immigrants have brought lots of important skills to Australia
• They have influenced trade and business in Australia; introduced transportation in remote areas
• They have introduced new cultures to Australia
• Immigrants have been in Australia for a very long time and therefore influenced Australia
• Australia has had people from many different countries coming here for a long time bringing their experience with them

Code 0: Provides a vague, incoherent, inaccurate or irrelevant response
• People came from the Middle east between 1840-1930
• Introduced camels
• Repeats information given
• Transport [too vague]
• Camel riding
• They have a better life [doesn't answer question]
</td></tr>
</table>

Further information about the marking of student responses, including the quality control measures implemented during the marking operation, is provided in Chapter 4.

STUDENT SURVEY

The NAP–CC student survey (sometimes referred to as the 'questionnaire' in previous cycles) addresses students' attitudes towards civic and citizenship issues and students' engagement in civic and citizenship activities.

The 2019 student survey was to a large extent the same as previous versions, except for a small number of statements that were added to pre-existing questions, some new items replacing items from the previous survey, and some minor changes to a small number of items.

The final version of the survey was developed following the Working Group review in consultation with ACARA. In addition to some minor revisions to individual items, the survey was revised to include additional information about digital citizenship participation and a new set of items relating to students' perceptions of the degree to which a given set of problems affect Australia.

Students' attitudes towards civic and citizenship issues were assessed with questions covering six constructs:

- Importance of conventional citizenship behaviour
- Importance of social movement-related citizenship behaviour
- Trust in civic institutions and processes
- Attitudes towards Indigenous culture
- Attitudes towards Australian diversity
- Perceptions of the severity of problems affecting Australia.

Each construct was measured using a set of Likert-type items typically consisting of four options (for example, 'strongly agree', 'agree', 'disagree' or 'strongly disagree').

Students' engagement in civic and citizenship activities was assessed with questions concerning the following areas:

- Participation in civic and citizenship related activities at school
- Participation in civic and citizenship related activities in the community
- Media use and participation in discussion of political or social issues
- Interest in political or social issues
- Confidence to actively engage in civic action
- Valuing civic action
- Intentions to promote important issues in the future
- Expectations of future civic engagement.

For the purposes of analysis and reporting, survey scales were created using the items from the aforementioned content areas. Detailed information about the scaling procedures conducted, as well as the psychometric properties of the scales created, are provided in Chapter 5. A copy of the student survey is included in Appendix A1.

Chapter 3 SAMPLING AND WEIGHTING

This chapter describes the NAP–CC 2019 sample design, the achieved sample, and the procedures used to calculate the sampling weights. The sampling and weighting methods were used to ensure that the data provided accurate and efficient estimates of the achievement outcomes for the Australian Year 6 and Year 10 student populations.

SAMPLING

The target populations for the study were Year 6 and Year 10 students enrolled in educational institutions across Australia.

A two-stage stratified cluster sample design was used in NAP–CC 2019, similar to that used in other Australian national sample assessments and in international assessments such as the Trends in International Mathematics and Science Study (TIMSS). The first stage consisted of a sample of schools, stratified according to state, sector, school type, performance in NAPLAN test[1], the Socio-Economic Indexes for Areas (SEIFA)[2], geographical location and school size. The second stage consisted of a sample of 20 random students from the target year level in sampled schools. Samples were drawn separately for each year level.

The sampling frame

Schools were selected from the sampling frame provided by ACARA, and complemented with data from the ACER School Frame, a comprehensive list of all schools in Australia, updated annually using information collected from multiple sources, including the Australian Bureau of Statistics and the Commonwealth, state and territory education departments. The enrolment figures in the sample frame are from the previous school year.

School exclusions

Schools excluded from the target population included: non-mainstream schools (such as schools for students with intellectual disabilities or hospital schools), schools listed as having fewer than five students in the target year levels, and very remote schools (except in the Northern Territory). These exclusions account for 2.6 per cent of the Year 6 student population and 2.9 per cent of the Year 10 student population.

The decision to include very remote schools in the Northern Territory sample was made because very remote schools constituted more than 25 per cent of the Year 6 population and close to 20 per cent of the Year 10 population in the

[1] Schools are grouped by quintiles based on the scores in the NAPLAN survey.
[2] This is a measure of socio-economic status based on the socio-economic conditions, such as education and employment, of the geographic location of the school.

Northern Territory (while this proportion was less than one per cent of the total student population of Australia). The same procedure was used for the 2016 study. The inclusion of very remote schools in the Northern Territory in the NAP–CC 2019 sample had only a negligible impact on the estimates for Australia and the other states.

The designed sample

For both Year 6 and Year 10 samples, sample sizes were chosen to provide accurate estimates of achievement outcomes for all states and territories. The expected 95 per cent confidence intervals for estimated means of the larger states were estimated in advance to be within approximately ±0.15 to ±0.2 of the population standard deviation. This level of precision was considered an appropriate balance between the analytical demands of the study, the burden on individual schools and the overall costs of the study. Confidence intervals of this magnitude require an effective sample size[3] of around 100–150 students in the larger states. The main requirement for achieving acceptable precision for a state or territory is to have a good-sized sample. Although a less important factor, sampling a larger proportion of the population will also improve precision. As the proportion of the total population surveyed becomes larger, the precision of the sample increases for a given sample size: This explains why the sample sizes for the smaller states and territories are smaller compared to the larger states and territories.

Table 3.1 shows the population of schools and students and the designed sample.

Table 3.1: Year 6 and Year 10 target population and designed samples by state and territory

	Year 6			Year 10		
	Enrolment	Schools in Population	Schools in Sample	Enrolment	Schools in Population	Schools in Sample
NSW	94,925	2,068	46	85,606	789	45
Vic.	75,256	1,650	46	69,538	562	45
QLD	65,284	1,151	46	56,944	471	45
SA	19,864	532	47	19,359	195	45
WA	32,891	725	46	28,186	248	45
Tas.	6,466	196	48	6,229	86	42
NT	3,154	116	30	2,661	47	27
ACT	5,387	96	30	4,907	38	30
Aust.	303,227	6,534	339	273,430	2,436	324

[3] The effective sample size is the sample size of a simple random sample that would produce the same precision as that achieved under a complex sample design.

First sampling stage

Stratification by state, sector and small schools was explicit: separate samples were drawn for each sector within states and territories. Stratification by school type, NAPLAN performance, SEIFA, geographic location and school size was implicit: schools within each state were ordered by size (according to the number of students in the target year level) within sub-groups defined by a combination of school type, NAPLAN performance, SEIFA index and geographic location.

The selection of schools was carried out using a systematic probability-proportional-to-size (PPS) method. For large schools, the measure of size (MOS) was equal to the enrolment at the target year. The sum of the measures of size of schools within a stratum is calculated, and divided into *n* equal-sized intervals, where *n* is the number of schools to be sampled from the stratum. The school selection probability is equal to the measure of size of the school divided by the interval size: $\Pr(\text{school selection}) = MOS_{school} / (\sum MOS_{all\ scools\ in\ stratum}) / n)$

The number of students to be sampled from the school is known as the 'target cluster size' (TCS). Students are sampled from the school with equal probability and so the selection probability of a student from a larger school is:

$$\Pr(\text{student selection within school}) = TCS / MOS_{school}$$

The combined effect of this two-stage process is that most students are sampled with equal probability:

$$\Pr(\text{student selection}) = \Pr(school\ selection) * \Pr(student\ selection\ within\ school)$$

$$= MOS_{school} / (\sum MOS_{all\ scools\ in\ stratum}) / n) * TCS / MOS_{school}$$

$$= TCS\ / (\sum MOS_{all\ scools\ in\ stratum}) / n)$$

If a school is selected with target year enrolment less than the TCS (denoted as a 'small school'), all students from that school will be in certain selections and the second term in the above expression becomes 1:

$$\Pr(\text{student selection}) = \Pr(school\ selection) * \Pr(student\ selection\ within\ school)$$

$$= MOS_{small\ school} / (\sum MOS_{all\ scools\ in\ stratum}) / n) * 1$$

$$= MOS_{small\ school} / (\sum MOS_{all\ scools\ in\ stratum}) / n)$$

In order to make the selection probability for these students the same as above, the starting point in the sample design is to set the measure of size for the smaller schools to TCS:

$$MOS_{small\ school} = TCS$$

$$Pr_{(student\ selection)} = TCS\ / (\sum MOS_{all\ scools\ in\ stratum}) / n)$$

For NAP–CC the TCS was set at 20 students. The starting point in the sample design is that all small schools with enrolments from 1 to 19, and all students from those schools, are sampled with equal probability.

This approach minimises variation in weights which is desirable. Large variations in weights can have a major impact on the precision of survey estimates.

The approach described above is used when small schools represent only a very small proportion of the total enrolment in the stratum. When the proportion of the total enrolment in small schools is larger, the number of schools to be sampled from the stratum is increased to cater for the fact that the yield from these smaller schools will be less than the target cluster size. In addition, the smallest of these smaller schools have their selection probabilities reduced, through a reduction in their measure of size, so that fewer of them are included in the sample, that is, they are under-sampled.

To under-sample small schools, all schools in the stratum are classified into one of the following groups based on their enrolment size:

- P1 ('extremely small'): enrolment of 2 or less
- P2 ('very small') enrolment between 3 and half the TCS
- Q ('moderately small'): enrolment from TCS/2 +1 to less than the TCS
- R ('large'): enrolment of TCS or larger

If the proportion of students in P1 and P2 schools in a stratum was 1% or more, or if the proportion of students in Q Schools was 4% or more, then the following adjustments were made:

- The MOS for 'P1' schools was reduced to 0.25 TCS. In this case, with TCS = 20, the MOS for these extremely small schools is reduced to 5
- The MOS for 'P2' schools was reduced to 0.5 TCS (i.e. MOS = 10)
- The total number of schools to be sampled from the stratum is increased, to preserve the desired sample yield from the stratum to close to the product of the TCS and the number of schools to be sampled from the stratum (TCS * n).

The first two adjustments mean that the extremely small and very small schools are sampled at lower rates, to minimise the operational burden of having too many of these very small schools in the sample.

The net effect of these adjustments is that the desired yield from the sample is preserved, variation in weights is kept to a minimum, and the operational burden of having a large number of small schools included in the sample is reduced.

After applying this procedure, the actual numbers of schools sampled for Year 6 and Year 10 were increased to 339 and 324, respectively, as shown in Table 3.1. The actual sample drawn is referred to as the 'implemented sample'.

Within each explicit stratum, the standard process for the selection of schools with PPS was as follows:

The MOS was accumulated from school to school and the running total was listed next to each school. The total cumulative MOS was a measure of the size of the population of students. Dividing this figure by the number of schools to be sampled provided the sampling interval.

The first school was sampled by choosing a random number between one and the sampling interval. The school whose cumulative MOS contained the random number was the first sampled school. By adding the sampling interval to the random number, a second school was identified. This process of consistently adding the sampling interval to the previous selection number resulted in a PPS sample of the required size.

As each school was selected, the next school in the sampling frame was designated as a replacement school to be included in cases where the sampled school did not participate. The school previous to the sampled school was designated as a second replacement. It was used if neither the sampled nor the first replacement school participated. In some cases (such as secondary schools in the Northern Territory) there were not enough schools available for replacement schools to be assigned. Due to the stratified sampling frame, the two replacement schools were generally similar (with respect to school type, NAPLAN performance, socio-economic status, geographical location and size) to the originally sampled school to which they were assigned.

After the school sample had been drawn, sampled schools might be identified as meeting a criterion for exclusion. When this occurred, the sampled school and its replacements were removed from the sample and removed from the calculation of response rates. For this cycle, no school was removed from the Year 6 sample but two schools were removed from the Year 10 sample.

Second sampling stage

The second stage of sampling consisted of the selection of 20 students within sampled schools. Students were sampled with equal probability across classes and were sorted by gender so that the male: female ratio of the selected sample was proportional to the ratio at the grade level. In small schools where the grade level comprised fewer than 20 students, *all* students in that grade level were selected for participation.

Student exclusions

Within the group of sampled students, individual students were excluded from the assessment on the basis of the criteria listed below.

- Functional disability: the student had a moderate to severe permanent physical disability such that they could not perform in the assessment situation.
- Intellectual disability: the student had a significant intellectual disability which severely limited their capacity to participate in the assessment situation.

- Limited English-language proficiency: the student was unable to read or speak the language of the assessment (English) and would be unable to overcome the language barrier in the assessment situation. Typically, a student who had received less than one year of instruction in English would be excluded.

Table 3.2 and Table 3.3 detail the numbers and percentages of students excluded from the NAP–CC 2019 assessment, according to the reason given for their exclusion. The number of student-level exclusions was 525 at Year 6 and 1,255 at Year 10[4]. This gives weighted exclusion rates of 2.6 per cent both for the sampled Year 6 and Year 10 students.

Table 3.2: Year 6 breakdown of student exclusions according to reason by state and territory

	Enrolment in Participating Schools	Reason for Exclusion				Proportion of Year 6 Students in Sampled Schools
		Functional Disability	Intellectual Disability	Limited English Proficiency	Total	
NSW	2,974	10	47	4	61	2.0
Vic.	2,724	11	35	30	76	2.8
QLD	4,000	25	74	22	120	3.0
SA	2,366	22	29	40	91	3.8
WA	2,927	21	12	6	40	1.4
Tas.	2,097	15	24	8	47	2.2
NT	1,145	16	9	6	31	2.7
ACT	2,138	4	42	14	61	2.8
Aust.	**20,370**	124	272	130	525	2.6

Table 3.3: Year 10 breakdown of student exclusions according to reason by state and territory

	Enrolment in Participating Schools	Reason for Exclusion				Proportion of Year 10 Students in Sampled Schools
		Functional Disability	Intellectual Disability	Limited English Proficiency	Total	
NSW	6,342	41	13	63	117	1.8
Vic.	8,308	49	29	238	316	3.8
QLD	8,328	60	89	60	209	2.5
SA	6,834	39	88	73	201	2.9
WA	7,723	45	70	21	136	1.8
Tas.	3,472	5	9	10	24	0.7
NT	2,063	54	37	31	122	5.9
ACT	4,546	39	21	70	130	2.9
Aust.	**47,616**	332	357	566	1,255	2.6

WEIGHTING

While the multi-stage stratified cluster design provides a very economical and effective data collection process in a school environment, oversampling of sub-populations and non-response cause differential probabilities of selection for the ultimate sampling elements, the students. Consequently, one student in the assessment does not necessarily represent the same number of students in the

[4] These exclusions are weighted by the student base weight within the school, so they represent the estimated total number of students in the sampled schools that are part of the segment of the population that would be excluded from the survey if they had been selected in the student sample in the first place, thus providing a more accurate estimate of the within school exclusion rate.

population as another. To account for differential probabilities of selection due to the design and to ensure unbiased population estimates, a sampling weight was computed for each participating student.

The overall sampling weight is the product of weights calculated at the two stages of sampling:

- the selection of the school at the first stage
- the selection of students within the sampled schools at the second stage.

First-stage weight

The first-stage weight is the inverse of the probability of selection of the school, adjusted to account for school non-response.

The probability of selection of the school is equal to its measure of size (MOS) divided by the sampling interval (SINT) or one, whichever is lower. A school with a MOS greater than the SINT is a certain selection and therefore has a probability of selection of one. Some very large schools were selected with certainty into the sample.

The sampling interval is calculated at the time of sampling, and for each explicit stratum it is equal to the cumulative MOS of all schools in the stratum, divided by the number of schools to be sampled from that stratum.

This factor of the first-stage weight, or the school base weight ($\mathrm{BW}_{\mathrm{sc}}$), was the inverse of this probability:

$$BW_{sc} = \frac{SINT}{MOS}$$

Following data collection, counts of the following categories of schools were made for each explicit stratum:

- the number of schools that participated (n_p^{sc})
- the number of schools that were sampled but should have been excluded (n_x^{sc})
- the number of non-responding schools (n_n^{sc}).

Note that $n_p^{sc} + n_x^{sc} + n_n^{sc}$ equals the total number of sampled schools from the stratum.

Examples of the second class (n_x^{sc}) were:

- a sampled school that no longer existed
- a school that, following sampling, was discovered to have fitted one of the criteria for school-level exclusion (e.g. very remote, very small), but which had not been removed from the frame prior to sampling.

In the case of a non-responding school (n_n^{sc}), neither the originally sampled school nor its replacements participated.

Within each explicit stratum, an adjustment was made to account for school non-response. This non-response adjustment (NRA) for a stratum was equal to:

$$NRA_{strt} = \frac{\left(n_p^{sc} + n_n^{sc}\right)}{n_p^{sc}}$$

The first-stage weight, or the final school weight, was the product of the inverse of the probability of selection of the school and the school non-response adjustment:

$$FW_{sc} = BW_{sc} * NRA_{strt}$$

Second-stage weight

Following data collection, counts of the following categories of students were made for each sampled school:

- the total number of students at the relevant year level (n_{tot}^{st})
- the number of students who participated (n_p^{st})
- the number of sampled students who were exclusions (n_x^{st})
- the number of non-responding sampled students (n_n^{st}).

Note that $n_{samp}^{st} = n_p^{st} + n_x^{st} + n_n^{st}$ equals the total number of sampled students from the sampled school.

The first factor in the second-stage weight was the inverse of the probability of selection of the student from the sampled school.

$$BW_{st} = \frac{n_{tot}^{st}}{n_{samp}^{st}}$$

The student-level non-response adjustment was calculated for each school as:

$$NRA_{sc} = \frac{n_p^{st} + n_n^{st}}{n_p^{st}}$$

The final student weight was:

$$FW_{st} = BW_{st} \times NRA_{sc}$$

Overall sampling weight

The full sampling weight (FWGT) was simply the product of the weights calculated at each of the two sampling stages:

$$FWGT = FW_{sc} \times FW_{st}$$

After computation of the overall sampling weights, the weights were checked for outliers, because outliers can have a large effect on the computation of the standard errors. A weight was regarded as an outlier if the value was more than four times the median weight within a subpopulation defined by year level, state or territory and sector (i.e. an explicit stratum). However, no outliers were found in the data for this cycle.

A final post stratification adjustment to the weights was carried out, so that the sum of the weights reflected student population estimates at the time of data collection. The population reference used were the population totals by state and sector obtained from the 2019 NAPLAN administration. The post-stratification adjustment scales the weights of all participating students in the stratum by the same factor. It therefore has no effect on the relative contribution of participating students, and in turn, the estimates within each stratum. It does however ensure that contributions to national estimates reflect the current population sizes for each state and sector. One adjustment factor was estimated for each combination of year level, state and school sector.

Table 3.4 shows the sum of the student final weights for all participating students in the sample by year level and state, both before and after the post-stratification adjustment explained in (39) and its comparison to the initial total population estimates (based on enrolment figures from the previous school year).

Table 3.4: Year 6 and Year 10 comparison of total population and sum of weights by State

	Year 6			Year 10		
		Sum of Student Final Weights			Sum of Student Final Weights	
	Target Population	Before post-stratification adjustment	After post-stratification adjustment	Target Population	Before post-stratification adjustment	After post-stratification adjustment
NSW	94,925	93,876	98,778	85,606	85,448	89,183
Vic.	75,256	75,369	76,716	69,538	68,568	69,284
QLD	65,284	64,210	67,464	56,944	56,162	60,675
SA	19,864	19,689	20,789	19,359	18,726	19,294
WA	32,891	33,006	34,317	28,186	27,524	30,184
Tas.	6,466	6,314	6,622	6,229	7,109	5,928
NT	3,154	3,060	3,410	2,661	3,121	2,916
ACT	5,387	5,253	5,359	4,907	4,820	5,176
Aust.	303,227	300,777	313,455	273,430	271,478	282,640

REPONSE RATES

Separate response rates were computed: (1) with replacement schools included as participants, and (2) with replacement schools regarded as non-respondents. In addition, each of these rates was computed using unweighted and weighted counts. In any of these methods, a school and a student response rate were computed and the overall response rate was the product of these two response rates. The differences in computing the four response rates are described below. These methods are consistent with the methodology used in TIMSS (Martin, Mullis & Hooper, 2015).

Unweighted response rates including replacement schools

The unweighted school response rate, where replacement schools were counted as responding schools, was computed as follows:

$$RR_1^{sc} = \frac{n_s^{sc} + n_{r1}^{sc} + n_{r2}^{sc}}{n_s^{sc} + n_{r1}^{sc} + n_{r2}^{sc} + n_{nr}^{sc}}$$

where n_s^{sc} is the number of responding schools from the original sample, $n_{r1}^{sc} + n_{r2}^{sc}$ is the total number of responding replacement schools, and n_{nr}^{sc} is the number of non-responding schools that could not be replaced.

The student response rate was computed over all responding schools. Of these schools, the number of responding students was divided by the total number of eligible, sampled students:

$$RR_1^{st} = \frac{n_r^{st}}{n_r^{st} + n_{nr}^{st}}$$

where n_r^{st} is the total number of responding students in all responding schools and n_{nr}^{st} is the total number of eligible, non-responding, sampled students in all responding schools.

The overall response rate is the product of the school and the student response rates.

$$RR_1 = RR_1^{sc} \times RR_1^{st}$$

Unweighted response rates excluding replacement schools

The difference of the second method from the first is that the replacement schools were counted as non-responding schools.

$$RR_2^{sc} = \frac{n_s^{sc}}{n_s^{sc} + n_{r1}^{sc} + n_{r2}^{sc} + n_{nr}^{sc}}$$

This difference had an indirect effect on the student response rate because fewer schools were included as responding schools, and student response rates were only computed for the responding schools.

$$RR_2^{st} = \frac{n_r^{st}}{n_r^{st} + n_{nr}^{st}}$$

The overall response rate was again the product of the two response rates.

$$RR_2 = RR_2^{sc} \times RR_2^{st}$$

Weighted response rates including replacement schools

For the weighted response rates, sums of weights were used instead of counts of schools and students. School and student base weights (BW) are the weight values before correcting for non-response, so they generate estimates of the population being represented by the responding schools and students. The full weights (FW) at the school and student levels are the base weights corrected for non-response.

School response rates are computed as follows:

$$RR_3^{sc} = \frac{\sum_i^{s+r1+r2}\left(BW_i \times \sum_j^{r_i}(FW_{ij})\right)}{\sum_i^{s+r1+r2}\left(FW_i \times \sum_j^{r_i}(FW_{ij})\right)}$$

where i indicates a school, $s + r1 + r2$ all responding schools, j a student, and r_i the responding students in school *i.* First, the sum of the student final weights FW_{ij} for the responding students from each school was computed. Second, this sum was multiplied by the school's base weight (numerator) or the school's final weight (denominator). Third, these products were summed over the responding schools (including replacement schools). Finally, the ratio of these values was the response rate.

As in the previous methods, the numerator of the school response rate is the denominator of the student response rate:

$$RR_3^{st} = \frac{\sum_i^{s+r1+r2}\left(BW_i \times \sum_j^{r_i}(BW_{ij})\right)}{\sum_i^{s+r1+r2}\left(BW_i \times \sum_j^{r_i}(FW_{ij})\right)}$$

The overall response rate is the product of the school and student response rates:

$$RR_3 = RR_3^{sc} \times RR_3^{st}$$

Weighted response rates excluding replacement schools

Practically, replacement schools were excluded by setting their school base weight to zero and applying the same computations as above. More formally, the parts of the response rates are computed as follows:

$$RR_4^{sc} = \frac{\sum_i^{s}\left(BW_i \times \sum_j^{r_i}(FW_{ij})\right)}{\sum_i^{s+r1+r2}\left(FW_i \times \sum_j^{r_i}(FW_{ij})\right)}$$

$$RR_4^{st} = \frac{\sum_i^{s}\left(BW_i \times \sum_j^{r_i}(BW_{ij})\right)}{\sum_i^{s}\left(BW_i \times \sum_j^{r_i}(FW_{ij})\right)}$$

$$RR_4 = RR_4^{sc} \times RR_4^{st}$$

Reported response rates

The Australian unweighted school response rate in Year 6 was 97 per cent when including replacement schools and 96 per cent when excluding replacement schools. In Year 10, the respective percentages were 88 per cent and 87 per cent. The lower response rates for Year 10 are mainly explained by the low response rates of the Northern Territory and especially Tasmania, where a significant portion of Government Schools were not contacted in time to conduct the assessment.

