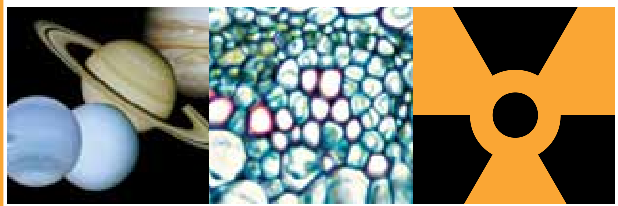


National Assessment Program – Science Literacy Year 6 Report

2012

acara

AUSTRALIAN CURRICULUM, ASSESSMENT AND REPORTING AUTHORITY

NAP–SL 2012 Project Staff

Dr Sofia Kesidou from Educational Assessment Australia (EAA) was the Project Director of NAP–SL 2012. Jennifer Cowing (EAA) was the Project Manager. The test development team was led by Joe Merlino (EAA). The Public Report was written by Dr Sofia Kesidou, Dr Rassoul Sadeghi and Nicholas Marosszeky. The School Release Materials were written by Dr Sofia Kesidou, Jennifer Cowing, Joe Merlino and Dr Rassoul Sadeghi (EAA).

The sampling and data analysis tasks were undertaken by Dr Rassoul Sadeghi, Nicholas Marosszeky and Dr Edward Li from EAA and Dr Margaret Wu and Mark Dulhunty from Educational Measurement Solutions (EMS). The Technical Report was written by Dr Rassoul Sadeghi, Nicholas Marosszeky and Dr Sofia Kesidou from EAA and Dr Margaret Wu and Mark Dulhunty from EMS.

Cover Photographs
Second and third from top: courtesy of Auburn West Public School and Our Lady of Fatima, Catholic Primary School, © Educational Assessment Australia

ACARA Copyright Administration, ACARA
Level 10, 255 Pitt Street
Sydney
NSW 2000
Email: info@acara.edu.au

Contents

List of Tables

List of Figures

Acknowledgements

National Assessment Program – Science Literacy Review Committee Members

The Science Literacy Review Committee (SLRC) was established in 2011 to support the development and implementation of the project. The contribution of the following jurisdictional representatives, specialists and others who assisted with the project is acknowledged and valued.

ACARA Chair	Michelle Robins
State and territory representatives	
Australian Capital Territory	Annie Termaat Sharon Fellows
New South Wales	Jane West
Northern Territory	Louise Nicholas Bindi Isis
Queensland	Michael Chamberlain
South Australia	Jan Brooks
Tasmania	Andrew Oakley
Victoria	Maria James
Western Australia	Marilyn McKee
ACARA educational measurement specialists	Goran Lazendic Andrew Smith Jennifer Lau
Language background other than English specialist	Hanya Stefaniuk
Indigenous education specialist	Joseph Sambono
Australian Government representative	Robyn MacMillan
Catholic school sector representative	Michael Doyle
Independent school sector representative	Glenda Leslie
Consultant	Mark Hackling
ACARA Executive Officers	Jennifer Lau Michael Boyle
Educational Assessment Australia (EAA)	Sofia Kesidou
EAA educational measurement specialist	Rassoul Sadeghi
Educational Measurement Solutions	Margaret Wu Mark Dulhunty

Foreword

The National Assessment Program – Science Literacy (NAP–SL) is one of a suite of national sample assessments (with Civics and Citizenship and Information and Communication Technology [ICT] Literacy) that are developed and managed by the Australian Curriculum, Assessment and Reporting Authority (ACARA). Conducted with a random sample of students on three-yearly cycles, these assessments are carried out under the auspices of the Standing Council on School Education and Early Childhood (SCSEEC). The assessments support measurement and reporting on progress towards the achievement of the objectives outlined in the *Melbourne Declaration on Educational Goals for Young Australians 2008*. They monitor the extent to which our schooling promotes equity and excellence; and the progress of young Australians towards becoming successful learners, confident and creative individuals, active and informed citizens.

The findings from the 2012 National Assessment Program – Science Literacy (NAP–SL) presented in this report, provide valuable information on the science literacy of Year 6 students in Australia. The report provides a national comparison against the science literacy scale and an analysis of findings across states and territories. The achievement data provide insight into the level of science-based knowledge, understandings and skills that our Year 6 students have developed. Together with these data, the student survey results highlight the extent of student interest in science, their engagement in science related activities and their understanding of how science is relevant in their lives.

This is the fourth report on Year 6 science literacy, following earlier ones in 2003, 2006 and 2009. The results over the four surveys show no change in national performance levels, in terms of both average student achievement and the proportion of students performing at or above the defined Proficient Standard in scientific literacy. The Proficient Standard is set at a challenging level and only just over half (51.4%) of Year 6 students reach or exceed it.

The challenges of international comparisons of students' performance in science learning have caused Australia to adopt a goal to be in the top five countries by 2025. The most recent results from the Trends in International Mathematics and Science Survey (TIMSS) conducted by the International Association for the Evaluation for Educational Achievement (IEA) show that, at Year 4, 18 countries perform on average significantly better than Australia (Martin, Mullis, Foy & Stanco, 2012, p.42) while at Year 8, nine countries do (Martin *et al*, 2012, p.44). Australia's results at Year 4 had declined significantly from those achieved in the first TIMSS in 1995 (Martin *et al*, 2012, p.50). At Year 8, where the 2012 results were stronger, there had been no change from the first results in 1995 (Martin *et al*, 2012, p.54).