Overall unweighted response rates including replacement schools were 87 per cent for Year 6 and 71 per cent for Year 10.

In terms of the coverage of the sampled population, weighted response rates are a more accurate indicator of the representativeness of the sample. In this regard, the overall response rate for Year 6 is 89 per cent when replacement schools are included and 88 per cent if only sampled schools are included. For Year 10, the numbers are 76 per cent and 75 per cent respectively.

Table 3.5 and Table 3.6 detail Year 6 and Year 10 response rates according to the four methods described above.

Table 3.5: Overall school and student response rates in Year 6

	Unweighted, including Substitute Schools			Unweighted, excluding Substitute Schools			Weighted, including Substitute Schools			Weighted, excluding Substitute Schools		
	Overall	School	Student	Overall	School	Student	Overall	School	Student	Overall	School	Student
NSW	0.88	0.98	0.90	0.88	0.98	0.90	0.89	0.98	0.90	0.89	0.98	0.90
Vic.	0.91	1.00	0.91	0.87	0.96	0.91	0.91	1.00	0.91	0.87	0.96	0.91
QLD	0.90	1.00	0.90	0.90	1.00	0.90	0.90	1.00	0.90	0.90	1.00	0.90
SA	0.85	0.96	0.89	0.81	0.91	0.89	0.85	0.96	0.89	0.82	0.92	0.89
WA	0.92	1.00	0.92	0.90	0.98	0.92	0.92	1.00	0.92	0.90	0.98	0.92
Tas.	0.89	1.00	0.89	0.89	1.00	0.89	0.89	1.00	0.89	0.89	1.00	0.89
NT	0.67	0.77	0.88	0.67	0.77	0.88	0.68	0.77	0.88	0.68	0.77	0.88
ACT	0.87	1.00	0.87	0.87	1.00	0.87	0.87	1.00	0.87	0.87	1.00	0.87
Aust.	0.87	0.97	0.90	0.86	0.96	0.90	0.89	0.99	0.90	0.88	0.97	0.90

Table 3.6: Overall school and student response rates in Year 10

	Unweighted, including Substitute Schools			Unweighted, excluding Substitute Schools			Weighted, including Substitute Schools			Weighted, excluding Substitute Schools		
	Overall	School	Student	Overall	School	Student	Overall	School	Student	Overall	School	Student
NSW	0.78	0.96	0.82	0.78	0.96	0.82	0.78	0.96	0.82	0.78	0.96	0.82
Vic.	0.76	0.96	0.79	0.76	0.96	0.79	0.76	0.96	0.80	0.76	0.96	0.80
QLD	0.79	0.96	0.83	0.76	0.91	0.83	0.79	0.96	0.83	0.76	0.91	0.83
SA	0.72	0.91	0.79	0.72	0.91	0.79	0.72	0.91	0.79	0.72	0.91	0.79
WA	0.81	0.95	0.84	0.79	0.93	0.85	0.80	0.95	0.84	0.79	0.93	0.85
Tas.	0.51	0.69	0.75	0.51	0.69	0.75	0.51	0.67	0.75	0.51	0.67	0.75
NT	0.39	0.54	0.73	0.39	0.54	0.73	0.36	0.50	0.73	0.36	0.50	0.73
ACT	0.76	0.93	0.82	0.76	0.93	0.82	0.76	0.93	0.82	0.76	0.93	0.82
Aust.	0.71	0.88	0.80	0.70	0.87	0.80	0.76	0.94	0.81	0.75	0.93	0.81

Chapter 4 DATA COLLECTION AND PROCESSING

The implementation of rigorous and quality-assured data collection and processing is crucial to the overall quality and reliability of the resulting data set. Over the course of many NAP sample cycles, ACER has refined these procedures in order to ensure that data collection practices are intuitive, well-designed and standardised across all aspects of field administration and that the data management processes implemented are rigorous, comprehensive and secure.

This chapter outlines the data management procedures implemented for NAP–CC 2019. This includes the various methods of data collection that were employed before, during and after the administration of the assessment, as well as the procedures applied in the transfer, tracking, verification and transformation of the data collected.

DATA MANAGEMENT PLAN

In line with best practice project management methodology, ACER created a detailed data management plan for the collection, transfer, processing and storage of data for the NAP–CC project. The plan firstly identified the data elements, or information assets, that were relevant to NAP–CC. It then detailed where each of the information assets were stored and described how they were to be secured over the life of the project. This plan was referred to and, where necessary, updated over the course of the project so that it would accurately describe the most current NAP–CC data management practices implemented by the project team.

DATA SECURITY

In the context of collecting, transferring and storing school- and student-level data, it is important to ensure that all systems, staff and processes are handling those information assets securely for the life of the project. Given that many of the NAP–CC information assets contained a level of Personally Identifiable Data of Australian school children, all assets were marked as protected in accordance with both ACER's Data Classification Policy and its Cryptographic Policy.

In addition, the team at ACER ensured that all policies and procedures implemented in the conduct of NAP–CC complied with the following three standards:

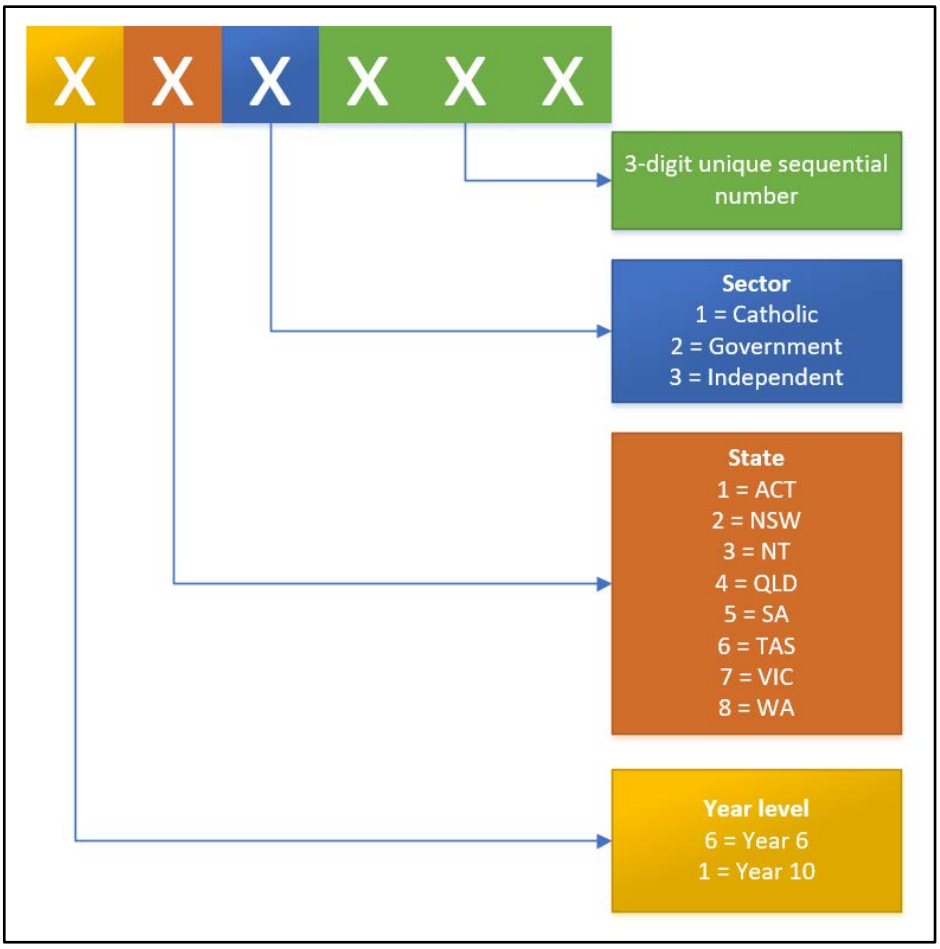
- ISO 27002:2015 Information technology - Security techniques - Code of practice for information security controls
- The Australian Government Information Security Manual (ISM) produced by the Australian Signals Directorate, and
- The Australian Government Protective Security Policy Framework.

DATA IDENTIFICATION

In order to track and monitor data throughout the life of the NAP–CC project, a system of identification (ID) codes was implemented. At the school level, a unique ID was created for each school at the time the sample was finalised. This school ID was six digits in length and comprised a concatenation of codes relating to year level, state, sector as well as a unique sequential number.

The specific codes used for each variable are outlined in Figure 4.1

Figure 4.1: Breakdown of codes used in unique school ID



At a student level, an ID was created that comprised the 6-digit school ID followed by a two-digit student number (01–20) that was unique to each sampled student within the school. This student ID was included in the student cognitive, contextual and student background data files so that data could be accurately matched and tracked throughout the data capture, cleaning and analysis stages.

Five spare IDs were created for each school and were distributed if additional test login credentials were required. The spare ID comprised the 6-digit school ID followed by a two-digit student number (21–25).

DATA COLLECTION FROM SCHOOLS AND JURISDICTIONS

As the NAP–CC assessment is administered to students within Australian schools, the contribution of both educational authorities and school staff in the data collection process is an essential part of the field administration.

In the lead up to the administration of the NAP–CC assessment, several stages of school liaison were necessary to collect school- and student-level information that would ensure the smooth running of the assessment on the scheduled date. Key personnel at each of the schools were nominated by the principal so that administrative and technical information could be collected in a timely manner. The roles of these nominated school personnel are outlined below:

The School Contact (SC). The SC was the main point of contact for ACER at the school and was responsible for coordinating and overseeing the assessment. SCs provided ACER with information about the school's preferred assessment dates, student cohort lists and, if this could not be provided by the jurisdiction, student background data for the selected students.

The School Technical Support Officer (STSO). The STSO was responsible for ensuring that the school's computer system was 'test ready' by the scheduled assessment date. Primarily, the role involved conducting a series of technical checks on a sample of computers that were to be used for the assessment and helping to troubleshoot any issues ahead of assessment day.

The Test Administrator (TA). The TA was responsible for administering the assessment to the sampled students, according to the standardised administration procedures provided in the TA Manual. The SC at the school would often also perform the duties of TA, though they could also choose to nominate another staff member for this role.

An overview of the school liaison and data collection processes is provided in Table 4.1**.**

Table 4.1: School liaison and data collection processes

Stage	Jurisdictional Activity	ACER Project Team Activity	School Activity
1	Educational authorities inform sampled schools of their selection in the assessment. If the jurisdiction confirms that a sampled school is unable to participate, the relevant replacement school is contacted	ACER contacts principals of sampled schools to request the nomination of a school contact person and school technical support officer	Principals of contacted schools supply requested contact information via secure online form
2		ACER contacts nominated School Contacts and requests preferred assessment dates and student lists for target year level (either Year 6 or Year 10 cohort)	School Contacts submit preferred assessment dates and student list via School Administration Website
3		ACER contacts nominated School Technical Support Officers (STSOs) and provides technical check instructions. ACER provides technical support and troubleshooting advice to STSOs via the Helpdesk	STSOs undertake technical checks to ensure the school's computer resources are test-ready
4		ACER notifies School Contacts of finalised assessment date and selected students via the School Administration Website	School Contact makes relevant school-level test day arrangements (including room bookings and informing sampled students of their selection)
5	Educational authorities provide SBD for students in schools for which this information is held centrally	Where SBD cannot be provided by the jurisdiction, ACER requests this information from School Contacts for all sampled students	School Contacts provide SBD for all sampled students via the School Administration Website
6		ACER provides detailed test administration manual and test login credentials to all nominated Test Administrators. ACER continues to provide support to schools via the Helpdesk	Test Administrators familiarise themselves with the processes and procedures outlined in the test administration manual and consult with ACER Helpdesk staff to confirm understanding of protocol and circumvent any perceived issues prior to the scheduled assessment date.

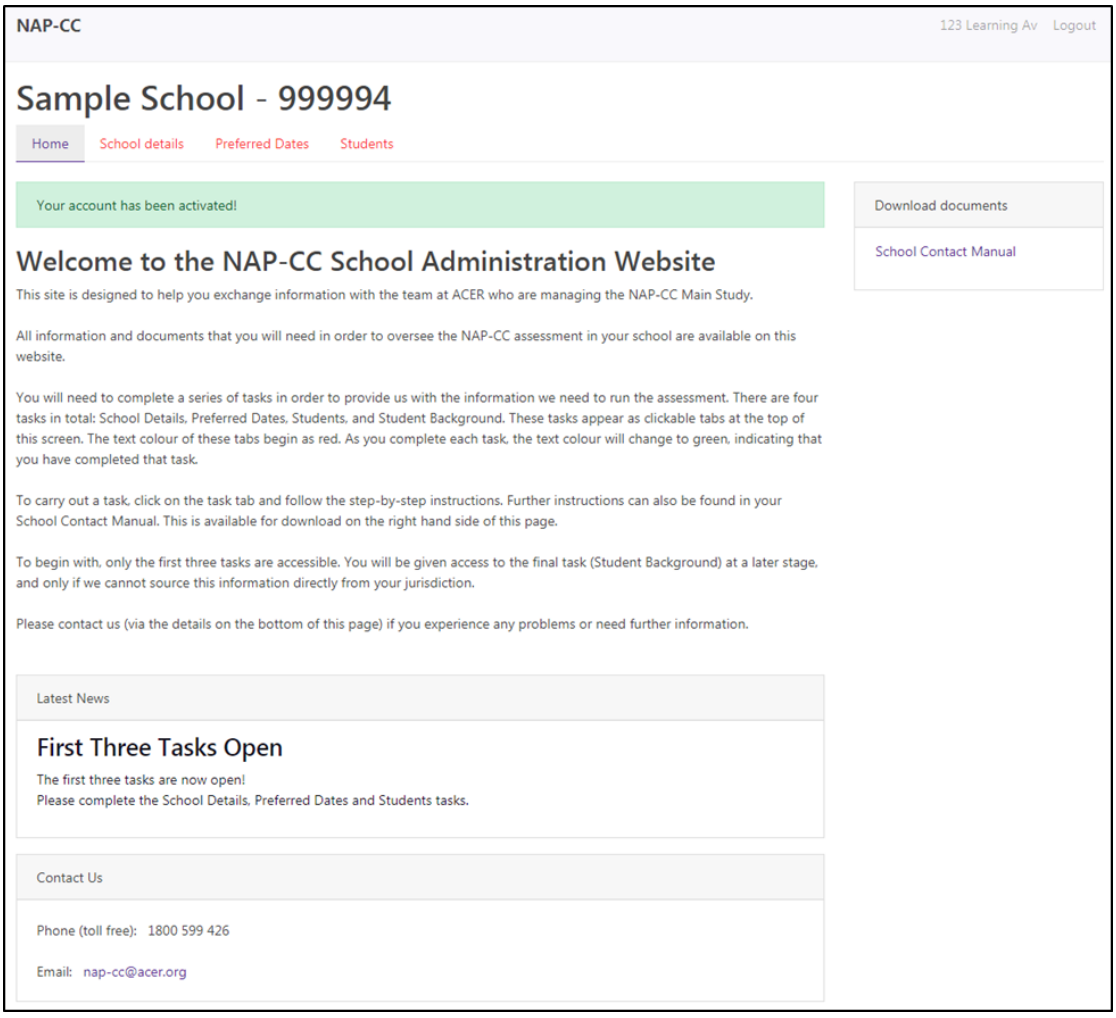
The NAP–CC online school administration website

All information provided by SCs to ACER was submitted via a secure website. The benefits of the NAP–CC online school administration website were two-fold: it eased the administrative burden on the selected schools, as well as providing a convenient, intuitive and secure repository for all school data relating to the study.

To access the website, SCs needed to create a secure password and activate their school-specific account. Once their account was activated, they were able to download all relevant administrative materials from this site, as well as use it to provide information to ACER regarding school contact details, assessment date preferences, and student-related information as required.

Figure 4.2 shows a screenshot from the homepage of the website.

Figure 4.2: NAP-CC online school administration website



The STSO technical checks

To ensure the smooth running of the assessment, it was necessary for STSOs to perform a series of technical checks on the computers that were selected for use.

An excerpt from the STSO Handbook detailing the technical checks STSOs were required to perform is provided in Appendix A2.

After the technical checks were undertaken, the ACER Project Team liaised with any STSOs who had reported issues. Technical issues were resolved through a process of troubleshooting with the ACER Project Team. This sometimes involved referring the matter to the test delivery system engineers or, in the case of access/security protocols, to the relevant central education authority of the applicable school.

Helpdesk provision and online support

An 1800 helpdesk support number and a dedicated email address were made available to schools for the entire Main Study administration phase (July – December 2019). Using these means, the ACER Project Team supported schools through all administrative, technical and operational tasks related to the administration of the NAP–CC assessment. Project staff were also on hand to provide any urgent assistance required during, or immediately preceding, the assessment session itself.

The helpdesk hours of operation during the assessment window were 8am-6pm AEST so that school hours across Australia's various time zones could be accommodated.

Collection of student background data

As per NAP protocol, student background data were collected for all participating students and matched to students' cognitive assessment and survey responses for analysis and reporting purposes.

The variables collected for participating students are set out in the Data Standards Manual (ACARA, 2019) and included:

- gender
- date of birth
- Indigenous status
- parents' school education
- parents' non-school education
- parents' occupation group
- students' and parents' home language
- geolocation of the students' school[5]

[5] Geolocation data was not collected for each student. Rather, location data of the school was sourced from the sampling frame and applied to all sampled students within a school.

Schools are required to collect this information from the time of student enrolment and the data are often held centrally by a school's educational authority. Where data were held centrally, ACER sought the student background data from the relevant educational authority so that schools were not unnecessarily burdened with this administrative task. This occurred in half (12 out of 24) of the jurisdictions across the country.

The source of student background data for each of the jurisdictions is outlined in Table 4.2.

Table 4.2: Student background data provision

State/Territory	Sector	Source
ACT	Govt	ACT DET
	Cath	ACT DET
	Ind	ACT DET
NSW	Govt	NSW DET
	Cath	School
	Ind	School
NT	Govt	NT DET
	Cath	School
	Ind	School
Qld	Govt	QLD DETE
	Cath	School
	Ind	School
SA	Govt	SA DECD
	Cath	SA CEO
	Ind	School
Tas	Govt	Tas DoE
	Cath	Tas CEO
	Ind	School
Vic	Govt	VIC DET
	Cath	School
	Ind	School
WA	Govt	WA DET
	Cath	School
	Ind	School

Where central data collection was not possible, ACER collected this information from the schools themselves. To do this, the ACER Project Team created a template into which schools could enter the coded background details for each sampled student. This template was then uploaded by each school onto the secure NAP–CC School Administration Website.

The code list for the student background data collected is presented in Table 4.3.

Table 4.3: Variable definitions for Student Background Data

Category	Description	Codes
Gender	Gender of student	F = female M = male
Date of Birth	Date of birth of student	Free response DD-MMM-YYYY
Indigenous Status	A student is considered to be 'Indigenous' if he or she identifies as being of Aboriginal and/or Torres Strait Islander (ATSI) origin.	1 = Aboriginal but not TSI origin; 2 = TSI but not Aboriginal origin; 3 = Both Aboriginal and TSI origin; 4 = Neither Aboriginal nor TSI origin; 9 = Not stated/unknown.
Parent School Education	The highest year of primary or secondary education a parent/guardian has completed.	1 = Year 9 or below; 2 = Year 10; 3 = Year 11; 4 = Year 12; 0 = Not stated/unknown/Does not have Parent 1.
Parent Non-School Education	The highest qualification attained by a parent/guardian in any area of study other than school education.	5 = Certificate I to IV (including Trade Certificate); 6 = Advanced Diploma/Diploma; 7 = Bachelor Degree or above; 8 = No non-school qualification; 0 = Not stated/unknown/Does not have Parent 1.
Parent Occupation Group	The occupation group which includes the main work undertaken by the parent/guardian.	1 = Senior management; professionals; 2 = Other management; associate professionals; 3 = Tradespeople; skilled office, sales and service; 4 = Unskilled workers; hospitality; 8 = Not in paid work in last 12 months; 9 = Not stated/unknown/Does not have Parent 1.
Student / Parent home language	The main language spoken in the home by the respondent.	1201 = English; Codes for all other languages as per the Australian Standard Classification of Languages (ASCI) Coding Index 2nd Edition

The ability of the ACER Project Team to collect student background data to the level required for data analysis purposes depends on how complete the records are kept at participating schools and central authorities. Where data variables were labelled as unknown or left blank by the school or jurisdiction, and the absence of data was confirmed upon follow up from the project team, these values were coded as missing. The percentage of missing values for the derived

background data variables, along with the percentages for all valid codes, are presented in the national report.

ASSESSMENT ADMINISTRATION

The NAP–CC 2019 assessment was conducted within a two-week window toward the beginning of Term 4 in each of the participating schools. The test window for each state and territory is outlined below:

- Qld, Vic.: Monday 14 October – Friday 25 October 2019
- ACT, NSW, NT, SA, Tas.[6] & WA: Monday 21 October – Friday 1 November 2019

Schools generally undertook the test session on one day within the testing window, though a small number nominated to run the test with smaller groups of students over several days for logistical or technical reasons.

Data capture

Student cognitive and survey data were captured via the Online National Assessment Platform program using the Locked Down Browser installed on school or student computers.

As all the student survey and achievement data were collected electronically, scanning and manual data entry of student responses were not required.

Student test experience

The NAP–CC assessment comprised a single test session of 60 minutes for Year 6 students, and 75 minutes for Year 10 students. The entire assessment administration time was no more than two hours in total. This two-hour period included time for settling the students into the test room, logging students into the assessment platform, reading the instructions to the students, administering the test and conducting a student survey.

Follow-up test sessions

If attendance on the scheduled day of assessment fell below 80 per cent, schools were asked to schedule a follow-up session later within the testing window with as many of the absent students as possible. To maximise participation for the follow-up sessions, an additional testing week was added to the original assessment window for schools in all states and territories. The conduct of follow up sessions in this way ensured a participation rate of at least 80 percent in most schools administering the NAP–CC assessment.

[6] Issues with test administration meant that the test window for Government Year 10 schools in Tasmania was delayed. The testing window for these 16 schools was 25/11/2019–6/12/2019.

Quality monitor visits

In order to document the quality and uniformity of the administrative procedures undertaken, a random selection of five per cent of schools across all sectors and jurisdictions were visited by National Quality Monitors on the scheduled day of the assessment. Selected schools were notified of the Quality Monitor's visit before the scheduled assessment day so that appropriate permissions could be obtained for the Quality Monitor's admission to the school.

National Quality Monitors were trained by the ACER project team in all aspects of test administration procedures and NAP-protocol prior to their deployment in schools. Their responsibility was to observe and record whether tasks in the procedural manuals were followed during the assessment session and to report their findings to the ACER project team via the completion of a structured online Quality Monitor Report.

In total, 35 schools from both year levels and a range of jurisdictions across Australia were visited by Quality Monitors.

DATA CLEANING AND VERIFICATION

Data cleaning and verification relate to processes of ensuring the integrity of the data collected. For NAP–CC, a series of data cleaning steps was undertaken on all data collected from jurisdictions, schools and students. With respect to student background data, the following steps were performed:

Student names (for the purposes of school reporting) were corrected where there was obvious first name/surname reversal, or where foreign characters (e.g. ?, !, %) were included. Some instances of correction had to be confirmed with the school directly.

Missing gender of the student was attributed where it could be inferred from the school type (e.g. where single-sex) or name of the student. Some instances of correction had to be confirmed with the school directly.

All dates of birth were converted to the standard dd/mm/yyyy format, and any auto-formatting executed by the spreadsheet template that rendered dates of birth illegible was reversed and corrected.

Any free text or abbreviated text was coded as per the variable coding schema presented in Table 4.3.

Any out of range, implausible or missing values were double-checked with the school or jurisdiction that provided the data. Where possible, the correct values were inputted. Where no further information was provided or available, the data were recoded to missing.

With respect to the student cognitive and survey data, the following preliminary data cleaning steps were performed:

Instances of invalid IDs were investigated and, after liaison with the test administration team, corrected where possible or else removed from the dataset.

Instances of spare IDs were matched with valid Student IDs and recoded accordingly. This often necessitated confirmation and cross-checking with the attendance roll data and notes from the test administration team.

Patterns of missing values were explored and, where appropriate, recoded to '9' for embedded missing, 'r' for not reached (cognitive data only) or 'n' for not administered.

Further information regarding the scaling procedures implemented for the cognitive achievement data and student survey data can be found in Chapter 5 of this report.

DATA TRANSFORMATION

With respect to the student background data collected, variables were also derived for the purposes of reporting achievement outcomes. The transformations undertaken by the analysis team followed the data rules outlined in ACARA's most recent version of the Data Standards Manual (ACARA 2019).

Table 4.4 shows the derived variables and the transformation rules used to recode them.

Table 4.4: Transformation rules used to derive student background variables for reporting

Variable	Name	Transformation rule
School location	ASGSRemote	The geographical classification of the school location according to the ABS remoteness classification (1= major cities, 2 = inner regional, 3 = outer regional, 4 = remote, 5 = very remote).
Gender	GENDER	Classified by response; missing data treated as missing unless the student was present at a single-sex school or unless deduced from student name.
Age	AGE	Derived from the difference between the Date of Assessment and the Date of Birth, transformed to whole years.
Indigenous Status	INDIG	Coded as Indigenous (1) if response was 'yes' to Aboriginal, OR Torres Strait Islander OR Both. Coded as non-Indigenous (0) otherwise.
LBOTE	LBOTE	Each of the three LBOTE questions (Student, Mother or Father) were recoded to 'LBOTE' (1) or 'Not LBOTE' (0) according to ASCL codes. The reporting variable (LBOTE) was coded as 'LBOTE' (1) if response was 'LBOTE' for any of Student, Mother or Father. If all three responses were 'Not LBOTE' then the LBOTE variable was designated as 'Not LBOTE' (0). If any of the data were missing then the data from the other questions were used. If all of the data were missing then LBOTE was coded as missing.
Parental Education	PARED	Parental Education equalled the highest education level (of either parent). Where one parent had missing data the highest education level of the other parent was used. Only if parental education data for both parents were missing, would Parental Education be coded as 'Missing' (0).
Parental Occupation	POCC	Parental Occupation equalled the highest occupation group (of either parent). Where one parent had missing data or was classified as 'Not in paid work', the occupation group of the other parent was used. Where one parent had missing data and the other was classified as 'Not in paid work', Parental Occupation equalled 'Not in paid work'. Only if parental occupation data for both parents were missing, would Parental Occupation be coded as 'Missing' (9).

MARKING OF STUDENT RESPONSES

In order to analyse the cognitive achievement data collected from participating students, responses were first scored appropriately. Depending on the nature and structure of an item, student responses were scored either automatically by the assessment system or, where extended text responses (constructed responses) were extracted, saved and marked later by groups of trained markers in a central marking location.

The following sections detail the various marking processes and quality control measures implemented during the marking operation.

Automated marking

Items that did not elicit open-ended responses from participating students were automatically scored as correct or incorrect by the assessment platform. These item types belonged to one of the following three response categories:

- **standard multiple choice**, for which students were asked to select the best answer from a list of typically four distinct options. For the purpose of analysis, the selection made by the student was recorded by the assessment system and scored as correct (one score point) or incorrect (zero score points).
- **multiple choices response**, for which students were asked to select *all* possible answers from a list of four or more distinct options. To receive one score point, all the selections a student made had to be correct. If all selections were not made correctly, the score awarded was zero.
- **interactive match**, which required students to provide their response to an item by using 'drag and drop' or hotspot functions. The selections made by the student were recorded in the system and to receive one score point, all selections (maximum of three) had to be correct.

As a quality control measure, students' raw responses for these items were also extracted from the system and compared to the item key in the codebook to ensure there were no anomalies with the automated scoring algorithm.

Scores of all three item types were reported in the school reports. The raw responses of standard multiple-choice and multiple choices items and the score of interactive match item were used in the psychometric analysis.

Marking of constructed response items

Items that required students to provide typed responses (ranging from one word to a maximum of three sentences) were saved by the assessment system and marked at a later date by a team of trained, human markers. This marking operation was conducted over a period of two weeks in November 2019[7]. Marking was based in the ACER Sydney Marking Centre and the ACER Marking System (AMS) was used as the marking platform.

In total, 36 markers were recruited with almost all having prior experience in marking NAP–CC responses (either from the Field Trial in June 2019, and/or from previous NAP–CC cycles).

Markers were organised into four groups, with each group overseen by an experienced Group Leader who reported to the Chief Marker. Each group of markers was trained by the Chief Marker on one item at a time, with the entire pool of responses for that item being marked before training in the next item

[7] The exception to this was in regard to the responses collected from some Tasmanian Year 10 students. Assessment administration for Year 10 Government schools in Tasmania took place in early December, and a supplementary marking operation was conducted at this time in order to mark those students' constructed responses. Data were extracted and supplied to the analysis team on 16 December 2019.

commenced. This train-mark, train-mark model meant that markers were able to focus on a single item at a time, making it easier to internalise the marking rubric and recall scoring criteria so that markers were able to mark a large set of data rapidly and accurately. Sample student answers for an item were given to the markers prior to marking that item. The scores for the sample answers were discussed and the scoring categories clarified before marking commenced in each instance.

As an important quality control measure, both warm-up scripts and control scripts were deployed to all markers for each item. These warm-up and control scripts were pre-selected and given a 'true score' by the Chief Marker before being assigned to each of the markers in their response pool. If a marker gave a score that was inconsistent with the score given by the Chief Marker, the scoring criteria were clarified with the marker before marking resumed.

In general, the results from these quality control measures were highly accurate with overall discrepancy between markers and the set controls being less than seven per cent. Only one item elicited an unacceptable compliance rate which meant that, as per NAP–CC marking protocol, the initial ratings of all markers were eliminated. Markers were then retrained and further discussion of the types of responses that were being observed by the markers was conducted. All responses for that item were then remarked and a satisfactory compliance rate was achieved.

In addition to the use of warm-up and control scripts, spot checking was performed as an ongoing quality control measure for the duration of the marking operation. For each item marked, approximately 10 per cent of responses were spot checked (i.e. marked again) by the designated Group Leader or the Chief Marker. The spot-checking process provided another opportunity to identify when items were being marked inconsistently, either by the whole group or by an individual marker. If inconsistent marking was identified, the markers were retrained on the specific item and the responses were re-marked.

Finally, to ensure the consistent application of marking criteria between the 2019 NAP–CC cycle and the previous cycle in 2016, a reliability check was undertaken on 12 of the items common to both assessments. A total of 595 scripts were blind marked by the marker pool. For nine of the 12 items marked, inter-cycle reliability in excess of 90 per cent was found, with the three remaining items eliciting lower reliability rates of 83 per cent, 79 per cent and 76 per cent, all of which are acceptable reliabilities. In all cases, the discrepancies between the 2016 and 2019 ratings were investigated by the Chief Marker. Whilst some instances of aberrant application of the rubric by the 2019 markers were found, in the majority of cases, the ratings provided in 2019 were consistent with the discussions and applications of the rubrics implemented in 2019.

Further information about the development of the scoring guides for constructed response items, including an example of an item and its scoring criteria, is provided in Chapter 2.

DATA PROCESSING FOR SCHOOL REPORTING

Once all student responses were marked, the following data processing steps were implemented in order to produce the summary reports that were distributed to the participating schools:

- Collation of all marked student data and creation of a single data file for each year level
- Removal of introductory practice items for each student and separation of student survey data (which was not included in the analysis for school summary reports)
- Checking of the student response data file against the codebook to ensure no major data anomalies
- Computation of item per cent correct, weighted by preliminary student weights and excluding not reached responses
- For partial credit items, computation of item per cent correct for each item in standard NAP sample format (e.g. 75,23 where 0,1,2 item becomes 75 (facility of 1 and 2), 23 (facility of 2 only))
- Formatting of data file to required specifications for import into the ACER Online Assessment and Reporting System (OARS).

After all student test data underwent the data processing steps, the final data set was imported into ACER OARS to create and distribute the online summary reports to participating schools.

School summary reports

The NAP–CC 2019 School Summary Reports provided schools with information about the specific items each student was administered, the level of credit each student received for every item they were administered, and the weighted proportion of students who received different levels of credit for each item. The reports were interactive in that users could filter and sort data to view information grouped by categories of interest, such as by student gender or item type. Furthermore, the reports were password-protected so that only the designated School Contact person could access them on the OARS platform and could then disseminate to other staff and/or students in line with their school's specific policy in this regard.

Whilst preliminary student weights were applied for the per cent correct analysis, scaled scores were not provided in the school reports. Provision of weighted, unscaled scores to schools is in line with school reporting protocol for other NAP sample assessments due to the rapid turnaround of reports for participating schools.

Appendix A3 provides the instructional guide that was sent to School Contacts at participating schools. The guide outlined how schools were to access the NAP–

CC reports via ACER OARS and provided guidance to school staff on how to interpret the contents.

Chapter 5 SCALING PROCEDURES

Both cognitive and survey items were scaled using item response theory (IRT) scaling methodology. The cognitive items were used to derive a one-dimensional NAP–CC achievement scale, while a number of scales were constructed based on different sets of survey items. This chapter outlines the procedures implemented to create these scales as well as providing a description of the associated processes of DIF analysis, item calibration, horizontal equating and the creation of plausible values.

THE SCALING MODEL

Test items were scaled with the one-parameter model (Rasch, 1960). In the case of dichotomous items, the model predicts the probability of selecting a correct response (value of one) instead of an incorrect response (value of zero), and is modelled as:

$$P_i(\theta_n) = \frac{\exp(\theta_n - \delta_i)}{1 + \exp(\theta_n - \delta_i)}$$

where $P_i\ (\theta_n)$ is the probability of person n scoring 1 on item i, θ_n is the estimated ability of person n, and δ_i is the estimated location of item i on this dimension. For each item, item responses are modelled as a function of the latent trait θn.

For items with more than two (k) categories (Likert-type items, for instance), the more general Rasch partial credit model (Masters & Wright, 1997) was applied, which takes the form of:

$$P_{x_i}(\theta_n) = \frac{\exp \sum_{k=0}^{x} (\theta_n - \delta_i + \tau_{ik})}{\sum_{h=0}^{m_i} \exp \sum_{k=0}^{h} (\theta_n - \delta_i + \tau_{ik})} \qquad x_i = 0,1,\text{K}\ ,m_i$$

where $P_{xi}\ (\theta_n)$ denotes the probability of person n scoring x on item i, θ_n denotes the person's ability, the item parameter δ_i gives the location of the item on the latent continuum, and τ_{ik} denotes an additional step parameter for each step k between adjacent categories.

The analysis of item characteristics and the estimation of model parameters were carried out with the ACER ConQuest software package (Version 5 software: see Adams, Wu, Macaskill, Haldane, Sun & Cloney, 2020).

SCALING COGNITIVE ITEMS

This section outlines the procedures for analysing and scaling the cognitive test items. The procedures are somewhat different from scaling the survey items, which are discussed later in the chapter.

The model fit of cognitive test items was assessed using a range of item statistics. The weighted mean-square statistic (infit), which is a residual-based fit statistic, was used as a global indicator of item fit. Infit statistics were reviewed both for item and step parameters.

In addition to this, item characteristic curves (ICCs) were also used to review item fit. ICCs provide a graphical representation of item fit across the range of student abilities for each item (including dichotomous and partial credit items). The functioning of the partial credit score guides was further analysed by reviewing the proportion of responses in each response category and the correct ordering of mean abilities of students across response categories.

Final decisions on removing test items were based on a range of different criteria. Generally, items were flagged for review if first item calibrations showed a considerably higher infit statistic (e.g. infit > 1.2) as well as low item-rest correlation (0.2 or lower). The ACER project team considered both item-fit criteria as well as the content of the item prior to a decision about removing or retaining flagged items for scaling.

Of the 179 items in the test, two were removed from the scale due to poor fit statistics at both year levels (CG_63 and HP_61). In addition, one item was removed at Year 6 only (AF31) and another three were removed at Year 10 only (CL_61, CS_61 and TC_63). Consequently, these items were not used to estimate student achievement.

DIFFERENTIAL ITEM FUNCTIONING

The quality of the items was also explored by assessing differential item functioning (DIF) by gender. DIF occurs when groups of students with the same ability have different probabilities of responding correctly to an item. For example, if boys have a higher probability of success than girls with the same ability on an item, the item shows DIF in favour of boys. This constitutes a violation of the model, which assumes that the probability is only a function of ability and not of any other variable. Substantial item DIF with respect to gender may result in bias of performance estimates across gender groups. No instances of substantial gender DIF were encountered so no items were removed for this reason.

ITEM CALIBRATION

Missing student responses, likely caused by issues with test length (‘not reached’ items)[8], were omitted from the calibration of item parameters but were treated as incorrect for the scaling of student responses. All other missing responses were included as incorrect responses for the calibration of items (except for the ones that were not administered).

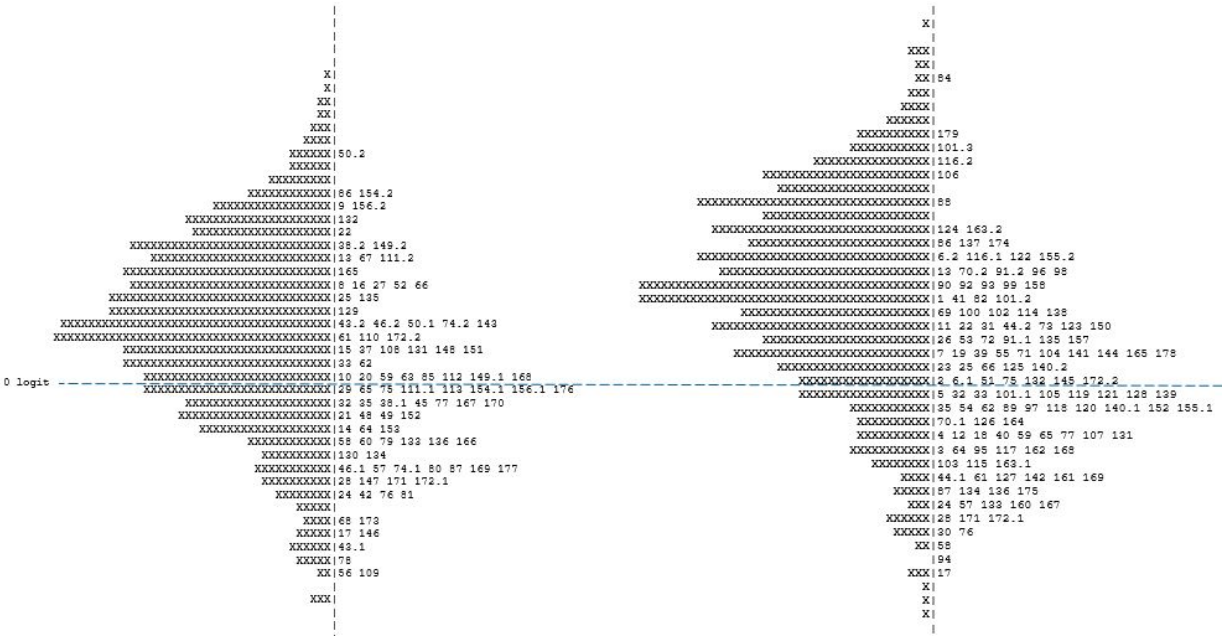
[8] ‘Not reached’ items were defined as all consecutive missing values at the end of the test except the first missing value of the missing series, which was coded as ‘embedded missing’, like other items that were presented to the student but which did not receive a response.

Item parameters were calibrated using all sampled student data, except for students who were identified as non-respondents. The student weights were rescaled to ensure that each state or territory was equally represented in the sample. The items were calibrated separately for Year 6 and Year 10. After removing items with unsatisfactory scaling characteristics, a total of 138 civics and citizenship items and 36 history items were used for scaling. Out of a total of 174 items, thirty-five items were administered to both year levels.

An investigation was conducted to look into the possibility of scaling the civics and citizenship items and the history items as a unidimensional scale. A two-dimensional model was created by year level. The correlation between the two dimensions was 0.98 for Year 6 and 0.96 for Year 10, which suggested a unidimensional scale at each year level. It was therefore decided to use a unidimensional model to scale the civics and citizenship items and history items together for each year level.

Figure 5.1 presents item maps for the two year levels. The crosses represent students, the numbers represent items, and in the case of a partial credit item the threshold is included. The vertical line represents the measured CC literacy scale with high-performing students and difficult items at the top and low-performing students and easy items at the bottom. The two scales are not directly comparable because they have been calibrated separately, but they have been lined up approximately for this report. The response probability in this figure is 0.5, which means that students with an ability equal to the difficulty (or threshold) of an item have a 50 per cent chance of responding correctly to that item. At each year level, the alignment of the student and item distributions in the figure shows that the test was well targeted for Year 6 and slightly easy for Year 10.

Figure 5.1: Item maps for Year 6 and Year 10



The overall reliability of the test, as obtained from the scaling model, was 0.84 for Year 6 and 0.87 for Year 10 (ACER ConQuest estimate). Appendix A4 shows the

item thresholds on the NAP–CC scale with a response probability of 0.50 and of 0.629 in logits and their equated reporting scale score. It also shows the respective percentages of correct responses for each year sample (giving equal weight to each jurisdiction). The weighted fit statistics are included in the last column and column three indicates if an item was used as a horizontal link (trend) item, see next section.

HORIZONTAL EQUATING

Test items at both year levels consisted of new and old items. The old items were developed for and used in previous cycles. As the items had been kept secure, they could be used as horizontal link items to equate the results of the 2019 assessment with the established NAP–CC scale.

To ensure that the link items had the same measurement properties across cycles, the relative difficulties in 2019 and 2016 were compared. Four out of 30 common items for Year 6 and two out of 40 common items for Year 10 showed large DIF between 2019 and 2016 and were not used for equating. For each year level, the set of link items showed similar average discrimination (item–rest correlation was 0.37 in 2016 and 0.36 in 2019 for Year 6 and 0.38 in 2016 and 0.40 in 2019 for Year 10). The average DIF with respect to gender in both cycles was also close to zero (0.03 logits in 2016 and 0.04 logits in 2019 for Year 6; and 0.02 logits in 2016 and 0.00 logits in 2019 for Year 10).

Figure 5.2 and Figure 5.3 show scatter plots of item difficulties for horizontal link items in 2016 and 2019 for Year 6 and Year 10, respectively. The average difficulty of each set of link items was set to zero and each dot represents one link item. The expected location under the assumption of complete measurement equivalence across both assessments is the identity line ($y = x$). The thick broken lines represent the 95 per cent confidence interval around the expected values and items outside of these lines had statistically significant deviations from the green identity line. The pink, broken line is the line of best fit between the item difficulties of the two cycles. The graphs show that the slope of this line is close to one.

The original standard errors provided by ACER ConQuest were adjusted by multiplying them by the square root of six, the approximate design effect in 2016. This correction was made because data were collected from a cluster sample design, whereas the scaling software assumes simple random sampling of data (see Chapter 3 for further information about sampling procedures).

Historical items were not used as link items if the difference between relative item difficulties was significant and more than 0.5 logits. Using this criterion, four items in Year 6 and two items in year 10 were excluded from equating.

[9] This means that a student with a scale score equal to the item difficulty parameters has 62% probability of giving a correct response to the test question.

Figure 5.2: Relative item difficulties in logits of Year 6 horizontal link items between 2016 and 2019

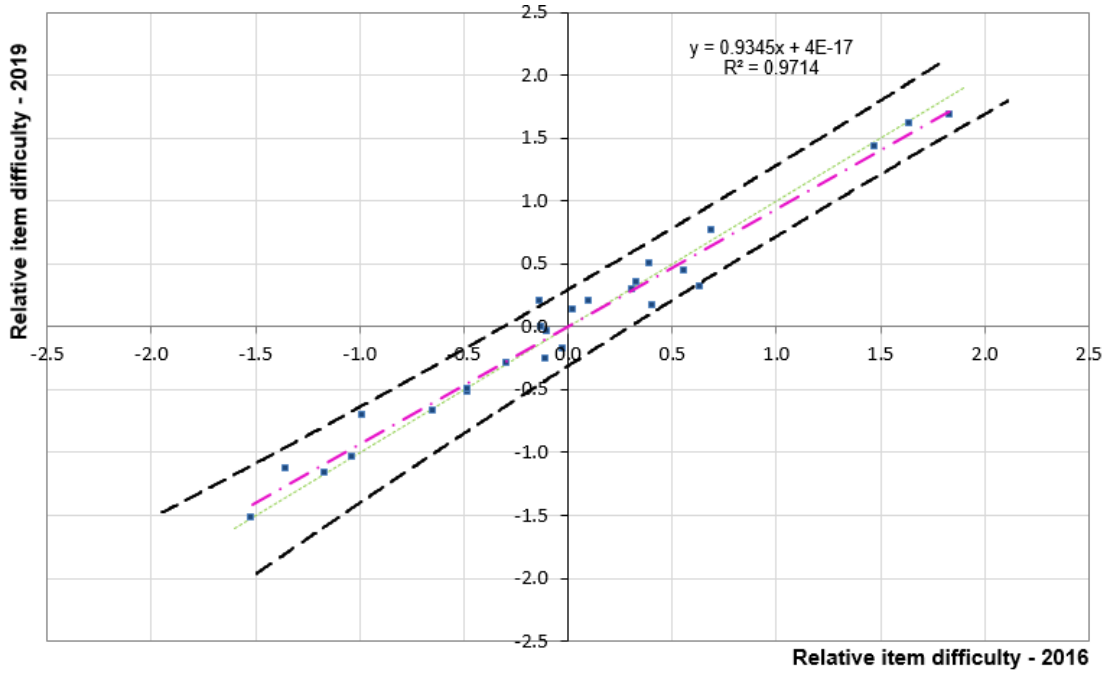
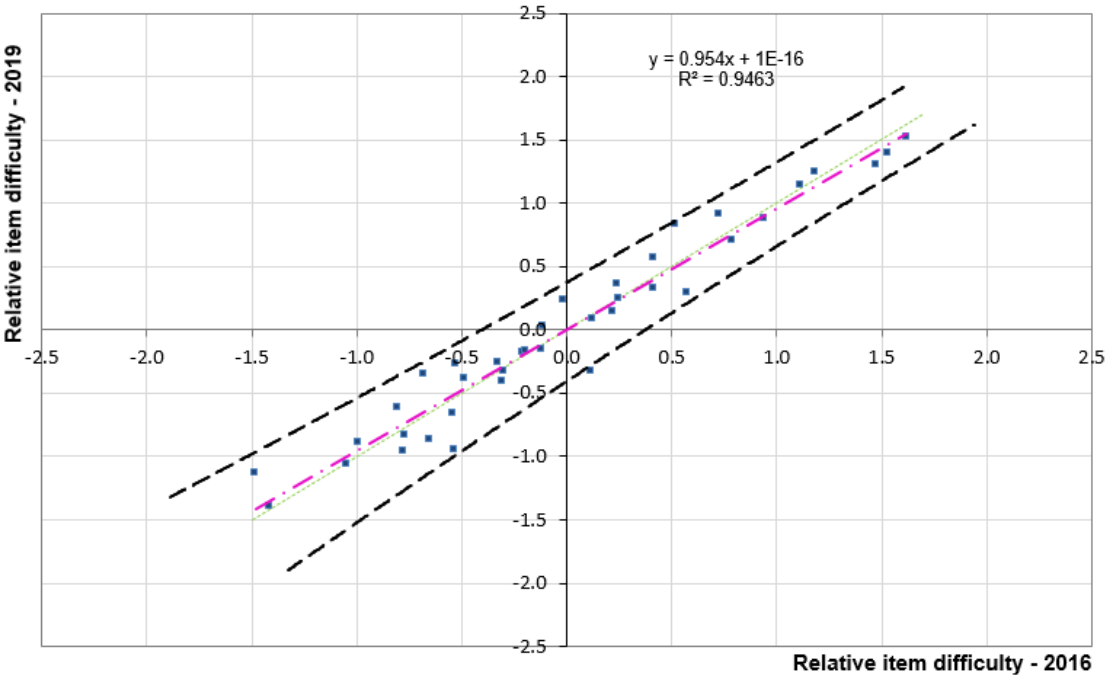


Figure 5.3: Relative item difficulties in logits of Year 10 horizontal link items between 2016 and 2019



Item–rest correlation is an index of item discrimination, which is computed as the correlation between the scored item and the raw score of all other items in a test form. It indicates how well an item discriminates between high- and low-performing students, similar to the item fit statistic. The 2016 and 2019 values of these discrimination indices are plotted in Figure 5.4 and Figure 5.5.