Results from the Programme for International Student Assessment (PISA) conducted by the Organisation for Economic Co-operation and Development (OECD) among 15-year-olds provides a similar picture to the TIMSS results for Year 8 students. Only six countries were significantly ahead of Australia in science in PISA in 2009 (OECD, 2010b, p.151) and there had been no decline from earlier years (OECD, 2010a, p.64).

These international comparisons provide some comfort about science achievement at secondary level though no evidence of improvement. They suggest, however, that Australia could do much better at primary school. We clearly need more than 51.4% of Year 6 students at or above the Proficient Standard in our domestic NAP–SL if we are to see our results moving up, instead of remaining stationary or even declining.

The report also provides analyses of the achievement of various sub-groups of students including Indigenous students, those living in remote and very remote areas, and those from non-English speaking backgrounds. Teachers and schools focused on promoting science education look for innovative ways to engage their students in learning about scientific concepts, and in investigating and analysing data so that they can make informed decisions in a world in which science and technology are increasingly shaping the lives of young Australians. What the student survey results show is that the great majority of students (over 80%) appear to be interested in learning new things in science, learning about science and doing science-based activities. This is a strong foundation on which to build student awareness of the importance of science in their everyday lives, build confidence, inspire excellence and encourage students to consider rewarding future careers in the field of science!

I take this opportunity to acknowledge the collaboration and dedication of senior educators across all states and territories and all sectors of Australian schooling who have contributed to the development of the 2012 NAP–SL Sample Assessment. ACARA acknowledges the work of the Science Literacy Review Committee in the development and implementation of this assessment. ACARA also thanks the principals, teachers and students at government, Catholic and independent schools around Australia who through their participation in the assessment provided valuable information about science literacy in schools.

I commend this report to teachers, educators, and community members engaged in the collective responsibility of achieving improved educational outcomes for all young Australians, and to those with a specific interest in developing young Australians who can confidently participate in a society which is increasingly dependent on science.

Professor Barry McGaw AO
Chair
Australian Curriculum, Assessment and Reporting Authority Board

References

Martin, M.O., Mullis, I.V.S., Foy, P. & Stanco, G.M. (2012) *TIMSS international results in science*. Amsterdam: IEA and Chestnut Hill, MA, USA: TIMSS and PIRLS International Study Center.

OECD. (2010a). *PISA 2009 results: learning trends – changes in student performance since 2000*. Vol V. Paris: Author.

OECD. (2010b). *PISA 2009 results: what students know and can do – student performance in reading mathematics and science. Vol. I.* Paris: Author.

Executive Summary

The National Assessment Program – Science Literacy (NAP–SL) assesses scientific literacy in the context of a student's ability to apply broad conceptual understandings of science in order to make sense of the world, to understand natural phenomena and interpret media reports about scientific issues. It also includes asking investigable questions, conducting investigations, collecting and interpreting data and making decisions. This construct evolved from the definition of scientific literacy used by the Organisation for Economic Co-operation and Development (OECD) – Programme for International Student Assessment (PISA):

> *... the capacity to use scientific knowledge, to identify questions and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity.*
>
> (OECD 1999, p. 60)

NAP–SL is one of a suite of three national sample assessments (with Civics and Citizenship, and ICT Literacy) which are conducted with random samples of students in three-year cycles. The results contribute to an understanding of students' progress towards the achievement of the Educational Goals for Young Australians specified in the Melbourne Declaration.

In July 2001, the Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA, now the Standing Council on School Education and Early Childhood [SCSEEC[1]]) agreed to the development of assessment instruments and key performance measures for reporting on student skills, knowledge and understandings in primary science. The development and implementation of this national assessment in science literacy is undertaken by the Australian Curriculum, Assessment and Reporting Authority (ACARA).

1 SCSEEC was previously known as the Ministerial Council for Education, Early Childhood and Youth Affairs (MCEECDYA) and, prior to that, the Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA).

The first Science Literacy assessment was conducted in 2003. The assessment is repeated with a new sample of Year 6 students every three years in order to identify trends over time. The findings in this report describe the scientific literacy of Year 6 students from the latest 2012 assessment, with comparisons made to the 2003, 2006 and 2009 cohorts.

ACARA established a national science literacy committee known as the Science Literacy Review Committee (SLRC) to advise it on critical aspects of the assessment program and ensure that the assessments and results were valid across the states and territories. The main function of the science literacy committee was to ensure that the scientific literacy assessment domain was inclusive of the different state and territory curricula and that the items comprising the assessments were fair for all students, irrespective of where they attended school.

Assessment domain

The scientific literacy assessment domain was developed in the first assessment cycle in consultation with curriculum experts from each state and territory and representatives from the Catholic and independent school sectors.

The assessment domain outlines the development of scientific literacy across three main strands:

Strand A: formulating or identifying investigable questions and hypotheses, planning investigations and collecting evidence.