Figure 5.4: Discrimination of Year 6 link items in 2016 and 2019

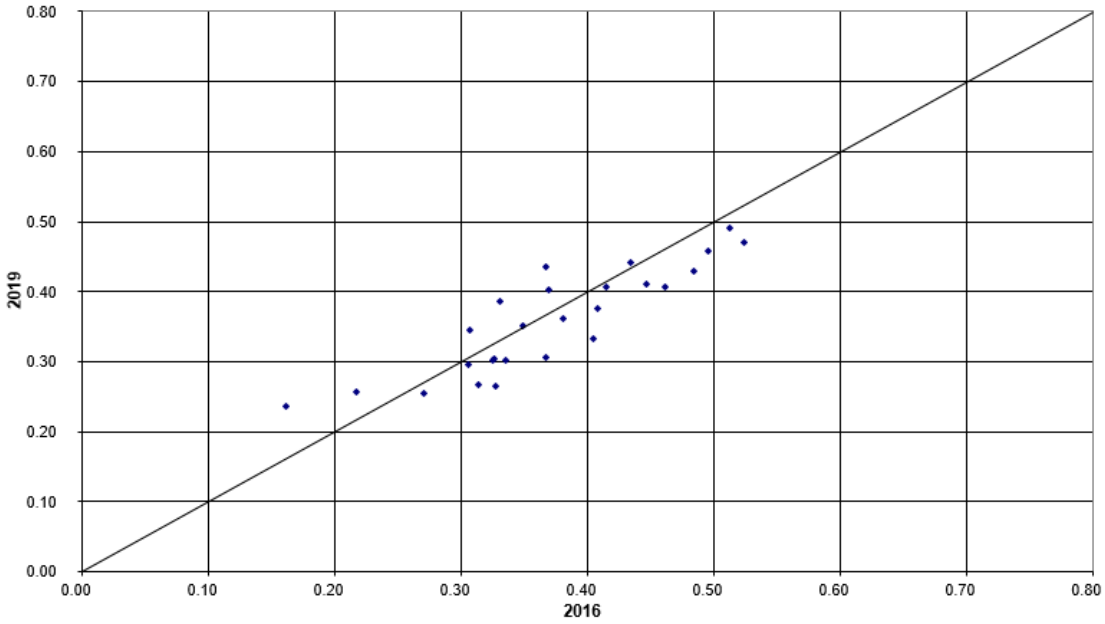
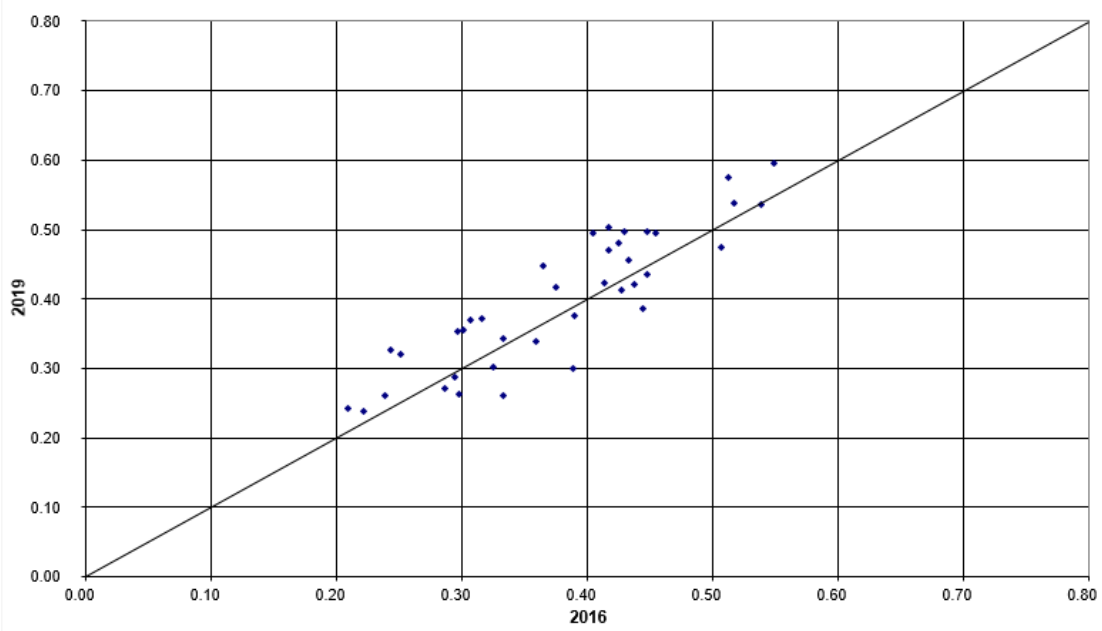


Figure 5.5: Discrimination of Year 10 link items in 2016 and 2019



After the selection of link items, common item equating was used to shift the 2019 scale onto the historical scale. The value of the shift is the difference in average difficulty of the link items between 2016 and 2019, 0.197 for Year 6 and 0.170 for Year 10). After applying this shift, the same transformation was applied as in 2016. The original scale scores (logits) for the Year 6 students were converted as:

$$\theta_n^* = \{(\theta_n + 0.197 - 0.193 - 0.063 - 0.473 - 0.547 - 0.189 - \bar{\theta}_{04})/\sigma_{04}\} \times 100 + 400$$

and for the Year 10 students:

$$\theta_n^* = \{(\theta_n + 0.170 - 0.168 - 0.208 - 0.777 - 0.057 + 0.119 - \bar{\theta}_{04})/\sigma_{04}\} \times 100 + 400$$

where θ_n^* is the transformed knowledge estimate for student *n*, θ_n is the original knowledge estimate for student *n* in logits, $\bar{\theta}_{04}$ is the mean ability in logits of the Year 6 students in 2004 (-0.6993), and σ_{04} is the standard deviation in logits of the Year 6 students in 2004 (0.7702).

Uncertainty in the link

The shift that equates the 2019 data with the 2016 data depends upon the change in difficulty of each of the individual link items. As a consequence, the sample of link items that have been chosen will influence the estimated shift. This means that the resulting shift could be slightly different if an alternative set of link items had been selected. As a result, there is an uncertainty associated with the equating that is due to the choice of link items, similar to the uncertainty associated with the sampling of schools and students.

The uncertainty that results from the selection of a sub-set of link items is referred to as a linking or equating error. This error should be taken into account when making comparisons between the results from different data collections across time. Just as with the error that is introduced through the process of sampling students, the exact magnitude of this equating error cannot be determined. We can, however, estimate the likely range of magnitudes for this error and take this error into account when interpreting results. As with sampling errors, the likely range of magnitude for the combined errors is represented as a standard error of each reported statistic.

The following approach has been used to estimate the equating error. Suppose we have a total of *L* score points in the link items in *K* modules. Use *i* to index items in a unit and *j* to index units so that $\hat{\delta}_{ij}^{y}$ is the estimated difficulty of item *i* in unit *j* for year *y*, and let:

$$C_{ij} = \hat{\delta}_{ij}^{2019} - \hat{\delta}_{ij}^{2016}$$

The size (number of score points) of unit *j* is m_j so that:

$$\sum_{j=1}^{K} m_j = L \quad \text{and} \quad \bar{m} = \frac{1}{K}\sum_{j=1}^{K} m_j$$

Further, let:

$$c_{\bullet j} = \frac{1}{m_j}\sum_{i=1}^{m_j} c_{ij} \quad \text{and} \quad \bar{c} = \frac{1}{N}\sum_{j=1}^{K}\sum_{i=1}^{m_j} c_{ij}$$

Then the link error, taking into account the clustering, is as follows:

$$LinkError_{2019,2016} = \sqrt{\frac{\sum_{j=1}^{k} m_j^2 (c_{.j} - \bar{c})^2}{K(K-1)\bar{m}^2}} = \frac{\sum_{j=1}^{k} m_j^2 (c_{.j} - \bar{c})^2}{L^2} \frac{K}{K-1}$$

The link error between 2016 and 2019 is 2.968 scale score points for Year 6 and 3.146 for Year 10. The equating error between 2019 and 2013 is the sum of the two equating errors between adjacent cycles for each year level. For example, the link error between 2019 and 2013 for Year 6 is:

$$error_{2019-2013} = \sqrt{2.968^2 + 4.424^2} = 5.327$$

The equating error between 2019 and 2010 is the square root of the sum of the three squared equating errors between the four cycles and the equating error between 2019 and 2007 is square root of the sum of the four squared equating errors between the five cycles.

$$error_{2019-2010} = \sqrt{2.968^2 + 4.424^2 + 4.848^2} = 7.203$$

$$error_{2019-2007} = \sqrt{2.968^2 + 4.424^2 + 4.848^2 + 5.28^2} = 8.931$$

PLAUSIBLE VALUES

Plausible values methodology was used to generate estimates of students' civic and citizenship achievement. Using item parameters anchored at their estimated values from the calibration process, plausible values were randomly drawn from the marginal posterior of the latent distribution (Mislevy, 1991; Mislevy & Sheehan, 1987; von Davier, Gonzalez, & Mislevy, 2009). Here, 'not reached' items were included as incorrect responses, just like other (embedded) missing responses. Estimations are based on the conditional item response model and the population model, which includes the regression on background and survey variables used for conditioning (see a detailed description in Adams & Wu, 2002). The ACER ConQuest software was used for drawing plausible values.

In previous cycles, plausible values were drawn by jurisdiction separately for each year level. A new approach of drawing plausible values nationally by year level was investigated. To evaluate the new approach, a new set of plausible values was drawn for the 2016 NAP–CC dataset regressing on dummy variables of explicit strata of jurisdiction by sector. For each year level, the subgroup means of the new plausible values such as gender and jurisdiction were then compared with the 2016 reported means. The comparison suggested that the results of the new approach are consistent with the reported result. It was therefore decided to apply the new conditioning approach for this cycle.

Some variables were used as direct regressors in the conditioning model for drawing plausible values. The variables included dummy variables of explicit strata of jurisdiction by sector, school mean performance adjusted for the

student's own performance[10], the school's geolocation and the student-level variables of gender, Indigenous status, language background other than English (LBOTE), highest parental education (PARED) and highest parental occupation group (POCC). Principle component analysis (PCA) was used to extract component scores from all other student-background variables and responses to questions in the student survey. The principle components were estimated separately by year level. Subsequently, the components that explained 99 per cent of the variance in the original variables were included as regressors in the final conditioning model for each year level. Details of the coding of variables included directly in the conditioning model or included in the PCA are listed in Appendix A5.

SCALING SURVEY ITEMS

The survey included items measuring constructs within two broad areas of interest: students' attitudes towards civics and citizenship issues (five scales) and students' engagement in civics and citizenship activities (five scales). The content of the constructs was described in Chapter 2. This section describes the scaling procedures and the psychometric properties of the survey scales.

Most of the survey scaling procedures remain the same as for the 2016 cycle. A few changes were made to the survey in 2019, including modification of two existing items and the addition of 15 new items. There were some differences in the composition of the derived survey scales, as detailed below.

- The scale relating to the confidence to engage in civic action (CIVCONF) includes one additional item in comparison to the 2016 cycle (new item).
- The scales relating to the perceptions of problems affecting Australia (PROBLEM) include two additional items in comparison to the 2016 cycle (new items).
- The scale relating to attitudes towards Australian diversity (ATAUSDIF) includes one additional item, and three items reversely worded in comparison to the 2016 cycle (new item and heavily modified items).

Exploratory factor analyses were carried out on newly developed or heavily modified scales (CIVCONF, PROBLEM, and ATAUSDIF) to provide evidence of the factor structure (suggesting a one-factor solution to the scale that fits the conceptual model).

Before estimating student scale scores for the survey indices, confirmatory factor analyses were carried out for all scales to evaluate the dimensionality of each set of items. Factorial analyses largely confirmed the expected dimensional structure of item sets and the resulting scales had satisfactory reliabilities. For example, there were eight items designed to measure intentions to promote important issues in the future (PROMIS) and five items reflecting student Intentions to

[10] So called *weighted likelihood estimates* (WLEs) were used as ability estimates in this case (Warm, 1989).

engage in civic action (CIVACT). The analyses confirmed the expected one-dimensional factor structure of each of these item sets.

Two items, originally expected to measure trust in civic institutions and processes (CIVTRUST), had relatively low correlations with the other items in the same scale and were therefore excluded from scaling.

Table 5.1 shows scale descriptions, scale names and number of items for each derived survey scale. In addition, the table includes scale reliabilities (Cronbach's alpha) as well as the correlations with student test scores for each year level.

Table 5.1: Description of survey scales

	Index name	Question number	Number of items	Scores	Cronbach's alpha		Correlation with achievement	
					Year 6	Year 10	Year 6	Year 10
Students' engagement in civic and citizenship activities								
Civic-related participation at school	No scale	a-i	9	0-1	-	-	-	-
Civic-related participation in the community	No scale[3]	a-f	6	0-2	-	-	-	-
Participation in civic-related communication	No scale	a-g	7	0-3	-	-	-	-
Intentions to promote important issues in the future	PROMIS	a-h	8	0-3	0.84	0.88	0.19	0.30
Student Intentions to engage in civic action	CIVACT[3]	a-e	5	0-3	-	0.82	-	0.19
Civic Interest	CIVINT	a-f	6	0-3	0.82	0.86	0.19	0.34
Confidence to engage in civic action	CIVCONF	a-h	8	0-3	0.89	0.90	0.25	0.35
Valuing civic action	VALCIV	a-e(f)[1]	5(6)	0-3	0.78	0.86	0.24	0.25
Students' attitudes towards civic and citizenship issues								
The importance of conventional citizenship	IMPCCON	a-e, k	6	0-3	0.79	0.85	0.12	0.27
The importance of social movement related citizenship	IMPCSOC	f-j	5	0-3	0.84	0.89	0.18	0.29
Trust in civic institutions and processes	CIVTRUST	a-f[2]	6	0-3	0.89	0.90	0.09	0.12
Attitudes towards Australian Indigenous culture	ATINCULT	a-e	5	0-3	0.88	0.93	0.29	0.31
Attitudes towards Australian diversity	ATAUSDIF[3]	a-g	7	0-3	-	0.92	-	0.32
Perceptions of problems affecting Australia	PROBLEM	a-k	11	0-3	0.90	0.89	-0.12	-0.03

[1] *Five questions for Year 6, six for Year 10*

[2] *Two items (g and h) were excluded from the scale.*

[3] *Indices only available for Year 10*

Student and item parameters were estimated using the ACER ConQuest software. Items were scaled using the Rasch partial credit model (Masters & Wright, 1997). Item difficulty parameters and students' attitudes (WLEs) were estimated for Year 6 and Year 10 separately on the full sample, weighting all states and territories equally.

When calibrating the item parameters, for each scale the average item difficulty was fixed to zero. Then, horizontal equating was conducted to put the student scores on to the same scale as last cycle so that the results could be compared. The transformation was applied as follows:

$$WLE^T = ((WLE + d + c - b) / a) * 10 + 50$$

Where WLE^T is the transformed student score for student T, WLE is the original attitude estimate in logits, d is the horizontal equating shift for Year 6 or Year 10 from 2019 to 2010., c is the vertical equating shift for Year 6 or Year 10 student scores established in 2010 for CIVCONF, CIVTRUST, PROMIS and VALCIV, b is the 2010 mean estimate in logits of the Year 10 students and a is the 2010 standard deviation in logits of the Year 10 students. The scales were converted to a metric with a mean score of 50 and a standard deviation of 10 for the Year 6

sample. A detailed description about the 2010 vertical equating shift is given in the NAP–CC 2010 Technical Report (Gebhardt, Fraillon, Wernert & Schulz, 2011).

Table 5.2 lists the transformation parameters for each of the survey scales. Note that

the transformation parameter d is computed using 2019 and 2013 attitude item estimates in logits rather than using 2019 and 2010 estimates. However, since the item and scaling parameterization of 2013 were exactly the same as in 2010, the horizontal equating shifts from 2013 to 2010 were zeros, which yields the horizontal equating shifts from 2019 to 2013 stay the same as those from 2019 to 2010.

the 2016 mean and standard deviation in logits were used as b for PROBLEM as this scale was developed and included in 2016, and

ATAUSDIF was considered to be a new scale in 2019 as it had been heavily modified. Therefore, it was not equated back to the historical scale.

Table 5.2 Transformation parameters for survey scales

	2019 to 2010 Horizontal Shift (d)		2010 Vertical Shift (c)		2010 Mean (b)	2010 SD (a)
	Year 6	Year 10	Year 6	Year 10	Year 10	Year 10
ATINCULT	0.00	0.00			2.42	2.50
CIVACT	-	0.00			-0.98	1.56
CIVCONF	0.15	0.04	-0.14	0.02	0.10	1.74
CIVINT	-0.10	-0.06			0.28	1.69
CIVTRUST	0.02	0.03	0.00	-0.13	-0.07	1.92
IMPCCON	-0.10	-0.13			0.55	1.63
IMPCSOC	-0.16	-0.20			1.03	2.15
PROMIS	-0.07	-0.09	0.05	-0.03	-0.15	1.46
VALCIV	-0.02	-0.03		0.03	1.41	1.63
PROBLEM	-0.11	-0.02			0.783[†]	1.53[†]

[†] *2016 Mean and SD were used for PROBLEM*

Similar to the equating process of the cognitive scale, equating errors need to be applied when comparing results of 2019 with results from 2016. For the survey scales, all items were within the same units and had the same maximum score. Therefore, a less complicated formula was used to compute the equating errors. After adjusting the item difficulties by applying the shifts so that the average difficulty of the items in a scale is equal in 2019 and 2016, the following formula was applied

$$EqErr = \frac{SD(d_i)}{\sqrt{N}},$$

where d_i is the difference between the adjusted difficulties of item *i* in 2016 and 2019 and *N* is the number of items in each scale.

The equating errors are presented in Table 5.3.

Table 5.3 Equating errors for survey scales

	Equating Error					
	2019 with 2016		2019 with 2013		2019 with 2010	
	Year 6	Year 10	Year 6	Year 10	Year 6	Year 10
ATINCULT	0.16	0.24	0.184	0.270	0.31	0.35
CIVACT	-	0.44	-	0.547	-	0.78
CIVCONF	0.10	0.17	0.106	0.318	0.24	0.36
CIVINT	0.308	0.27	0.481	0.307	0.54	0.38
CIVTRUST	0.07	0.12	0.218	0.414	0.35	0.79
IMPCCON	0.523	0.38	0.541	0.382	0.56	0.43
IMPCSOC	0.40	0.23	0.548	0.518	0.57	0.60
PROMIS	0.139	0.39	0.345	0.414	0.38	0.57
VALCIV	0.20	0.25	0.248	0.281	0.35	0.30
PROBLEM	0.608	0.52	-	-	-	-

Chapter 6 PROFICIENCY LEVELS AND THE PROFICIENT STANDARDS

One of the key objectives of NAP–CC is to monitor trends in civics and citizenship achievement over time. The NAP–CC scale forms the basis for the empirical comparison of student achievement. In addition to the metric established for the scale, a set of proficiency levels with substantive descriptions form the NAP–CC described proficiency scale.

One mechanism for monitoring trends in the NAP sample assessments (ICT literacy, civics and citizenship, and science literacy) is the reporting of student attainment of key performance measures (KPMs) defined for each area. The proportion of students achieving at or above the proficient standard for each of Year 6 and Year 10 is the national KPM for civics and citizenship specified in the Measurement Framework for Schooling in Australia (ACARA 2019).

This chapter describes the establishment and subsequent revision of the NAP–CC proficiency levels and summarises the process used as part of NAP–CC to establish the Year 6 and Year 10 NAP–CC proficient standards.

PROFICIENCY LEVELS

Assumptions underpinning the proficiency levels

The proficiency levels were established in 2004 and were based on an approach developed for the OECD's Project for International Student Assessment (PISA). For PISA, a method was developed that ensured that the notion of being at a level could be interpreted consistently and in line with the fact that the achievement scale is a continuum. This method ensured that there was some common understanding about what being at a level meant and that the meaning of being at a level was consistent across levels. Similar to the approach taken in the PISA study (OECD 2005, p.255), this method takes the following three variables into account:

- the expected success of a student at a particular level on a test containing items at that level
- the width of the levels in that scale, and
- the probability that a student in the middle of a level would correctly answer an item of average difficulty for that level.

To achieve this for NAP–CC, the following two parameters for defining proficiency levels were adopted by the PMRT:

- setting the response probability for the analysis of data at $p = 0.62$; and
- setting the width of the proficiency levels at 1.00 logit.

With these parameters established, the following statements can be made about the achievement of students relative to the proficiency levels.

- A student whose result places him/her at the lowest possible point of the proficiency level is likely to get approximately 50 per cent correct on a test made up of items spread uniformly across the level, from the easiest to the most difficult. In other words, any student whose performance is within a level is expected to respond correctly to at least 50 per cent of the items that are located within the same level and is therefore regarded as being able to demonstrate skills required to answer items at that level.
- A student whose result places him/her at the lowest possible point of the proficiency level is likely to get 62 per cent correct on a test made up of items similar to the easiest items in the level.
- A student at the top of the proficiency level is likely to get 82 per cent correct on a test made up of items similar to the easiest items in the level.

Establishing the position of and describing the proficiency levels

The positioning of the proficiency levels on the NAP–CC scale was done together with a standards setting exercise in which a proficient standard was established for each year level. The Year 6 proficient standard was set at 405 scale points, the cut-point between Level 1 and Level 2 on the NAP–CC scale, and the Year 10 proficient standard was set at 535 scale points the cut-point between Level 2 and Level 3 (details of the standard-setting procedures are reported later in this chapter).

Clearly, other solutions with different parameters defining the proficiency levels and alternative inferences about the likely per cent correct on tests could also have been chosen. The approach used in PISA, and adopted for NAP–CC, attempted to balance the notions of mastery and 'pass' in a way that is likely to be understood by the community.

Proficiency level cut points

Six proficiency levels were generated for reporting student achievement. The levels were generated following the establishment of the boundary of levels 1 and 2 at 405 scale points. Table 6.1 shows these levels and shows the percentage of Year 6 and Year 10 students in each level in NAP–CC 2019.

Table 6.1: Proficiency level cut-points and percentage of Year 6 and Year 10 students in each level in 2019

Proficiency Level	Lower level boundary (scale points)	Percentage	
		Year 6	Year 10
Level 5	795	1 (±0.3)*	1 (±0.4)
Level 4	665		8 (±1.5)
Level 3	535	14 (±1.4)	29 (±2.1)
Level 2	405	37 (±1.7)	36 (±2.5)

Level 1	275	33 (±1.7)	18 (±2.0)
Below Level 1		15 (±1.5)	7 (±1.3)

**Levels 4 and 5 for Year 6 are reported together*

Describing proficiency levels

The proficiency levels were described using a combination of expert descriptions of the knowledge of the skills required to answer each civics and citizenship item and information from the analysis of students' responses. Each level description provides a synthesised overview of the civics and citizenship and history knowledge and understanding that a student working within the level is able to demonstrate as evidenced by the assessment items within that level.

Summary descriptors for levels 1 to 5 of the NAP–CC scale were established in the first cycle of NAP–CC in 2004. A descriptor for 'below level 1' achievement was developed in 2007 when more test material was available to support this description.

Routinely as part of each NAP–CC cycle, the proficiency level descriptors are reviewed with respect to new item content and consequently revised if warranted. New examples of achievement at each level are also added to supplement the level descriptors as appropriate.

The proficiency level descriptors were updated in 2013 to reflect the larger pool of items that had been developed over the cycles since 2004. In 2019, the scale descriptors were further revised to reflect the inclusion of items from the NAP–CC history sub-strand of the revised NAP–CC Assessment Framework.

The NAP–CC scale represents a hierarchy of students' knowledge, skills and understanding associated with civics and citizenship content. The scale describes a developmental learning progression in the sense that students are assumed to be typically able to demonstrate achievement of the content and cognitive processes described at the level below, as well as at their measured level of achievement.

The proficiency level descriptors are provided in Appendix A6.

SETTING THE PROFICIENT STANDARDS

The proficient standards "represent a 'challenging but reasonable' expectation of student achievement at a year level with students needing to demonstrate more than elementary skills expected at that year level" (ACARA 2019, p. 5). This is different from the definition of either a benchmark or a national minimum standard, which refer to minimum competence.

The process for setting standards in areas such as primary science, information and communications technologies, civics and citizenship and secondary (15-year-old) reading, mathematics and science was endorsed by the PMRT at its 6 March 2003 meeting and is described in the paper, Setting National Standards (PMRT 2003).

The Year 6 and Year 10 proficient standards for NAP–CC were set in March 2005, with an expert group of civics and citizenship educators from all Australian jurisdictions using a combination of a modified Angoff (yes/no) and Bookmark standards-setting procedures. A description of this process is given in the NAP–CC 2004 Technical Report.

To access the NAP–CC public report and technical report documents from previous cycles visit www.nap.edu.au > 'Results and reports' section > 'National reports' page.

By referring to the proficient standards, Year 6 students performing at level 2 and above, and Year 10 students performing at level 3 and above have consequently met or exceeded their relevant proficient standard.

The proficient standards for Year 6 and Year 10 civics and citizenship achievement were endorsed by the Key Performance Measures subgroup of the PMRT in 2005. These standards have remained unchanged as the KPMs for civics and citizenship across all subsequent cycles (ACARA 2019, p 12).

Chapter 7 REPORTING OF RESULTS

The students assessed in NAP–CC 2019 were selected using a two-stage cluster sampling procedure. At the first stage, schools were sampled from a sampling frame with a probability proportional to their size as measured by student enrolments in the relevant year level. In the second stage, 20 students at each year level were randomly sampled within schools (see Chapter 3 for further information on sampling and weighting).

Applying cluster sampling techniques is an efficient and economical way of selecting students in educational research. However, as these samples were not obtained through (one-stage) simple random sampling, standard formulae to obtain sampling errors of population estimates are not appropriate. In addition, NAP–CC estimates were obtained using plausible value methodology (see Chapter 5 on scaling procedures), which allows for estimating and combining the measurement error of achievement scores with their sampling error.

This chapter describes the method applied for estimating sampling as well as measurement error. In addition, it contains a description of the types of statistical analyses and significance tests that were carried out for reporting of results in the NAP–CC 2019 National Report.

COMPUTATION OF SAMPLING AND MEASUREMENT VARIANCE

Unbiased standard errors from studies should include both sampling variance and measurement variance. One way of estimating sampling variance on population estimates from cluster samples is by utilising the application of replication techniques (Wolter 1985; Gonzalez & Foy 2000). The sampling variances of population means, differences, percentages and correlation coefficients in NAP–CC studies were estimated using the jackknife repeated replication technique (JRR). The other component of the standard error of achievement test scores, the measurement variance, can be derived from the variance among the five plausible values for NAP–CC. In addition, for comparing achievement test scores with those from previous cycles (2004, 2007, 2010, 2013 and 2016), an equating error was added as a third component of the standard error.

REPLICATE WEIGHTS

When applying the JRR method for stratified samples, primary sampling units (PSUs) – in this case schools – are paired into pseudo-strata, also called sampling zones. The assignment of schools to these sampling zones needs to be consistent with the sampling frame from which they were sampled (to obtain pairs of schools that were adjacent in the sampling frame) and zones are always constructed within explicit strata of the sampling frame. This procedure ensures

that schools within each zone are as similar to each other as possible[11]. For NAP–CC 2019 there were 86 sampling zones each in Year 6 and Year 10.

Within each sampling zone, one school was randomly assigned a value of two, whereas the other one received a value of zero. To create replicate weights for each of these sampling zones, the jackknife indicator variable was multiplied by the original sampling weights of students within the corresponding zone so that one of the paired schools had a contribution of zero and the other school a double contribution, whereas schools from all other sampling zones remained unmodified.

At each year level, 86 replicate weights were computed. This was done in order to have a consistent number of replicate weight variables in the final database.

STANDARD ERRORS

In order to compute the sampling variance for a statistic t, t is estimated once for the original sample S and then for each of the jackknife replicates J_h. The JRR variance is computed using the formula:

$$Var_{jrr}(t) = \sum_{h=1}^{H} \left[t(J_h) - t(S) \right]^2$$

where H is the number of replicate weights, $t(S)$ is the statistic t estimated for the population using the final sampling weights, and $t\ (J_h)$ is the same statistic estimated using the weights for the hth jackknife replicate. For all statistics that are based on variables other than student test scores (plausible values), the standard error of t is equal to:

$$\sigma(t) = \sqrt{Var_{jrr}(t)}$$

The computation of JRR variance can be obtained for any statistic. However, many standard statistical software packages such as SPSS® do not generally include any procedures for replication techniques. Therefore, specialist software, the SPSS® Replicates add-in, was used to run tailored SPSS® macros to estimate JRR variance for means and percentages[12].

Population statistics for NAP–CC scores were always estimated using all five plausible values, with standard errors reflecting both sampling and measurement error. If t is any computed statistic and t_j is the statistic of interest computed on one plausible value, then:

[11] In the case of an odd number of schools within an explicit stratum on the sampling frame, the remaining school is randomly divided into two halves and each half assigned to the two other schools in the final sampling zone to form *pseudo-schools*.

[12] Conceptual background and application of macros with examples are described in the *PISA Data Analysis Manual* SPSS®, 2nd edn (OECD, 2009b).

$$t = \frac{1}{M}\sum_{i=1}^{M} t_i$$

with M being the number of plausible values.

The sampling variance U is calculated as the average of the sampling variance for each plausible value U_i :

$$U = \frac{1}{M}\sum_{i=1}^{M} U_i$$

Using five plausible values for data analysis allows the estimation of the error associated with the measurement of NAP–CC due to the lack of precision of the test instrument. The measurement variance or imputation variance Bm was computed as:

$$B_m = \frac{1}{M-1}\sum_{i=1}^{M}\left(t_i - t\right)^2$$

To obtain the final standard error of NAP–CC statistics, the sampling variance and measurement variance were combined as:

$$SE = \sqrt{U + \left(1 + \frac{1}{M}\right)B_m}$$

with U being the sampling variance.

The 95 per cent confidence interval, as presented in the NAP–CC 2019 National Report, was computed as 1.96 times the standard error, which is actually the range of the confidence interval. The actual 95 per cent confidence interval of a statistic is between the value of the statistic minus 1.96 times the standard error and the value of the statistic plus 1.96 times the standard error.

REPORTING OF DIFFERENCES IN AVERAGE ACHIEVEMENT

This report includes comparisons of average achievement across states and territories; that is, averages of scales and percentages were compared in graphs and tables. Each population estimate was accompanied by its 95 per cent confidence interval. In addition, tests of significance for the difference between estimates were provided, in order to describe the probability that differences were just a result of sampling and measurement errors.

The following types of significance tests for differences in average achievement population estimates were reported:

- between states and territories;
- between student background subgroups; and

- across the six assessment cycles (2004, 2007, 2010, 2013, 2016 and 2019).

Differences in average achievement between states and territories and year levels

Pairwise comparison charts allow the comparison of population estimates between one state or territory and another or between Year 6 and Year 10. Differences in averages were considered significant when the test statistic *t* was outside the critical values ±1.96 (α = 0.05). The t value is calculated by dividing the difference in averages by its standard error that is given by the formula:

$$SE_{dif\_ij} = \sqrt{SE_i^2 + SE_j^2}$$

where SE_{dif_ij} is the standard error on the difference and SE_i and SE_j are the standard errors of the compared averages $i$and j. The standard error on a difference can only be computed this way if the comparison is between two independent samples like states and territories or year levels. Samples are independent if they were drawn separately.

Differences in average achievement between dependent subgroups

The formula for calculating the standard error provided above is only suitable when the subsamples being compared are independent (see OECD 2009 for more detailed information). In the case of dependent subgroups, the covariance between the two standard errors needs to be taken into account and the Jackknife repeated replication (JRR) technique should be used to estimate the sampling error for average differences.

As subgroups other than 'state or territory' and 'year level' are dependent subsamples (for example, gender and language background subgroups), the difference between statistics for subgroups of interest and the standard error of the difference were derived using the specialist software *SPSS® Replicates Add-in* that runs macros to apply JRR. Differences between subgroups were considered significant when the test statistic *t* was outside the critical values ±1.96 (α = 0.05). The value *t* was calculated by dividing the average difference by its standard error.

Differences in average achievement between assessment cycles

This report also includes comparisons of assessment results across cycles. As the process of equating the tests across the cycles introduces some additional error into the calculation of any test statistic, an equating error term was added to the formula for the standard error of the difference (between cycle averages, for example).

The value of the equating error between 2019 and 2016 is 2.968 units of the civics and citizenship scale for Year 6 and 3.146 for Year 10. When testing the

difference of a statistic between the two assessments, the standard error of the difference is computed as follows:

$$SE(\mu_{19} - \mu_{16}) = \sqrt{SE_{19}^2 + SE_{16}^2 + EqErr_{19-16}^2}$$

where μ can be any statistic in units on the NAP–CC scale (average, percentile, gender difference, but *not* percentages), SE_{19}^2 is the respective standard error of this statistic in 2019, SE_{16}^2 is the respective standard error of this statistic in 2016,.and $EqErr_{19-16}^2$ is the equating error for comparing 2019 with 2016 results.

When comparing population estimates between 2019 and the third assessment in 2013, two equating errors (between 2019 and 2016 and between 2016 and 2013) had to be taken into account. This was achieved by applying the following formula for the calculation of the standard error for differences between statistics from 2019 and 2013:

$$SE(\mu_{19} - \mu_{13}) = \sqrt{SE_{19}^2 + SE_{13}^2 + EqErr_{19-13}^2}$$

For Year 6, $EqErr_{19-16}^2$ reflects the uncertainty associated with the equating between the assessment cycles of 2019 and 2016 (2.97 score points), as well as between 2016 and 2013 (4.42 score points). This combined equating error was equal to 5.33 score points and was calculated as:

$$EqErr_{19\_13} = \sqrt{\mathrm{EqErr}_{19_{16}}^2 + \mathrm{EqErr}_{16_{13}}^2}$$

Similarly, for comparisons between 2019 and the first cycle in 2004, the equating errors between each adjacent pair of assessments had to be taken into account and standard errors for differences were computed as:

$$SE(\mu_{19} - \mu_{04}) = \sqrt{SE_{19}^2 + SE_{04}^2 + EqErr_{19-04}^2}$$

The combined equating error for Year 6 was equal to 9.92 score points, and was calculated as:

$$EqErr_{19\_13} = \sqrt{\mathrm{EqErr}_{19_{16}}^2 + \mathrm{EqErr}_{16_{13}}^2 + \mathrm{EqErr}_{13_{10}}^2 + \mathrm{EqErr}_{10_{07}}^2 + \mathrm{EqErr}_{07_{04}}^2}$$

The equating errors for comparing averages between 2019 and each previous NAP–CC cycle are provided in Table 7.1.

Table 7.1: Equating errors for comparing averages between NAP–CC 2019 and each previous assessment cycle

Assessment cycle years	Equating error of average difference (scale points) Year 6	Year 10
2019-2016	2.97	3.15
2019-2013	5.33	5.39
2019-2010	7.20	7.17
2019-2007	8.93	8.36
2019-2004	9.92	8.66

Differences in percentages between assessment cycles

To report the significance of differences between percentages at or above proficient standards, the equating error for each year level could not be applied directly. Therefore, the following replication method was applied to estimate the equating error for percentages at proficient standards.

For each year level cut-point that defines the corresponding proficient standard (405 for Year 6 and 535 for Year 10), a number of *n* replicate cut-points were generated (5,000) by adding a random error component with an average of 0 and a standard deviation equal to the estimated equating error. Percentages of students at or above each replicate cut-point (ρ_n) were computed and an equating error for each year level was estimated as

$$EquErr(\rho) = \sqrt{\frac{(\rho_n - \rho_o)^2}{n}}$$

where ρ_o is the percentage of students at or above the (reported) proficient standard. The standard errors for the differences between percentages at or above proficient standards were calculated as:

$$SE(\rho_{19} - \rho_{16}) = \sqrt{SE(\rho_{19})^2 + SE(\rho_{16})^2 + EqErr(\rho)^2}$$

ρ_{16} and ρ_{19} are the percentages at or above the proficient standard in 2016 and 2019 respectively.

The equating errors for comparing percentage achievement between 2019 and each previous NAP–CC cycle are provided in Table 7.2 and Table 7.3 for Year 6 and Year 10 respectively.

Table 7.2: Equating errors for comparing percentages between NAP–CC 2019 and each previous assessment cycle (Year 6)

Year	Group	Equating Error 2019 with 2016	2013	2010	2007	2004
6	Aust.	0.86	1.56	2.55	2.71	3.35
6	NSW	0.91	1.62	2.66	2.84	3.53
6	Vic.	0.76	1.46	2.49	2.65	3.30
6	QLD	1.01	1.76	2.72	2.86	3.45
6	WA	0.78	1.50	2.45	2.60	3.22
6	SA	1.06	1.76	2.72	2.87	3.47
6	Tas.	0.95	1.55	2.32	2.44	2.94
6	ACT	0.87	1.67	2.59	2.72	3.23
6	NT	0.63	1.11	1.96	2.11	2.72
6	Female	0.87	1.59	2.60	2.76	3.43
6	Male	0.89	1.58	2.52	2.67	3.28
6	Non-Indigenous	0.88				
6	Indigenous	0.83				
6	English	0.88				
6	Language other than English	0.83				
6	Senior managers and professionals	0.67				
6	Other managers and associate professionals	1.08				
6	Tradespeople & skilled office, sales and service staff	1.05				
6	Unskilled workers; hospitality	0.52				
6	Not in paid work in last 12 months	1.00				
6	Not stated or unknown	0.99				
6	Year 9 or equivalent or below	0.47				
6	Year 10 or equivalent or below	0.83				
6	Year 11 or equivalent	0.82				
6	Year 12 or equivalent	1.11				
6	Certificate 1 to 4 (inc. trade cert.)	0.73				
6	Advanced diploma/diploma	1.14				
6	Bachelor degree or above	0.94				
6	Not stated or unknown	1.08				

Table 7.3: Equating errors for comparing percentages between NAP–CC 2019 and each previous assessment cycle (Year 10)

		Equating Error 2019 with				
Year	Group	2016	2013	2010	2007	2004
10	Aust.	0.74	1.36	2.24	2.23	2.64
10	NSW	0.72	1.31	2.13	2.12	2.49
10	Vic.	0.83	1.51	2.46	2.45	2.90
10	QLD	0.70	1.26	2.19	2.17	2.64
10	WA	0.75	1.30	2.09	2.08	2.42
10	SA	0.78	1.46	2.34	2.33	2.73
10	Tas.	0.64	1.48	2.67	2.65	3.15
10	ACT	1.04	1.55	2.19	2.18	2.54
10	NT	1.11	1.99	3.01	3.00	3.42
10	Female	0.59	1.16	2.03	2.02	2.46
10	Male	0.91	1.59	2.48	2.47	2.85
10	Non-Indigenous	0.76				
10	Indigenous	0.30				
10	English	0.68				
10	Language other than English	0.93				
10	Senior managers and professionals	0.90				
10	Other managers and associate professionals	0.90				
10	Tradespeople & skilled office, sales and service staff	0.63				
10	Unskilled workers; hospitality	1.01				
10	Not in paid work in last 12 months	0.69				
10	Not stated or unknown	0.40				
10	Year 9 or equivalent or below	1.03				
10	Year 10 or equivalent or below	0.57				
10	Year 11 or equivalent	0.40				
10	Year 12 or equivalent	0.41				
10	Certificate 1 to 4 (inc. trade cert.)	0.92				
10	Advanced diploma/diploma	0.71				
10	Bachelor degree or above	0.83				
10	Not stated or unknown	0.55				

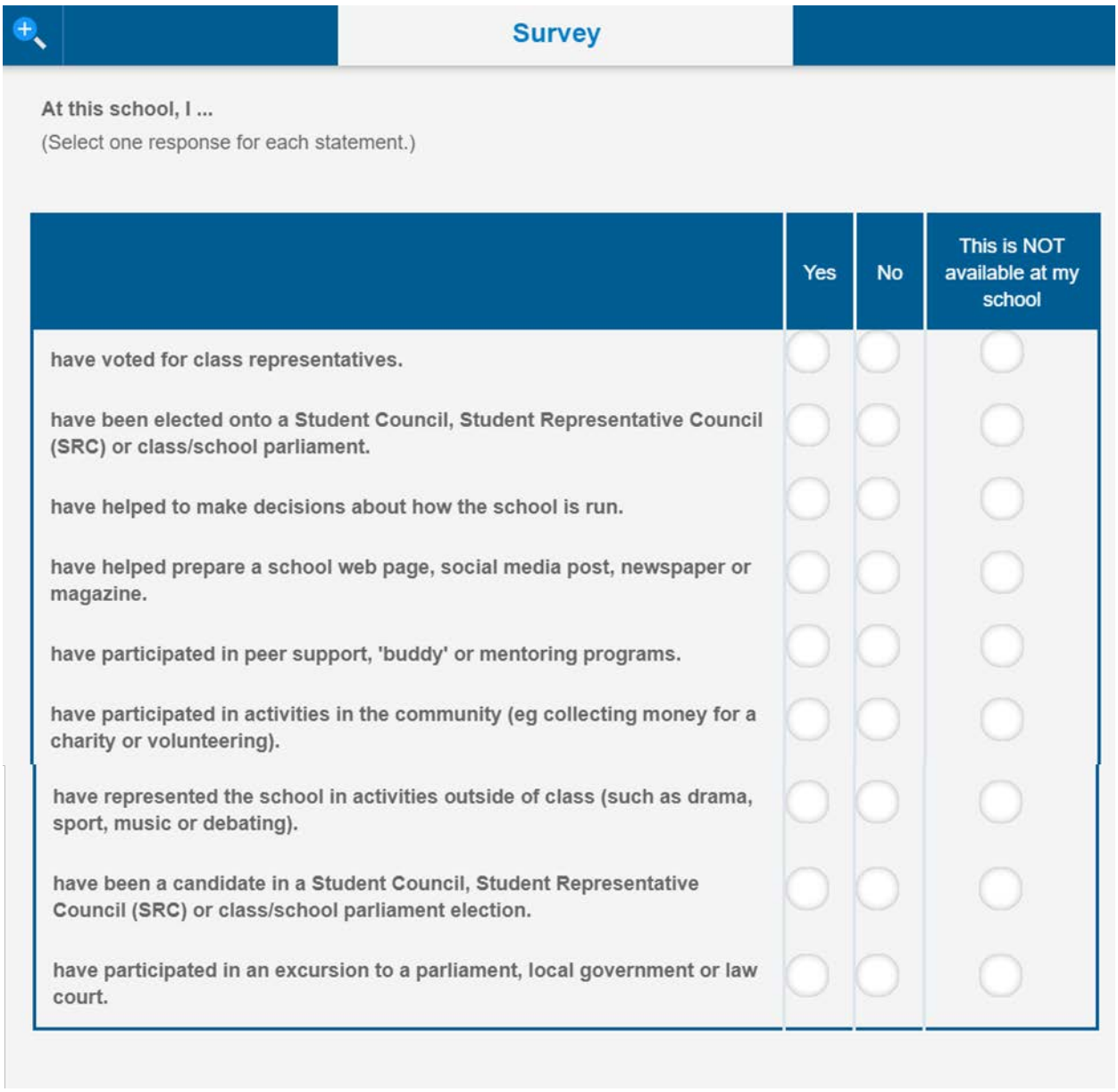
Appendix A1: Student survey

The Year 6 and Year 10 student survey instruments contained mostly the same questions. However, an alternative set of items was administered for each year level for item set 8, and Year 6 students were not administered the following item sets at all:

- item set 2
- item set 5
- item set 12.

All student survey item sets are presented on the following pages.

ITEM SET 1



Survey

At this school, I ...

(Select one response for each statement.)

	Yes	No	This is NOT available at my school
have voted for class representatives.			
have been elected onto a Student Council, Student Representative Council (SRC) or class/school parliament.			
have helped to make decisions about how the school is run.			
have helped prepare a school web page, social media post, newspaper or magazine.			
have participated in peer support, 'buddy' or mentoring programs.			
have participated in activities in the community (eg collecting money for a charity or volunteering).			
have represented the school in activities outside of class (such as drama, sport, music or debating).			
have been a candidate in a Student Council, Student Representative Council (SRC) or class/school parliament election.			
have participated in an excursion to a parliament, local government or law court.			

ITEM SET 2 (YEAR 10 ONLY)

Have you ever participated in activities associated with each of the following?

(Select one response for each statement.)

	Yes, I have done this within the past 12 months	Yes, I have done this but not within the past 12 months	No, I have never done this
collecting money for a charity or social cause	◉	○	○
a voluntary group doing something to help the community	○	○	○
an environmental organisation	○	○	○
a human rights organisation	○	○	○
a youth development organisation (eg Scouts, Australian Services Cadets, Police and Community Youth Clubs)	○	○	○
an animal rights or protection organisation	○	○	○

ITEM SET 3

Outside of school, how often do you…

(Select one response for each statement.)

	More than three times a week	At least once a week	At least once a month	Never or hardly ever
use the internet (including social media) to get news of current events?	○	○	○	○
watch the news on television?	○	○	○	○
listen to news on the radio?	○	○	○	○
read about current events in a paper or online newspaper?	○	○	○	○
post or share a comment or image about a political or social issue on the internet (including social media)?	○	○	○	○
talk about political or social issues with your family?	○	○	○	○
talk about political or social issues with your friends?	○	○	○	○

ITEM SET 4

There are many different ways to express your opinions about important issues.
Would you do any of the following in the future?
(Select one response for each statement.)

	I would certainly do this	I would probably do this	I would probably NOT do this	I would certainly NOT do this
sign an online petition	○	○	○	○
write a letter or an email to a newspaper	○	○	○	○
write your opinion about an issue on the internet (eg on social media, a blog or web forum)	○	○	○	○
wear a badge, hat or T-shirt expressing your opinion	○	○	○	○
contact a member of parliament or a local council	○	○	○	○
take part in a peaceful march or rally	○	○	○	○
collect signatures for a petition	○	○	○	○
choose NOT to buy certain products or brands of product as a protest	○	○	○	○

ITEM SET 5 (YEAR 10 ONLY)

There are many different ways people can participate in the community.
Which of the following will you do in the future?
(Select one response for each statement.)

	I will certainly do this	I will probably do this	I will probably NOT do this	I will certainly NOT do this
find information about candidates before voting in an election	○	○	○	○
help a candidate or party during an election campaign	○	○	○	○
join a political party	○	○	○	○
join a trade union or other union	○	○	○	○
stand as a candidate in local council or shire elections	○	○	○	○

ITEM SET 6

How interested are you in the following?

(Select one response for each statement.)

	Very interested	Quite interested	Not very interested	Not interested at all
what is happening in your local community	○	○	○	○
Australian politics	○	○	○	○
social issues in Australia	○	○	○	○
environmental issues in Australia	○	○	○	○
what is happening in other countries	○	○	○	○
global (worldwide) issues	○	○	○	○

ITEM SET 7

How well do you think you could do each of the following?

(Select one response for each statement.)

	Very well	Fairly well	Not very well	Not at all
discuss news about a conflict between countries	◉	○	○	○
argue your opinion about a political or social issue	○	○	○	○
be a candidate in a school or class election	○	○	○	○
organise a group of students in order to achieve changes at school	○	○	○	○
express your opinion on a current issue in a letter or email to a newspaper	○	○	○	○
give a speech to your class about a social or political issue	○	○	○	○
present information about a political or social issue on social media	○	○	○	○
express your opinion in a comment you post on a website	○	○	○	○

ITEM SET 8 (YEAR 6 VERSION)

How much do you agree or disagree with each of the following statements?

(Select one response for each statement.)

	Strongly agree	Agree	Disagree	Strongly disagree
If students act together at school they can make real change happen.	○	○	○	○
Elected student representatives (such as members of the Student Council or Student Representative Council) contribute to school decision making.	○	○	○	○
Student participation in how schools are run can make schools better.	○	○	○	○
Organising groups of students to express their opinions could help solve problems in schools.	○	○	○	○
It is important for students to vote in school elections.	○	○	○	○

ITEM SET 8 (YEAR 10 VERSION)

How much do you agree or disagree with each of the following statements?

(Select one response for each statement.)

	Strongly agree	Agree	Disagree	Strongly disagree
If students act together at school they can make real change happen.	○	○	○	○
Elected student representatives (such as members of the Student Council or Student Representative Council) contribute to school decision making.	○	○	○	○
Student participation in how schools are run can make schools better.	○	○	○	○
Organising groups of students to express their opinions could help solve problems in schools.	○	○	○	○
It is important for students to vote in school elections.	○	○	○	○
Citizens can have a strong influence on government policies in Australia.	○	○	○	○

ITEM SET 9

How important do you think the following are for being a good citizen in Australia?

(Select one response for each statement.)

	Very important	Quite important	Not very important	Not important at all
supporting a political party	○	○	○	○
learning about Australia's history	○	○	○	○
learning about political issues in the newspaper, on the radio, on TV or on the internet	○	○	○	○
learning about what happens in other countries	○	○	○	○
discussing politics	○	○	○	○
participating in peaceful protests about important issues	○	○	○	○
participating in activities to benefit the local community	○	○	○	○
taking part in activities promoting human rights	○	○	○	○
taking part in activities to protect the environment	○	○	○	○
making personal efforts to protect natural resources (eg water saving, recycling)	○	○	○	○
voting in elections	○	○	○	○

ITEM SET 10

How much do you trust each of the following groups or institutions in Australia?

(Select one response for each statement.)

	Completely	Quite a lot	A little	Not at all
the Australian parliament	○	○	○	○
your state or territory parliament	○	○	○	○
your local government (eg local council or shire)	○	○	○	○
law courts	○	○	○	○
the police	○	○	○	○
Australian political parties	○	○	○	○
the media (ie television, newspapers, radio)	○	○	○	○
social media (eg Twitter, blogs, YouTube, Facebook, Instagram)	○	○	○	○

ITEM SET 11

How much do you agree or disagree with the following statements about Aboriginal and Torres Strait Islander peoples?

(Select one response for each statement.)

	Strongly agree	Agree	Disagree	Strongly disagree
Australia should support the cultural traditions and languages of Aboriginal and Torres Strait Islander peoples.	○	○	○	○
Australia has a responsibility to improve the quality of life of Aboriginal and Torres Strait Islander peoples.	○	○	○	○
It is important to recognise the traditional ownership of their land by Aboriginal and Torres Strait Islander peoples.	○	○	○	○
All Australians have much to learn from Aboriginal and Torres Strait Islander peoples' cultures, traditions and people.	○	○	○	○
All Australians should be given the chance to learn about reconciliation between Aboriginal and Torres Strait Islander peoples and other Australians.	○	○	○	○

ITEM SET 12 (YEAR 10 ONLY)

How much do you agree or disagree with the following statements about Australian society?

(Select one response for each statement.)

	Strongly agree	Agree	Disagree	Strongly disagree
Immigrants should be encouraged to keep their cultural beliefs, practices and languages.	○	○	○	○
Australia will remain a peaceful country as more people from different backgrounds come to live here.	○	○	○	○
Australia benefits greatly from having people from many cultures and backgrounds.	○	○	○	○
At school, all students should learn about different cultural beliefs and practices.	○	○	○	○
All Australians should accept different cultural beliefs and practices.	○	○	○	○
Having people from many different cultures and backgrounds makes it easier for a country to be united.	○	○	○	○
Australia will be a better place in the future as more people with different backgrounds come to live here.	○	○	○	○

ITEM SET 13

Below is a list of problems affecting countries across the world in different ways.
In your view, to what extent is Australia affected by each of these problems?
(Select one response for each problem.)

	to a large extent	to a moderate extent	to a small extent	not at all
pollution	○	○	○	○
unemployment	○	○	○	○
terrorism	○	○	○	○
poverty	○	○	○	○
climate change	○	○	○	○
water shortages	○	○	○	○
lack of access to high-quality education	○	○	○	○
crime	○	○	○	○
lack of access to adequate health services	○	○	○	○
racism and discrimination	○	○	○	○
lack of cyber security and privacy	○	○	○	○

Appendix A2: Technical Checks – Excerpts from the STSO Manual

The nominated School Technical Support Officer at each school was tasked with completing a number of technical checks in order to ensure the school's technical set up for the assessment was 'test-ready'. STSOs were asked to complete these tasks in the weeks leading up to the scheduled assessment at their school. The instructions reproduced below are excerpts from the STSO Manual that have been modified slightly to improve readability as an Appendix.

RUN A BANDWIDTH TEST

You must ensure that your school's bandwidth capabilities meet the minimum requirements for the NAP-CC Assessment delivery system. Please make a note of the upload and download speed of the bandwidth test you complete so you can include the results in the STSO technical preparations questionnaire. If possible, do more than one bandwidth test and take an average across all tests.

To conduct the bandwidth test, please navigate to any free online speed test tool. There are many bandwidth tests available online but two are provided below:

https://speedof.me/

http://www.speedtest.net/

The bandwidth test should be done on a student computer that will be used for the assessment. For accuracy, you should also conduct the bandwidth test during normal school hours, if possible.

If your school's internet connection **does not** meet the following **minimum** requirements:

- 2 - 3 Mbps download and
- 100 – 150 Kbps upload

we may need to contact you to discuss running two or more smaller test sessions.

DOWNLOAD AND INSTALL THE LOCKED DOWN BROWSER (LDB) ON STUDENT DEVICES

Students access the NAP-CC assessment via the Locked Down Browser, so this must be installed on all devices used by students to take the assessment.

If you experience any issues when installing the LDB please see *Section 3 – Getting help*, which provides details about where to find further documents to assist you on the Assessform website.

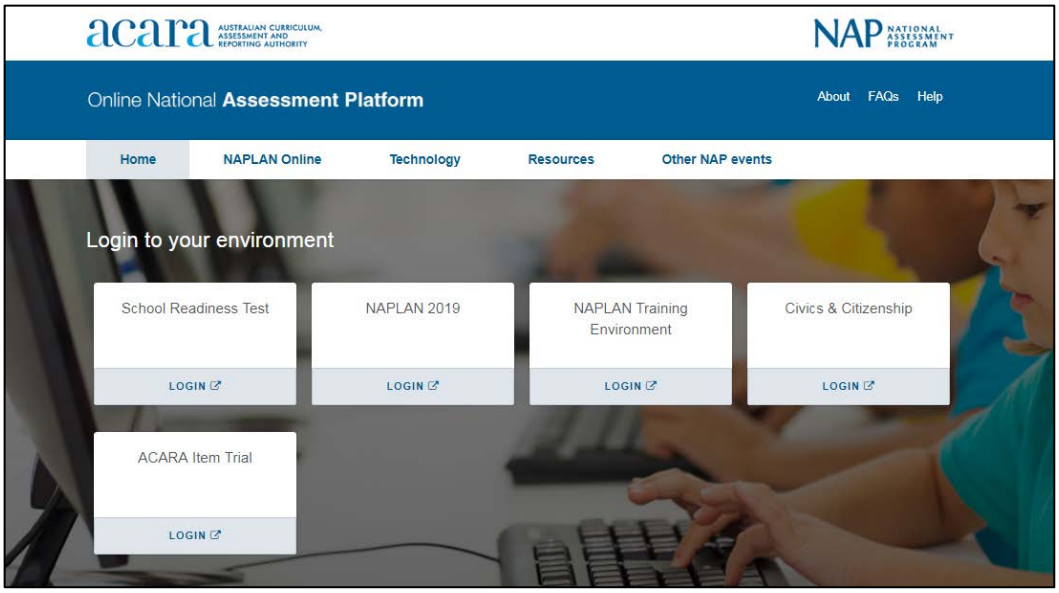
IMPORTANT NOTE: The most recent version of the LDB is needed to access the NAP-CC assessment. If any device already has the LDB installed, you should check that it is not out of date. You can do this by launching the LDB. If the system alerts you that your LDB is out of date, you will need to download a new version.

It is also imperative that the Locked Down Browser is installed on a profile that students will be able to access on the day of the test. The Device Check must

***also be conducted using this profile whilst accessing the internet connection* available to students.**

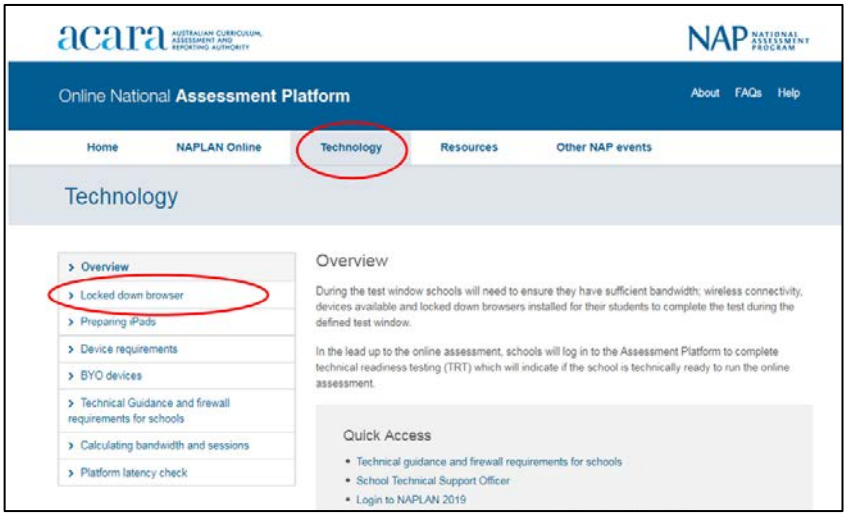
1. Open a browser, navigate to https://www.assessform.edu.au/ (Figure 1)

Figure 1: The Assessform website



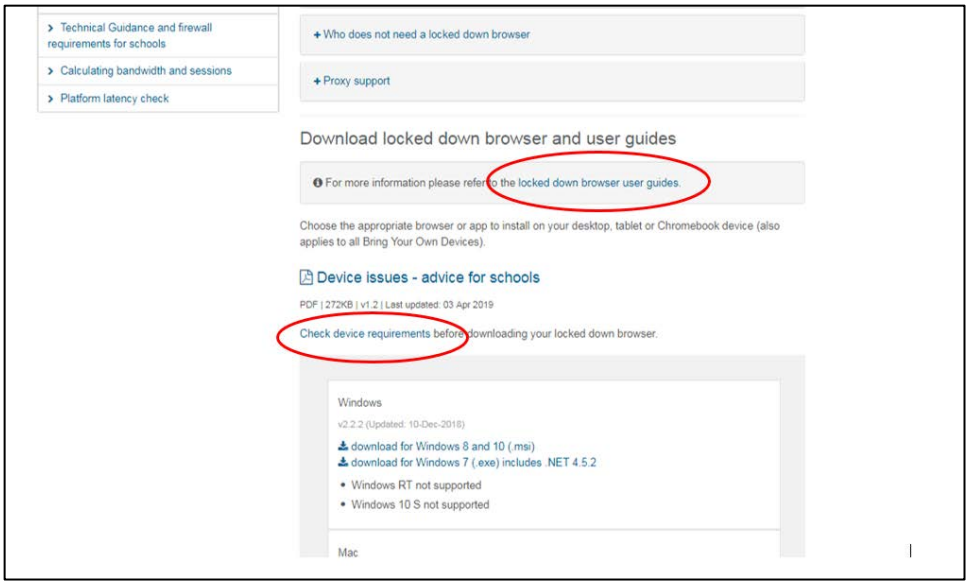
2. Click on Technology on the navigation bar and then click on Locked down browser in the left hand menu (Figure 2).

Figure 2: Technology tab



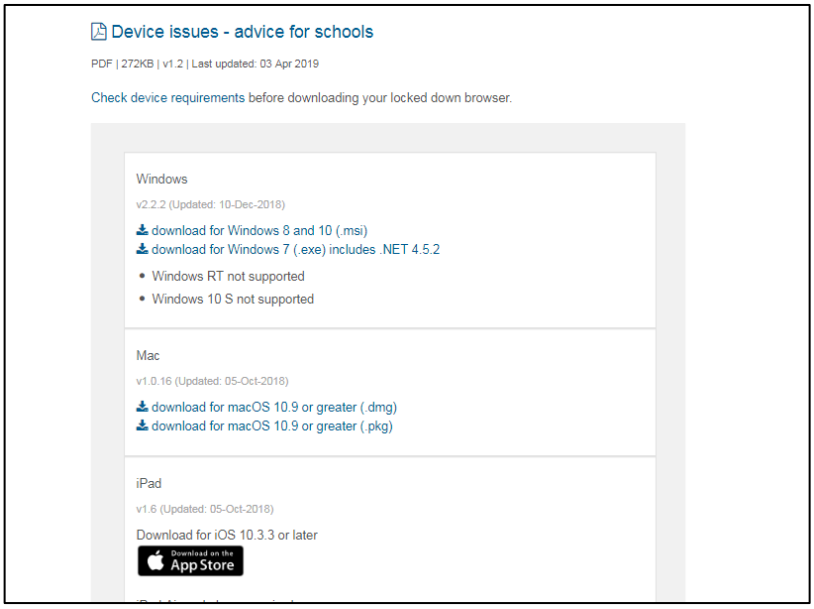
3. On the Locked down browser page (Figure 3), you will find a link to the locked down browser user guides and device requirements information. The LDB user guides provide detailed instructions for installing the LDB on a range of different devices. The Device requirements page, accessed via the Check device requirements link (Figure 3) outlines the minimum specifications a device must meet to interact successfully with the online assessment platform. You should check that student devices meet these requirements before downloading the locked down browser onto them.

Figure 3: User guides and device requirements



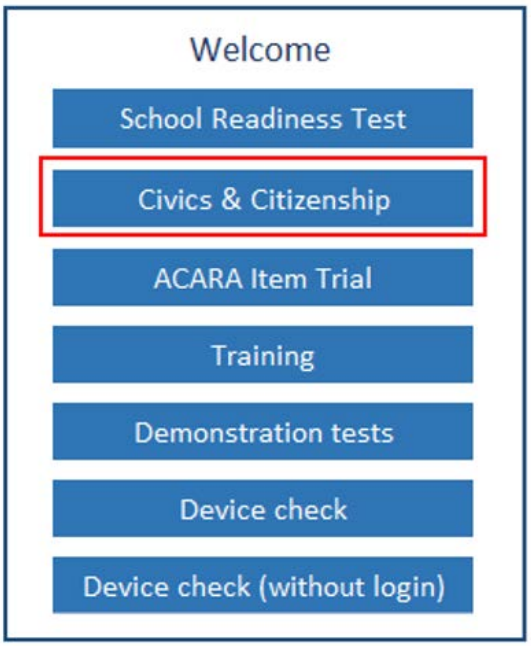
4. On the Locked down browser page you will also find download links to the LDB installation files (Figure 4). Click on the appropriate link for the device you are using and install the LDB.

Figure 4: LDB download links



5. Once installed, you must check that the installation has completed correctly. To do this, launch the LDB and select the Civics & Citizenship button from the list the system presents (Figure 5)

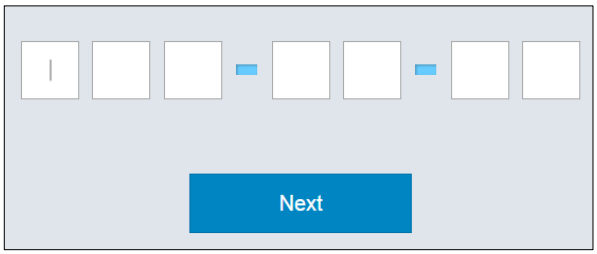
Figure 5: Welcome options



6. You should now see an Audio Check screen. **Please note: there is no audio requirement for the NAP-CC Assessment. To move to the next screen you should click the *I can hear the sound on the headphones* radio button.** The system will then present a *Start Test* button. When this is clicked, the system will present a session code screen (this is where a student will enter a test session code on test day).

   If you can see the *session code screen* (Figure 6), the installation has been successful and you can exit the LDB. You must perform this check on all machines on which the LDB is installed.

Figure 6: Test session code screen



7. Ensure you install the LDB and check its installation on all devices students will use to take the assessment.

PERFORM THE DEVICE CHECK ON STUDENT DEVICES

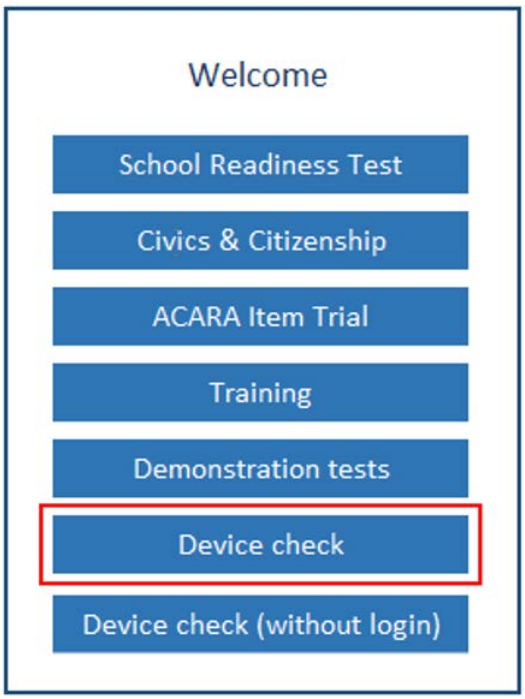
To ensure that all student devices will be able to successfully run the assessment, you must perform a device check on each machine. *Note: the platform offers a number of ways to perform a device check. For all student computers used in the NAP-CC Assessment, the device check must be performed via the locked down browser. If the check is not performed this way there is a risk that computers/devices may not be able to access the test event on test day.*

IMPORTANT NOTE: even if a machine already had the LDB installed and you did not need to download a new version, you must still perform the Device Check.

1. Launch the locked down browser.

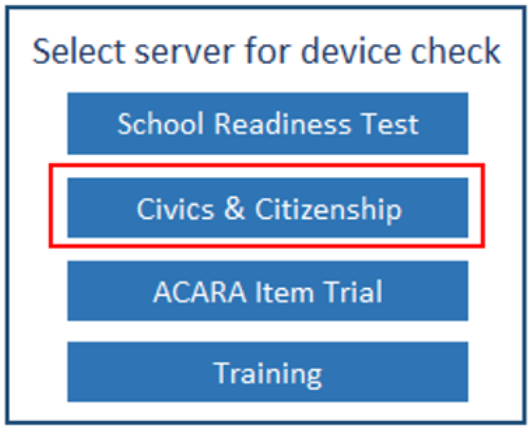
2. Select *Device check* from the list of options provided (Figure 7). Do not select *Device check (without login)*.

Figure 7: Starting the device check



3. Select *Civics & Citizenship* from the list of server options provided (Figure 8).

Figure 8: Selecting server

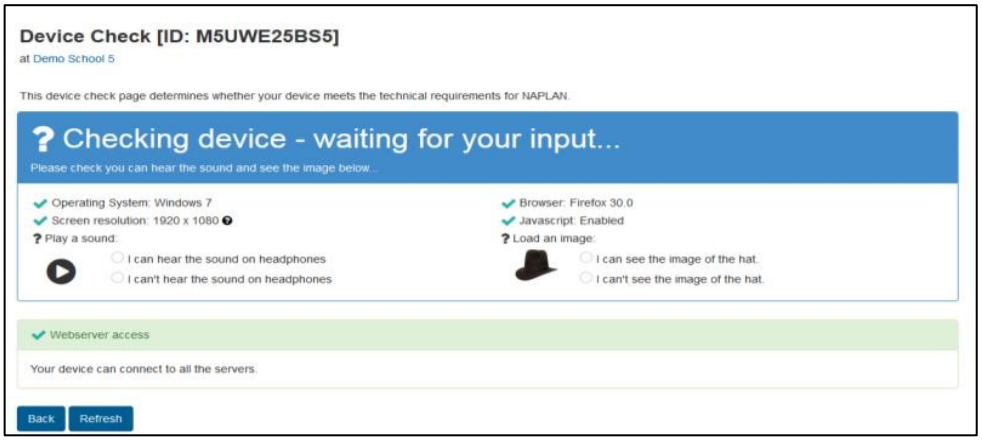


4. In the login boxes that appear (Figure 9) enter your STSO username and password (provided in the email sent to you by the NAP-CC team).

Figure 9: Device check login screen

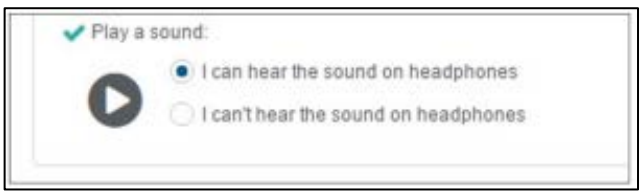
5. The Device Check will now run for one to two seconds. Once complete, you will see a screen similar to the one below (Figure 10).

Figure 10: Device check



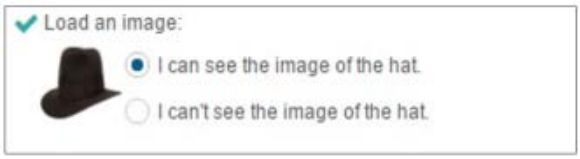
6. As previously stated, there is no audio component to the NAP-CC Assessment. However, to complete the device check please click on *I can hear the sound on headphones* (Figure 11).

Figure 11: Device check for sound



7. Indicate whether the device is able to load an image by selecting the appropriate response (Figure 12).

Figure 12: Device check for images



8. The Device Check is now complete. The device and browser you are using have been checked against the technical requirements for NAP assessments. You should see one of the two screens below (Figure 13 and Figure 14).

Figure 13: Device check pass

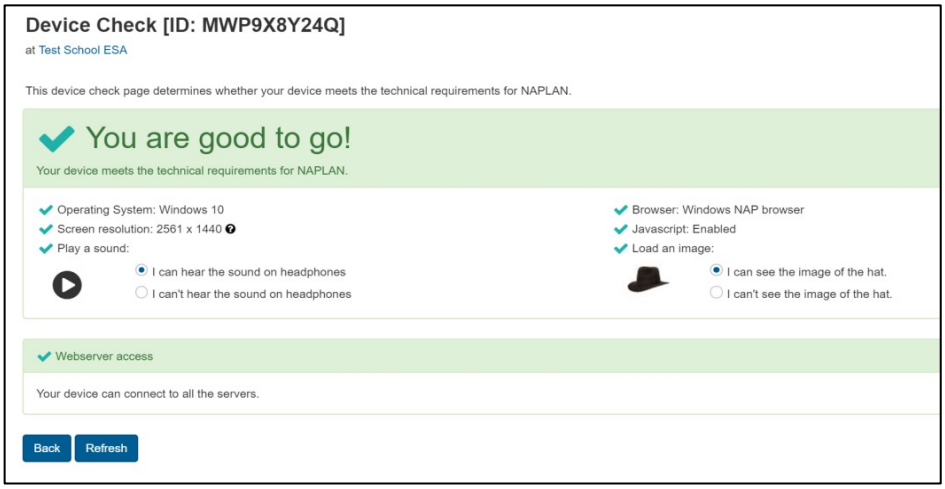
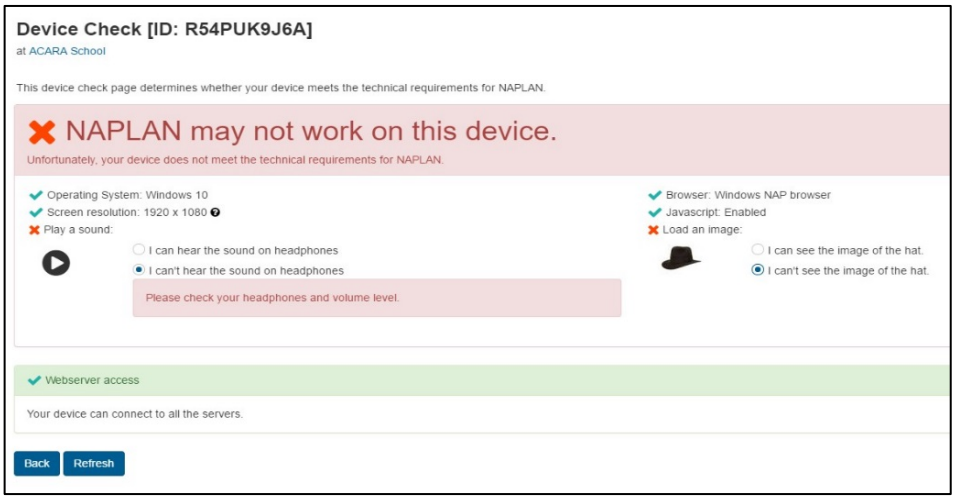


Figure 14: Device check fail



If your Device Check was successful, please proceed to step 9.

If you receive a ✘ fail against an element of the test, please see Section 3 – *Getting help to assist you in rectifying the problem*. Once the device, network or LDB has been updated, please run the device check again.

9. Click the Back button on the device check screen (Figure 15). Your result will be saved. Please note: if you exit the Device Check by using the grey X in the bottom right corner, your Device Check will not register in the Device List for your school.

Figure 15: Finishing the device check

10. Exit the app.

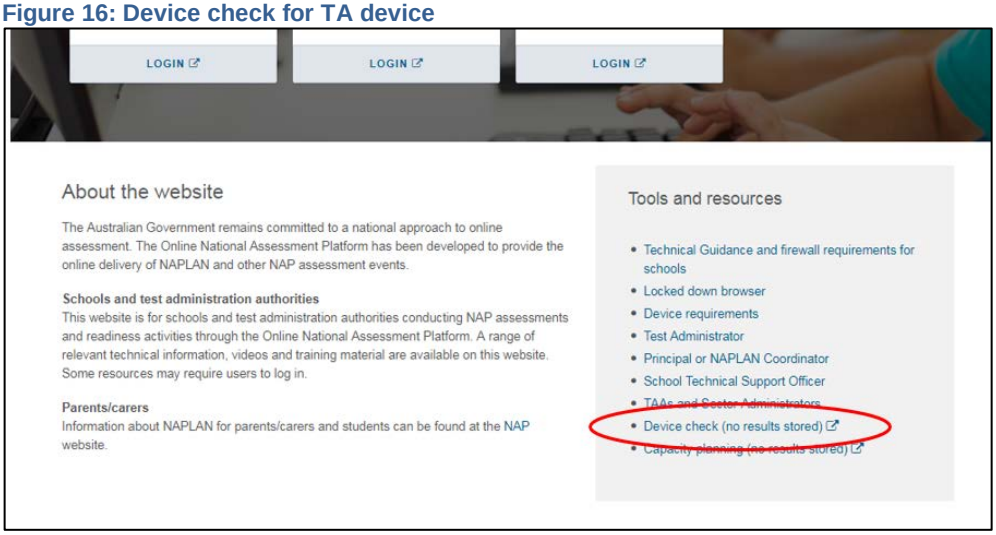
11. Repeat steps 1-10 **for every device** that will be used for the NAP-CC Assessment.

ENSURE A DEVICE FOR THE TEST ADMINISTRATOR (TA) IS PREPARED

You will need to ensure a computer has been set aside for the TA to use on assessment day. This device does not need to have the LDB installed, and the Device Check should be performed *outside* the LDB.

To run the Device Check on the TA machine, open the home page of the Assessform website assessform.edu.au and click on *Device Check (no results stored)* in the Tools and resources section on the right (Figure 16) and follow the instructions.

Figure 16: Device check for TA device



COMPLETE THE STSO TECHNICAL PREPARATIONS QUESTIONNAIRE

Once you have performed all technical readiness steps (speed test; download, install and checking of the LDB on all student devices; student device check; TA device check) please complete the 'STSO technical preparations questionnaire'. The specific link to your school's questionnaire can be found in the email that also contained your login details to the Assessform website.

Appendix A3: School Reports

NAP NATIONAL ASSESSMENT PROGRAM

acara AUSTRALIAN CURRICULUM, ASSESSMENT AND REPORTING AUTHORITY

2019 NAP-Civics and Citizenship School Summary Report: Instructional Guide

Accessing the report

1. Navigate to the school report webpage for the required year level (Year 6 or Year 10):
   - Year 6 reports: https://oars.acer.edu.au/nap-cc-2019-year-6/
   - Year 10 reports: https://oars.acer.edu.au/nap-cc-2019-year-10/

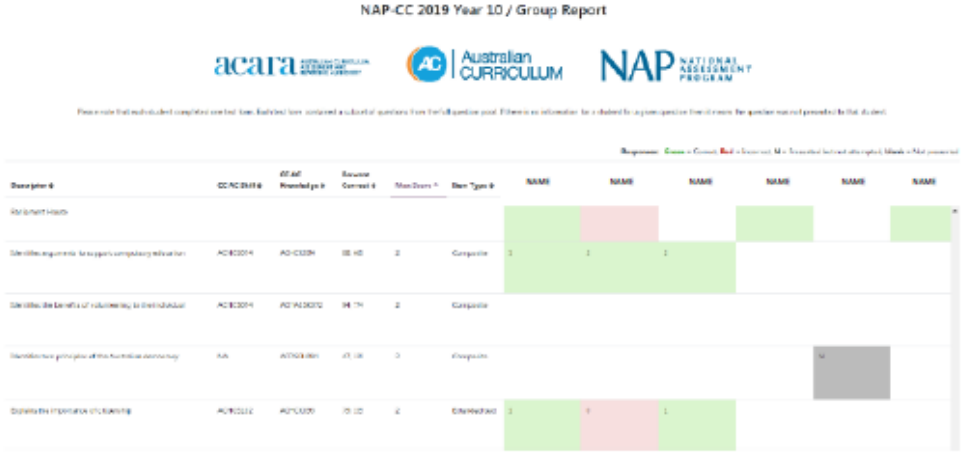
2. Enter your username and password, and then click on the green *Log in* button. Please note: your designated username is provided in the email to which these instructions were attached. Your password has been sent by separate email for security purposes.

Login page

Viewing the school report

There are two ways to view the school summary report. You can export the data as an excel file by clicking on the purple *Download* button in the top right hard corner.

Alternatively, you can view the interactive report online, as shown below. The report shows the results for all students in your school on all tasks included in the NAP-CC assessment. Moving horizontally on the screen can be achieved using the scroll bar on the bottom of the page



School report

ACER

Below is a brief description of the contents of each of the columns shown in this report.

a) **Descriptor**: This contains a brief description of what students needed to do in order to answer an item correctly. Each row refers to a single item in the assessment.

b) **CC AC Skill:** This is the Australian Curriculum Skills code for Civics and Citizenship or History, to which the item is mapped.

c) **CC AC Knowledge:** This is the Australian Curriculum Knowledge and Understanding code for Civics and Citizenship or History, to which the item has been mapped.

d) **Percent Correct**: This shows an estimate of the national percentage of students who responded to the task correctly. For tasks with a maximum score of more than 1, you will see more than one percentage. Each percentage reflects the number of students that reached each score or higher. For example, if a task has a maximum score of 2, the first number is the percentage of students that received a score of 1 out of 2, the second number is the percentage of students that received a score of 2.

e) **Max Score**: This shows the maximum score available for each task.

f) **Item type:** This denotes the type of online test type for the item. The item types were multiple choice, multiple choices, composite, interactive match, and extended text.

The scores for each task are listed under the names of each student. There are four possible displays of the score for each task:

i. Green: The student responded to the task correctly (or partially correctly). The number refers to the score the student received for their response to the task. This can be compared to the maximum score for that task.

ii. Red (0): The student responded to the task incorrectly.

iii. Grey (N): The task was assigned to that student but the student did not provide a response.

iv. Blank: The task was not in an item set assigned to that student.

The online report has a set of clickable sorting features, so you can manipulate how you would like to view the data. For example, view data ordered by percent correct, or grouped by AC Code.

Logging out

At any time you can log out of the reporting system by clicking on your numerical School Username at the top right of the screen and selecting the *Log out* option.

Appendix A4: Item Difficulties

Table A4.1: Item difficulties – Year 6 (*Equating shifts were applied to the reporting scale scores only)

			Threshold 1			Threshold 2				
Item	Scores	Horizontal link	RP =0.5	RP =0.62	CC scale (equated)	RP =0.5	RP =0.62	CC scale (equated)	Correct	Weighted fit (MNSQ)
AD_61	1	No	1.07	1.56	529				40%	1.02
AD_62	1	No	1.84	2.33	628				26%	0.89
AE_61	1	No	0.16	0.65	411				59%	0.97
AF34	1	Yes	1.29	1.78	557				36%	1.08
AL_61	1	No	-0.39	0.10	339				70%	1.15
AL_62	1	No	0.37	0.86	437				55%	1.09
AL_63	1	No	1.13	1.62	536				39%	1.10
AP21	1	Yes	-1.45	-0.96	202				85%	0.90
BA41	1	No	0.12	0.61	406				60%	0.92
BC_61	1	No	-0.25	0.24	358				67%	0.87
BD_61	1	No	1.58	2.06	594				30%	0.95
BH_61	1	No	-1.13	-0.64	244				80%	0.97
BH_62	1	No	0.98	1.47	517				41%	0.98
BT_61	1	No	1.05	1.54	526				41%	1.05
CC_61	1	No	-0.99	-0.50	261				79%	0.85
CE_61	1	No	-0.02	0.47	388				62%	0.98
CG_61	1	No	-0.20	0.29	364				65%	1.02
CG_62	1	No	0.30	0.79	429				55%	0.98
CG31	1	Yes	-0.12	0.37	375				64%	1.07
CM_61	1	No	0.39	0.88	441				55%	0.86
CM_62	2	No	-0.20	0.29	364	1.46	1.95	579	50%	0.93
CO_61	1	No	-1.15	-0.66	241				80%	1.01
CO_62	2	No	-1.57	-1.08	185	0.62	1.11	471	69%	1.02
CR_61	1	No	-0.08	0.41	380				64%	0.97
CR_62	2	No	-0.79	-0.30	288	0.67	1.15	476	65%	1.08
CV_61	1	No	-0.22	0.27	362				66%	0.94
CV32	1	Yes	-0.33	0.16	347				67%	0.96
CW_61	2	No	0.66	1.15	476	2.37	2.86	697	31%	1.01
DR0232	1	Yes	1.11	1.60	534				40%	0.98
EQ41	1	Yes	-1.84	-1.35	150				88%	1.06
ER31	1	Yes	-0.85	-0.36	280				76%	1.03
ER32	1	Yes	-0.58	-0.09	314				72%	0.91
ER33	1	No	0.19	0.68	415				57%	1.20
ES_61	1	No	-0.51	-0.02	324				71%	0.89
EX_61	1	No	0.60	1.09	467				51%	0.99
EX_62	1	No	0.33	0.82	433				56%	1.03
FS41	1	No	0.18	0.67	413				59%	1.18
FT31	1	Yes	-0.36	0.13	344				68%	0.97
FT32	1	Yes	-0.03	0.46	386				62%	1.02
FT33	1	No	1.08	1.57	530				40%	0.93
FW41	1	Yes	1.36	1.85	566				34%	1.07
FW42	1	Yes	-1.36	-0.87	213				83%	0.91
GC_61	2	No	-0.79	-0.30	287	0.68	1.17	478	64%	0.99
GS31	1	Yes	-0.01	0.48	389				62%	1.11
GS32	1	Yes	-1.03	-0.54	257				79%	1.02
GS33	1	Yes	-0.11	0.38	375				64%	1.04
HD_61	1	No	-1.72	-1.23	166				87%	0.84
HD_62	1	No	-0.58	-0.09	315				72%	0.91

			Threshold 1			Threshold 2				
Item	Scores	Horizontal link	RP =0.5	RP =0.62	CC scale (equated)	RP =0.5	RP =0.62	CC scale (equated)	Correct	Weighted fit (MNSQ)
HD_63	1	No	-0.82	-0.33	284				76%	0.96
HD_64	1	No	-1.02	-0.53	257				79%	0.94
HS_61	1	No	0.13	0.61	406				59%	0.92
HU_61	1	No	2.04	2.53	655				23%	1.04
HU_62	1	No	-0.80	-0.31	286				75%	1.00
LA_61	1	No	0.35	0.84	436				54%	1.09
LA_62	1	No	-1.85	-1.36	149				88%	0.87
LC_61	1	No	0.56	1.05	462				51%	1.12
LC_63	2	No	0.01	0.50	391	1.36	1.85	566	49%	0.93
LG0231	1	Yes	0.13	0.62	406				59%	1.11
LP_61	1	No	-0.07	0.42	381				63%	0.85
PA_61	1	No	0.84	1.33	499				45%	1.15
PL_61	1	No	-0.66	-0.17	304				74%	0.99
PM41	1	Yes	0.45	0.94	448				52%	1.02
PR_61	1	No	1.76	2.24	618				27%	1.08
PROT31a	1	Yes	-0.50	-0.01	325				71%	1.07
PROT32	1	Yes	-0.61	-0.12	310				73%	0.96
PROT33	1	No	0.95	1.44	513				43%	0.98
PROT54	1	No	-0.50	-0.01	324				71%	0.91
RE_61	1	No	0.74	1.22	485				46%	0.99
RI41	1	Yes	-1.49	-1.00	196				85%	0.92
RO_61	1	No	-0.98	-0.49	263				78%	0.96
RO_63	1	No	0.43	0.92	446				52%	0.99
RO_64	2	No	0.18	0.67	414	1.51	2.00	585	44%	1.07
RU_61	1	No	0.43	0.92	445				53%	1.06
SA_61R	1	No	-0.28	0.21	353				68%	1.00
SC_61	1	No	-0.42	0.07	336				69%	1.10
SD_61	2	No	-0.04	0.45	384	1.96	2.45	644	43%	1.06
SS_61	2	No	-0.03	0.46	385	1.85	2.34	629	43%	1.08
TL_61	1	No	1.20	1.68	545				38%	0.98
TO_61	1	No	-0.52	-0.03	323				71%	0.87
TS41	1	Yes	-0.15	0.34	370				64%	0.88
TS42	1	Yes	0.18	0.67	413				57%	0.96
TS43	1	Yes	-0.82	-0.33	284				75%	0.94
UN31	1	Yes	-0.19	0.30	366				65%	1.08
VO41	1	Yes	-0.99	-0.50	261				78%	0.90
VO42	2	No	-0.92	-0.43	270	0.60	1.09	467	65%	0.99
WA_61	1	No	-1.31	-0.82	219				83%	0.91
WL43	1	Yes	0.03	0.52	393				61%	1.05
WT_61	1	No	-0.88	-0.39	276				77%	1.02

Table A4.2: Item difficulties – Year 10 (*Equating shifts were applied to the reporting scale scores only)

			Threshold 1			Threshold 2			Threshold 3				
Item	Scores	Horizontal link	RP =0.5	RP =0.62	CC scale (equated)	RP =0.5	RP =0.62	CC scale (equated)	RP =0.5	RP =0.62	CC scale (equated)	Correct	Weighted fit (MNSQ)
AA31	1	Yes	0.87	1.36	547							50%	0.96
AA32	1	Yes	-0.02	0.47	433							67%	0.94
AA33	1	Yes	-0.71	-0.22	343							78%	0.98
AB_61	1	No	-0.52	-0.03	367							75%	1.14
AB_62	1	No	-0.18	0.31	412							70%	1.04
AB_63	2	No	0.01	0.50	436	1.29	1.78	602				55%	1.06
AC0231	1	No	0.28	0.77	471							62%	1.16
AF31	1	Yes	0.49	0.98	499							58%	1.03
AF33	1	Yes	-0.65	-0.16	351							78%	0.96
AF34	1	Yes	1.15	1.63	583							45%	1.04
AP21	1	No	-2.06	-1.57	168							92%	0.94
AT_61	1	No	-0.57	-0.08	361							77%	0.80
AT_62	1	No	0.29	0.78	472							62%	0.83
BD_61	1	No	0.61	1.10	514							55%	0.94
BD41	1	Yes	0.18	0.67	458							64%	1.11
BH_61	1	No	-1.36	-0.87	258							86%	0.91
BH_62	1	No	0.14	0.63	453							65%	0.97
BS_61	1	No	0.39	0.88	486							60%	1.03
CC_61	1	No	-1.50	-1.01	241							87%	0.87
CF_61	1	No	-1.54	-1.05	234							88%	1.06
CF_63	1	No	0.49	0.98	499							58%	1.03
CG_61	1	No	-0.13	0.36	419							70%	1.06
CG_62	1	No	-0.11	0.38	421							69%	0.95
CG31	1	No	-0.30	0.19	396							72%	1.11
CN_61	1	No	0.24	0.73	466							63%	0.95
CN_62	1	No	-0.53	-0.04	366							76%	0.83
CN_63	1	No	0.77	1.26	535							52%	1.03
CP_61	2	No	-0.95	-0.46	312	0.61	1.10	514				70%	0.94
DB21	1	Yes	-0.09	0.40	424							69%	0.98
EC_61	1	No	0.46	0.95	495							58%	1.11
EH_61	1	No	-0.33	0.16	393							72%	1.03
EH_62	1	No	0.24	0.73	466							63%	0.87
ER31	1	Yes	-1.27	-0.78	270							85%	1.03
ER32	1	Yes	-1.78	-1.29	203							89%	0.85
ER33	1	Yes	-0.56	-0.07	362							76%	1.09
EX_61	1	No	-1.04	-0.55	300							83%	0.93
EX_62	1	No	-0.29	0.20	398							72%	1.03
FT31	1	Yes	-0.77	-0.28	335							79%	1.01
FT32	1	Yes	-0.54	-0.05	364							76%	1.11
FT33	1	No	0.09	0.58	447							65%	0.99
GA_61	1	No	0.74	1.23	531							53%	1.14
GB_61	2	No	-0.49	0.00	371	1.12	1.61	580				61%	1.08
GB_62	1	No	0.33	0.82	477							61%	1.01
GB_63	1	No	0.41	0.90	488							59%	0.84
GB_64	1	No	0.61	1.10	513							55%	0.90
GS31	1	Yes	-0.05	0.44	428							68%	1.11
GS32	1	Yes	-1.52	-1.03	238							88%	1.10
GS33	1	Yes	-0.64	-0.15	352							78%	1.14
HH_61	1	No	0.88	1.37	550							49%	1.09
HP_62	1	No	3.16	3.65	845							12%	0.98
HU_61	1	No	1.41	1.89	617							39%	1.06
HU_62	1	No	-1.10	-0.61	291							83%	1.01
HV_61R	1	No	1.77	2.26	665							32%	1.05
HV_62	1	No	-0.30	0.19	396							72%	0.98
HV_63	1	No	0.93	1.42	555							49%	0.98
IC_61	2	No	0.46	0.95	494	1.15	1.64	584				52%	1.22

			Threshold 1			Threshold 2			Threshold 3				
Item	Scores	Horizontal link	RP =0.5	RP =0.62	CC scale (equated)	RP =0.5	RP =0.62	CC scale (equated)	RP =0.5	RP =0.62	CC scale (equated)	Correct	Weighted fit (MNSQ)
IC_62	1	No	0.96	1.45	560							48%	0.87
IC_63	1	No	0.98	1.47	562							48%	0.84
IE_61	1	No	-1.84	-1.35	196							90%	0.94
IE_62	1	No	-0.68	-0.19	346							78%	0.92
IE_63	1	No	1.09	1.58	576							46%	1.23
IJ21	1	Yes	-0.29	0.20	397							71%	1.17
IN_61	1	No	1.15	1.64	584							44%	1.02
IQ11	1	Yes	0.93	1.41	555							49%	1.13
IQ12	1	Yes	0.76	1.25	534							52%	1.16
IQ13	3	Yes	-0.19	0.30	410	0.82	1.31	542	2.38	2.87	744	48%	1.12
IR_61	1	No	0.68	1.17	523							54%	1.03
IT_61	1	No	-0.88	-0.39	320							80%	0.90
IT_62	1	No	0.34	0.83	478							61%	0.94
JS_61	1	No	-0.21	0.28	408							70%	0.96
JS_62	1	No	2.10	2.59	708							26%	1.06
JS_63	1	No	-0.59	-0.10	358							76%	0.78
MC_61	1	No	0.71	1.20	527							53%	0.89
MC_62R	1	No	-0.93	-0.44	314							81%	0.88
MC_63	2	No	1.27	1.76	600	2.30	2.79	733				30%	1.04
MG31	1	Yes	-0.71	-0.22	343							78%	0.90
MP31	1	Yes	-0.24	0.25	403							71%	0.81
MP32	1	Yes	-0.15	0.34	416							70%	0.85
MP34	1	Yes	-0.35	0.14	389							73%	0.92
MP35	1	Yes	-0.13	0.36	418							69%	0.93
MR_61	1	No	1.33	1.82	608							40%	1.08
MT_61	1	No	0.59	1.08	512							56%	0.87
MT_62	1	No	1.54	2.03	635							37%	0.93
NI_61	1	No	0.06	0.55	443							66%	1.05
NI_62	1	No	-0.48	0.01	372							75%	0.89
OP_61	1	No	-1.07	-0.58	296							83%	0.84
OP_62	1	No	-0.22	0.27	406							71%	0.83
PM41	1	Yes	-0.55	-0.06	363							76%	1.03
PR_61	1	No	0.02	0.51	438							67%	1.03
PROT31a	1	Yes	-1.34	-0.85	260							86%	0.97
PROT32	1	Yes	-1.21	-0.72	277							84%	1.01
PROT33	1	Yes	0.46	0.95	494							59%	0.99
PROT54	1	No	-1.12	-0.63	289							83%	0.92
PW_61	1	No	1.47	1.96	625							39%	1.10
PW_62	1	No	0.70	1.19	525							54%	1.04
RB_61	1	No	-0.18	0.31	412							70%	0.88
RC_63	2	No	-0.24	0.25	404	0.10	0.59	448				74%	1.22
RD_61	1	No	0.20	0.69	461							64%	1.18
RD_62	1	No	-0.95	-0.46	311							82%	1.01
REF1_1	1	Yes	0.32	0.81	477							61%	1.10
RF_61	1	No	-0.08	0.41	425							68%	0.84
RP31	1	Yes	0.53	1.02	504							57%	1.04
SA_61R	1	No	-0.32	0.17	393							72%	1.00
SL_61	2	No	-0.36	0.13	388	1.33	1.82	607				57%	1.07
TC_61	1	No	0.37	0.86	483							60%	1.10
TC_62	1	No	0.95	1.44	558							48%	1.06
TD41	1	Yes	-1.25	-0.76	273							85%	0.84
TD42	1	Yes	-1.04	-0.55	300							82%	0.82
TE31	1	No	-0.67	-0.18	347							78%	1.03
TE32	2	No	-0.85	-0.36	324	1.49	1.98	628				60%	1.03
TE33	1	No	-0.35	0.14	389							73%	0.99
TL_61	1	No	0.29	0.78	472							61%	0.87
TS41	1	Yes	-1.33	-0.84	263							86%	0.92
TS42	1	Yes	-0.73	-0.24	340							79%	0.91
TS43	1	Yes	-1.00	-0.51	305							82%	0.89
VO41	1	Yes	-1.45	-0.96	247							87%	0.97
VO42	2	Yes	-1.48	-0.99	243	-0.11	0.38	421				81%	1.00
WF_61	1	No	1.39	1.88	615							40%	1.05
WF_63	1	No	-1.21	-0.72	278							84%	0.93
XE_61	1	No	0.23	0.72	465							62%	1.05
XE_63	1	No	2.57	3.05	768							20%	0.97

Appendix A5: Student background variables

Table A5.1: Student background variables

Variable	Name	Values	Coding	Regressor
Adjusted school mean achievement	sch_mn	Adjusted school mean	Logits	Direct
State and Sector	State_Sector	ACT_Catholic	10000000000000000000	Direct
		NSW_Catholic	01000000000000000000	
		NT_Catholic	00100000000000000000	
		QLD_Catholic	00010000000000000000	
		SA_Catholic	00001000000000000000	
		TAS_Catholic	00000100000000000000	
		VIC_Catholic	00000010000000000000	
		WA_Catholic	00000001000000000000	
		ACT_Government	00000000100000000000	
		NSW_Government (Reference category)	00000000000000000000	
		NT_Government	00000000010000000000	
		QLD_Government	00000000001000000000	
		SA_Government	00000000000100000000	
		TAS_Government	00000000000010000000	
		VIC_Government	00000000000001000000	
		WA_Government	00000000000000100000	
		ACT_Independent	00000000000000010000	
		NSW_Independent	00000000000000001000	
		NT_Independent	00000000000000000100	
		QLD_Independent	00000000000000000010	
		SA_Independent	00000000000000000001	
		TAS_Independent	00000000000000000000100	
		VIC_Independent	00000000000000000000010	
		WA_Independent	00000000000000000000001	

Variable	Name	Values	Coding	Regressor
School Location	ASGSRemote	Major Cities of Australia (Reference category)	0000	Direct
		Inner Regional Australia	1000	
		Outer Regional Australia	0100	
		Remote Australia	0010	
		Very Remote Australia	0001	
Gender	Gender	Male	1	Direct
		Female (Reference category)	0	
Language background other than English	LBOTE	No (Reference category)	00	Direct
		Yes	10	
		Missing	01	
Indigenous Status	ATSI	No (Reference category)	00	Direct
		Yes	10	
		Missing	01	
Highest parental education	PARED	Year 9 or equivalent or below	1000000	Direct
		Year 10 or equivalent	0100000	
		Year 11 or equivalent	0010000	
		Year 12 or equivalent	0001000	
		Certificate I to IV (including Trade Certificate)	0000100	
		Advanced Diploma/Diploma	0000010	
		Bachelor Degree or above (Reference category)	0000000	
		Not stated or unknown or does not have Parent 1/2	0000001	

Variable	Name	Values	Coding	Regressor
Highest parental Occupation Group	POCC	Senior management; professionals (Reference category) Other management; associate professionals Tradespeople; skilled office, sales and service Unskilled workers; hospitality Not in paid work in the last 12 months Not stated or unknown or does not have Parent 1/2	00000 10000 01000 00100 00010 00001	Direct
Age	AGE	Value Missing	Copy, 0 Mean, 1	PCA
Civic participation at school - vote Civic participation at school - elected Civic participation at school - decisions Civic participation at school - webpage/magazine Civic participation at school - buddy Civic participation at school - community Civic participation at school - co-curricular Civic participation at school - candidate Civic participation at school - excursion	ST01Q01 ST01Q02 ST01Q03 ST01Q04 ST01Q05 ST01Q06 ST01Q07 ST01Q08 ST01Q09	Yes No This is not available at my school Missing	Three dummies for each variable with the year level mode as the reference category	PCA

Variable	Name	Values	Coding	Regressor
Civic participation in community - collecting money	ST02Q01	Yes, I have done this within the past 12 months	Three dummies for each variable with the year level mode as the reference category.	PCA
Civic participation in community - help community	ST02Q02	Yes, I have done this but not within the past 12 months		
Civic participation in community - environmental	ST02Q03	No, I have never done this	Year 10 only.	
Civic participation in community - human rights	ST02Q04	Missing		
Civic participation in community - youth organisation	ST02Q05			
Civic participation in community - animal rights	ST02Q06			
Civic communication - internet	ST03Q01	More than three times a week	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing	PCA
Civic communication - television	ST03Q02	At least once a week		
Civic communication - radio	ST03Q03	At least once a month		
Civic communication - newspaper	ST03Q04	Never or hardly ever		
Civic communication - social media	ST03Q05	Missing		
Civic communication - family	ST03Q06			
Civic communication - friends	ST03Q07			

Variable	Name	Values	Coding	Regressor
Expected participation - sign petition	ST04Q01	I would certainly do this	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing.	PCA
Expected participation - write to newspaper	ST04Q02	I would probably do this		
Expected participation - write opinion on internet	ST04Q03	I would probably not do this		
Expected participation - wear an opinion	ST04Q04	I would certainly not do this		
Expected participation - contact an MP	ST04Q05	Missing		
Expected participation - rally or march	ST04Q06			
Expected participation - collect signature	ST04Q07			
Expected participation - choose not to buy	ST04Q08			
Expected active engagement -research candidates	ST05Q01	I will certainly do this	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing Year 10 only.	PCA
Expected active engagement - help on campaign	ST05Q02	I will probably do this		
Expected active engagement - join party	ST05Q03	I will probably not do this		
Expected active engagement - join union	ST05Q04	I will certainly not do this		
Expected active engagement - be a candidate	ST05Q05	Missing		

Variable	Name	Values	Coding	Regressor
Interest in civic issues - local community	ST06Q01	Very interested	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing.	PCA
Interest in civic issues - politics	ST06Q02	Quite interested		
Interest in civic issues - social issues	ST06Q03	Not very interested		
Interest in civic issues - environmental	ST06Q04	Not interested at all		
Interest in civic issues - other countries	ST06Q05	Missing		
Interest in civic issues - global issues	ST06Q06			
Confidence to engage - discuss a conflict	ST07Q01	Very well	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing	PCA
Confidence to engage - argue an opinion	ST07Q02	Fairly well		
Confidence to engage - be a candidate	ST07Q03	Not very well		
Confidence to engage - organise a group	ST07Q04	Not at all		
Confidence to engage - write a letter	ST07Q05	Missing		
Confidence to engage - give a speech	ST07Q06			
Confidence to engage - social media	ST07Q07			
Confidence to engage - website	ST07Q08			

Variable	Name	Values	Coding	Regressor
Belief in value of action - act together	ST08Q01	Strongly agree	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing.	PCA
Belief in value of action - elected reps	ST08Q02	Agree		
Belief in value of action - student participation	ST08Q03	Disagree	ST08Q06 – Year 10 only.	
Belief in value of action - organising groups	ST08Q04	Strongly disagree		
Belief in value of action - vote school election	ST08Q05	Missing		
Belief in value of action - citizens	ST08Q06			
Belief in civic responsibility - support a party	ST09Q01	Very important	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing	PCA
Belief in civic responsibility - learn history	ST09Q02	Quite important		
Belief in civic responsibility - learn politics	ST09Q03	Not very important		
Belief in civic responsibility - learn about other countries	ST09Q04	Not important at all		
Belief in civic responsibility - discuss politics	ST09Q05	Missing		
Belief in civic responsibility - peaceful protests	ST09Q06			
Belief in civic responsibility - local community	ST09Q07			
Belief in civic responsibility - human rights	ST09Q08			
Belief in civic responsibility - environmental	ST09Q09			
Belief in civic responsibility - protect natural resources	ST09Q10			
Belief in civic responsibility - vote in elections	ST09Q11			

Variable	Name	Values	Coding	Regressor
Trust in institutions - Australian parliament	ST10Q01	Completely	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing.	PCA
Trust in institutions - state parliament	ST10Q02	Quite a lot		
Trust in institutions - local government	ST10Q03	A little		
Trust in institutions - law courts	ST10Q04	Not at all		
Trust in institutions - police	ST10Q05	Missing		
Trust in institutions - political parties	ST10Q06			
Trust in institutions - media	ST10Q07			
Trust in institutions - social media	ST10Q08			
Attitudes towards Indigenous - support traditions	ST11Q01	Strongly Agree	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing	PCA
Attitudes towards Indigenous - improve quality of life	ST11Q02	Agree		
Attitudes towards Indigenous - traditional ownership	ST11Q03	Disagree		
Attitudes towards Indigenous - learn from traditions	ST11Q04	Strongly disagree		
Attitudes towards Indigenous - learn about reconciliation	ST11Q05	Missing		

Variable	Name	Values	Coding	Regressor
Attitudes towards Diversity - keep traditions	ST12Q01	Strongly Agree	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing.	PCA
Attitudes towards Diversity - remain peaceful	ST12Q02	Agree		
Attitudes towards Diversity - benefit greatly	ST12Q03	Disagree	Year 10 only.	
Attitudes towards Diversity - all should learn	ST12Q04	Strongly disagree		
Attitudes towards Diversity - accept differences	ST12Q05	Missing		
Attitudes towards Diversity - unity easy	ST12Q06			
Attitudes towards Diversity - better place with different background	ST12Q07			
Problems affecting Australia - pollution	ST13Q01	To a large extent	Recode to 3,2,1,0; missing replaced by the year level mode; dummies for missing	PCA
Problems affecting Australia - unemployment	ST13Q02	To a moderate extent		
Problems affecting Australia - terrorism	ST13Q03	To a small extent		
Problems affecting Australia - poverty	ST13Q04	Not at all		
Problems affecting Australia - climate change	ST13Q05	Missing		
Problems affecting Australia - water shortages	ST13Q06			
Problems affecting Australia - lack of access education	ST13Q07			
Problems affecting Australia - crime	ST13Q08			
Problems affecting Australia - lack of access health	ST13Q09			
Problems affecting Australia - racism and discrimination	ST13Q10			
Problems affecting Australia - lack of cyber security and privacy	ST13Q11			

Appendix A6: NAP–CC scale proficiency level descriptions

Table A6.1: NAP-CC scale proficiency level descriptions

Proficiency level	Proficiency level description	Examples of student achievement at this level
Level 5 ≥ 795	Students working at level 5 demonstrate precise knowledge and understanding of the workings of Australian democracy and the contexts in which it has developed. In general, they evaluate civic actions and recognise the potential for ambiguity in contested civics and citizenship concepts.	Students working at level 5, for example: • understand the underlying principles of elections in which a majority government is formed, and the role independent members can play in the formation of a majority government • analyse the reasons why a specified High Court decision may have been close and understand the federal/state division of powers • explain the significance of Anzac Day and relate Anzac Day to Australian national pride and identity • analyse the potential for tension between critical citizenship and abiding by the law • recognise the historical exclusion of Indigenous Australians from the electoral process and understand the shift in the policy towards inclusion.
Level 4 665–794	Students working at level 4 recognise the interaction between governmental policies and processes, and actions of civil and civic institutions and the broader community. They explain the benefits, motivations and outcomes of institutional policies and parliamentary processes. They demonstrate familiarity with the precise discipline-specific vocabulary associated with civics and citizenship and history content and concepts, both through interpreting text and in written responses.	Students working at level 4, for example: • understand why members of parliament are required to register their financial interests • explain the conflict inherent in resisting a 'bad' law while still remaining a 'good' citizen • understand the principles that are at the heart of our democratic system and can identify their historical origins • explain wartime propaganda and its use during times of conflict • provide a plausible explanation for a perception of the lack of representation of Indigenous Australian views in the Australian democracy • explain how having citizens learn about other cultures can benefit the community through encouraging social harmony.

Proficiency level	Proficiency level description	Examples of student achievement at this level
Level 3 **535–664**	Students working at level 3 demonstrate knowledge of specific details of the Australian democracy such as election processes. They make connections between the processes and outcomes of civil and civic institutions and demonstrate awareness of the common good as a potential motivation for civic action. Students working at level 3 demonstrate awareness that civic processes can be explained and justified in relation to their broader purposes.	Students working at level 3, for example: • understand why certain processes take place on election days • understand the effectiveness of certain protest strategies • recognise features of human rights • understand civic motivation in a historical context • identify different forms of government • understand the consequences of statelessness • recognise Australia's historical ties to Britain • understand the historical context for specific government wartime programs • identify one role of the High Court • identify some of the controversy surrounding Federation • identify a group that actively represents a sector within the community • justify reasons for restrictions to free speech • identify that sites of historical significance belong to the whole community • recognise some key functions and features of the parliament such as defining the role of the speaker of the House of Representatives • identify the value of participatory decision-making processes • identify the importance in democracies for citizens to engage with issues.

Proficiency level	Proficiency level description	Examples of student achievement at this level
Level 2 **405–534**	Students working at level 2 demonstrate knowledge of core aspects of the Australian democracy. They demonstrate awareness of the connection between fundamental principles (such as fairness) and their manifestation in rules and laws. They demonstrate awareness that citizenship rights and responsibilities are collective as well as individual, and make simple evaluations of given mechanisms of civic action.	Students working at level 2, for example: • identify historical immigration policies • recognise the value of education to society • recognise the importance of certain rules for a cohesive society • understand the contribution that can be made by refugees • understand the impact of government programs for the disadvantaged • identify the countries involved in a famous battle • suggest a disadvantage of consensus decision-making • identify the role of the Prime Minister • identify the origins of the Westminster system • give a reason explaining the contribution of aid to regional security • identify a correct statement about the federal system of government • identify a purpose for the existence of public records • recognise the definition of an independent member of parliament • understand the underlying principles of a referendum • recognise that respecting the right of others to hold differing opinions is a democratic principle • identify the role of the Governor-General • recognise changes in our national identity over time • recognise why a fair society needs to be based on rules and laws • recognise the role of the voter in a representative democracy • identify one way that colonisation affected Indigenous Australian self-governance.

Proficiency level	Proficiency level description	Examples of student achievement at this level
Level 1 **275–404**	Students working at level 1 demonstrate knowledge of broad features of the Australian democracy. They recognise the cultural significance of the land to Indigenous Australians and that cultural attitudes and values can change over time. They demonstrate familiarity with simple mechanisms of community engagement and how civic actions inform and influence change.	Students working at level 1, for example: • identify the main role of the Prime Minister • understand an example of freedom of expression • understand a limitation on freedom of expression • identify the names of the two houses of the Australian parliament • understand the reason for rules related to voting results • identify a benefit of belonging to the United Nations • identify that the federal government is responsible for the defence forces • suggest a lawful civic action to influence local government decisions • suggest the motivation behind an act of ethical consumerism • identify that learning about other cultures can benefit a community • identify that members of parliament represent the people in their electorates • recognise that attitudes to immigration in Australia have changed over time • describe ways of protesting in a democracy • identify and explain a principle that supports compulsory voting in Australia • identify qualities that are necessary for civic responsibilities • recognise the principle of equity when applied to employment opportunities.

Proficiency level	Proficiency level description	Examples of student achievement at this level
Below level 1 <275	Students working at below level 1 demonstrate knowledge of the notion of fairness and recognise some basic human rights. They demonstrate familiarity with basic aspects of democratic processes and legal systems and some familiarity with generalised characteristics of Australian identity.	Students working at below level 1, for example: • identify a basic right related to work • understand the explicit commitment made by new Australian citizens • identify a basic human right • recognise that taxes are a source of government revenue • recognise that members of parliament get their jobs by being voted for in elections • connect the separation of powers to the concept of fairness in a democracy • recognise that Australians have diverse origins • identify the importance of a gesture of cultural respect • identify the notion of good citizenship potential • recognise that Australia seeks to maintain close ties with other countries in the Asia-Pacific area • recognise that some schools encourage student participation in school decision-making • describe a fundamental democratic right related to age.

Appendix A7: Percentiles of achievement

Table A7.1: Percentiles of achievement – Year 6

			5th	10th	25th	Mean - 95% CI	Mean	Mean + 95% CI	75th	90th	95th
Year 6	Aust.	2004	229	270	334	393	400	407	470	525	558
		2007	220	266	339	400	405	410	479	534	565
		2010	207	254	330	401	408	415	489	559	602
		2013	190	239	322	397	403	409	490	555	594
		2016	181	234	326	400	408	416	498	563	601
		2019	188	241	327	402	408	413	495	565	609
	NSW	2004	241	286	350	402	418	433	491	546	576
		2007	259	306	373	421	432	443	499	553	581
		2010	228	277	348	413	426	439	506	576	619
		2013	191	244	333	404	418	432	510	583	621
		2016	178	237	328	395	413	431	506	571	606
		2019	183	230	326	396	407	418	494	568	614
	Vic.	2004	257	294	357	406	417	427	482	531	561
		2007	247	292	356	408	418	429	489	536	564
		2010	234	273	347	408	422	436	497	567	610
		2013	225	271	346	410	421	432	500	559	593
		2016	202	252	335	401	415	429	501	562	595
		2019	215	267	337	404	414	424	494	564	604
	Qld	2004	212	250	310	357	371	384	437	487	516
		2007	194	239	306	363	376	390	453	512	546
		2010	172	221	300	358	374	391	456	520	561
		2013	179	223	304	371	384	397	467	531	569
		2016	175	225	319	388	401	415	489	555	598
		2019	188	249	331	401	415	428	506	578	617
	WA	2004	203	242	305	358	371	385	439	497	532
		2007	181	229	305	358	369	380	445	498	529
		2010	194	240	320	387	402	417	486	556	596
		2013	183	222	303	367	383	399	468	534	569
		2016	180	226	314	387	403	419	492	562	600
		2019	189	245	328	396	407	417	494	560	601
	SA	2004	208	248	315	365	381	398	453	505	534
		2007	198	248	318	369	385	400	454	518	554
		2010	206	252	321	383	396	408	471	542	580
		2013	177	226	303	365	379	394	461	524	562
		2016	181	229	329	392	409	426	496	562	601
		2019	146	202	292	361	377	392	465	536	576
	Tas.	2004	210	256	327	378	393	408	466	519	551
		2007	201	242	323	383	401	419	481	546	580
		2010	197	249	331	396	411	425	495	570	613
		2013	182	225	307	370	383	396	465	522	557
		2016	183	231	315	384	400	416	484	552	591
		2019	153	201	295	373	385	398	477	553	595
	ACT	2004	243	290	361	412	423	434	494	543	574
		2007	246	288	357	405	425	446	499	558	584
		2010	252	297	364	425	442	458	522	585	625
		2013	236	289	369	418	433	447	507	561	594
		2016	213	269	351	410	426	442	509	573	605
		2019	226	282	367	425	444	464	530	597	629
	NT	2004	187	227	299	354	371	388	448	506	534
		2007	-131	-46	145	233	266	299	418	489	533
		2010	62	122	217	285	316	347	431	497	531
		2013	85	148	224	288	314	341	410	479	517
		2016	-17	45	145	269	302	335	442	513	549
		2019*	45	132	263	315	348	380	454	524	559

**The sample requirements were not achieved in the Northern Territory for Year 6. This may have resulted in a less representative sample and biased results. Therefore, their results should be interpreted with caution. More details can be found at the end of Chapter 1 and in the technical report.*

Table A7.2: Percentiles of achievement – Year 10

			5th	10th	25th	Mean - 95% CI	Mean	Mean + 95% CI	75th	90th	95th
Year 10	Aust.	2004	289	345	428	489	496	503	575	631	664
		2007	295	345	429	493	502	510	585	646	681
		2010	278	339	436	508	519	530	614	679	716
		2013	305	354	434	505	511	518	593	660	699
		2016	260	320	411	484	491	498	579	652	695
		2019	243	303	403	482	488	495	582	657	702
	NSW	2004	337	381	457	511	521	532	594	648	679
		2007	311	361	456	512	529	546	618	679	714
		2010	319	380	479	534	558	582	652	711	744
		2013	336	382	460	520	535	550	614	681	721
		2016	299	350	429	496	509	522	591	669	713
		2019	254	320	410	483	500	516	593	682	724
	Vic.	2004	284	338	424	475	494	513	577	634	665
		2007	288	337	424	477	494	511	577	634	665
		2010	292	350	443	495	514	533	597	657	690
		2013	318	368	443	507	521	535	599	666	709
		2016	248	309	408	474	489	504	579	650	693
		2019	240	301	407	471	485	499	577	640	677
	Qld	2004	259	318	400	452	469	487	549	602	635
		2007	298	341	415	467	481	495	554	610	641
		2010	225	287	390	454	482	511	586	652	685
		2013	290	334	408	472	484	496	564	624	664
		2016	251	309	392	452	471	491	559	628	666
		2019	228	279	392	461	476	490	575	639	681
	WA	2004	270	334	420	469	486	504	567	620	653
		2007	262	320	405	455	478	500	558	617	651
		2010	266	333	427	488	509	530	603	675	714
		2013	297	354	430	495	510	524	595	657	695
		2016	248	317	419	481	501	522	594	663	700
		2019	269	326	420	493	511	529	609	682	723
	SA	2004	242	307	401	449	465	481	546	597	624
		2007	304	358	443	481	505	528	581	639	673
		2010	284	328	412	469	487	506	571	640	679
		2013	274	326	408	470	486	503	571	638	673
		2016	237	300	406	461	476	492	561	629	669
		2019	254	302	386	450	466	482	550	620	671
	Tas.	2004	279	334	421	472	489	505	569	624	658
		2007	258	310	400	468	484	500	575	636	674
		2010	280	330	411	477	492	507	581	646	681
		2013	238	294	384	445	466	487	559	617	651
		2016	225	276	372	442	463	484	557	630	675
		2019†	176	220	321	400	428	456	537	609	653
	ACT	2004	305	370	452	497	518	540	595	654	687
		2007	285	358	458	504	523	543	608	669	703
		2010	298	358	444	499	523	547	613	673	702
		2013	317	376	458	511	525	539	599	677	720
		2016	294	345	437	502	518	534	603	682	722
		2019	285	352	449	509	525	541	617	680	719
	NT	2004	285	345	420	457	490	524	570	635	668
		2007	165	288	408	426	464	502	553	619	649
		2010	204	285	394	451	483	516	598	642	720
		2013	156	200	341	394	418	442	515	581	619
		2016	186	222	336	399	427	455	529	596	644
		2019†	207	289	391	443	460	477	546	611	655

†The sample requirements were not achieved in Tasmania and the Northern Territory for Year 10. This may have resulted in a less representative sample and biased results. Therefore, their results should be interpreted with caution. More details can be found at the end of Chapter 1 and in the technical report.

REFERENCES

ACARA (2018), *NAP Civics and Citizenship Assessment Framework*, https://www.nap.edu.au/docs/default-source/default-document-library/2019-nap-civics-and-citizenship-(cc)-assessment-framework-v1-1_december-2018_ed.pdf?sfvrsn=2 (accessed 5 June 2020)

ACARA (2019), *Measurement Framework for Schooling in Australia 2019*, https://www.acara.edu.au/docs/default-source/default-document-library/measurement-framework-for-schooling-in-australia-2019773213404c94637ead88ff00003e0139.pdf?sfvrsn=0 (accessed 20 June 2020).

ACARA (2020), *National Assessment Program – Civics and Citizenship 6 & 10 2019 report. Sydney, Australia: Australian Curriculum, Assessment and Reporting Authority*.

Adams, R. J., Wu, M. L., & Wilson, M. R. (2015). *ACER ConQuest: Generalised Item Response Modelling Software* [Computer software]. Version 4. Camberwell, Victoria: Australian Council for Educational Research.

Adams, R. J., Wu, M. L., Macaskill, G., Haldane, S. A., Sun, X. X. & Cloney, D. (2020). *ACER ConQuest Version 5.7.3: Generalised item response modelling software* [Computer software]. Melbourne: ACER.

Adams, R. J., & Wu, M. L. (2002). *PISA 2000 Technical Report*. Paris: OECD.

Gebhardt, E., Fraillon, J., Wernert, N., & Schulz, W. (2011). *National Assessment Program : Civics and Citizenship Education 2010 : year 6 and year 10 : technical report.* Australian Curriculum, Assessment and Reporting Authority (ACARA). https://research.acer.edu.au/civics/12.

Gonzalez, E. J., & Foy, P. (2000). 'Estimation of sampling variance.' In M.O. Martin, K.D. Gregory & S.E. Semler (eds), *TIMSS 1999 Technical Report*. Chestnut Hill, MA: Boston College.

Martin, M. O., Mullis, I. V., & Hooper, M. (2016). Methods and procedures in TIMSS 2015. *Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College*.

Masters, G. N., & Wright, B. D. (1997). 'The partial credit model.' In W.J. Van der Linden & R.K. Hambleton (eds), *Handbook of Modern Item Response Theory*, pp. 101–122. New York/Berlin/Heidelberg: Springer.

Mislevy, R. J. (1991). 'Randomization-based inference about latent variables from complex samples'. *Psychometrika*, 56, 177-196.

Mislevy, R. J., & Sheehan, K. M. (1987). 'Marginal estimation procedures'. In: A.E. Beaton (Ed.), *The NAEP 1983-1984 Technical Report*, 293-360. Princeton, NJ: Educational Testing Service.

OECD (2005). *PISA 2003 Technical Report*. Paris: OECD.

OECD (2009a). *PISA 2006 Technical Report*. Paris: OECD.

OECD (2009b). *PISA Data Analysis Manual SPSS*® (2nd edn). Paris: OECD.

OECD (2014). *PISA 2012 Technical Report*. Paris: OECD.

PMRT (2003). *Setting National Standards*. Paper presented at the March 2003 meeting of the Performance Measurement and Reporting Taskforce.

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen: Nielsen and Lydiche.

Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). 'What are plausible values and why are they useful?' *IERI Monograph Series,* (Vol. 2, pp 9-36). Hamburg and Princeton: IER Institute and ETS.

Warm, T. A. (1989). 'Weighted likelihood estimation of ability in Item Response Theory', *Psychometrika*, vol. 54, pp. 427–450.

Wolter, K. M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.