Strand B: interpreting evidence and drawing conclusions from students' own or others' data, critiquing the trustworthiness of evidence and claims made by others, and communicating findings.

Strand C: using science understandings for describing and explaining natural phenomena, and for interpreting reports about phenomena.

The assessment instruments draw on four concept areas:

- Earth and Space
- Energy and Force
- Living Things
- Matter.

The 2012 NAP–SL assessment instrument also drew on the concepts and content of the Statements of Learning in Science, which were endorsed in 2006 by ministers of education in all states and territories. Future NAP–SL tests will also draw on the Australian Curriculum: Science.

Assessment instrument

The assessment instrument was administered to a random sample consisting of five per cent of the total Australian Year 6 student population. The students' regular classroom teachers administered NAP–SL on the following dates:

- 17 October 2012 – New South Wales, Northern Territory, Queensland, South Australia, Tasmania and Victoria
- 24 October 2012 – Australian Capital Territory and Western Australia.

The assessment instrument consisted of seven pencil-and-paper tests, including multiple-choice and open-ended items, and two practical tasks. Each student completed one of the pencil-and-paper tests and one of the practical tasks. Students were allowed 60 minutes for the pencil-and-paper tests and 45 minutes for the practical task. The practical tasks required the students to conduct an experiment in groups of three and then respond individually to a set of questions about the experiment. Students also completed a 34-item Student Survey which sought to gather information about students' perceptions of and attitudes to science and their experiences of science learning at their school.

Student performance in scientific literacy

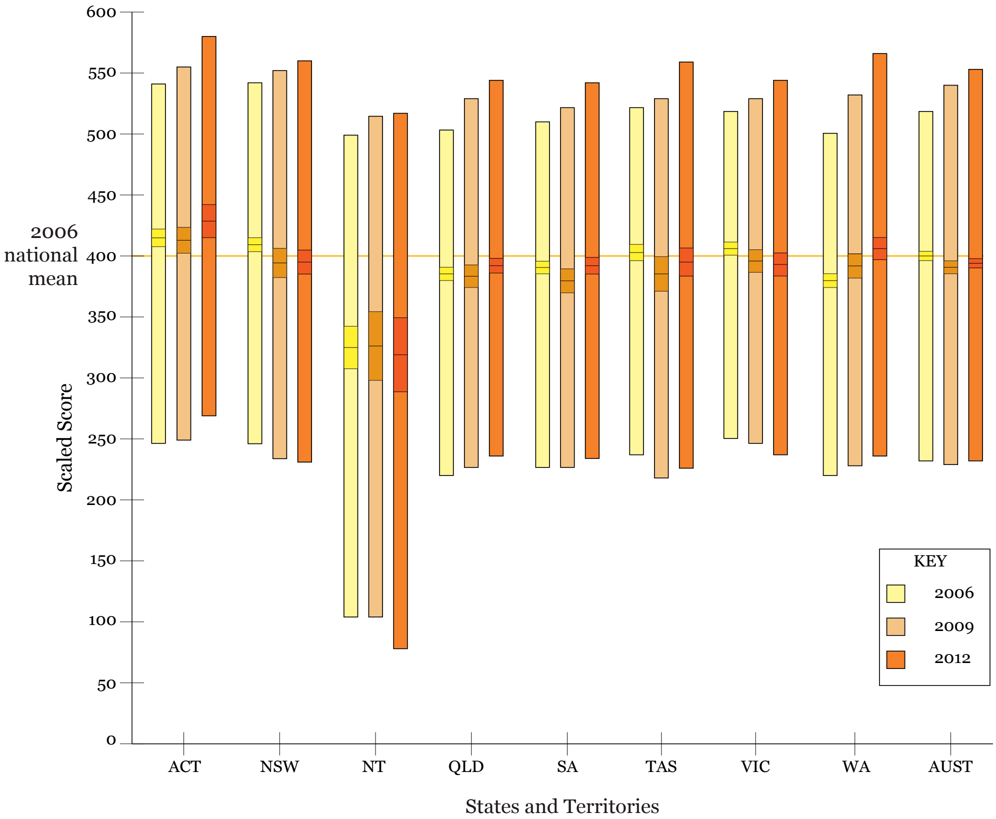
One of the main objectives of NAP–SL is to monitor trends over time. The scientific literacy scale was initially established in 2003. However, in 2006, a more robust test design was implemented. The sample was designed to be more inclusive of remote schools and included items that catered for a greater range of student ability. Consequently, the results of the 2006 assessment were used to establish a new baseline scientific literacy scale and the 2003 results were re-scaled onto it.

The 2012 student performance has also been scaled to the 2006 baseline. As a result, the majority of the trend commentary in this report is based on comparisons between the 2006, 2009 and 2012 assessments.

In order to test the statistical differences between the 2006 and 2012 results, and those of 2009 and 2012, it is necessary to take into account the variability in the data caused when the 2012 results are equated to the 2006 and 2009 scales.

Figure ES.1 shows the mean scores and distributions for 2006, 2009 and 2012.

Figure ES.1 Comparison of distributions of student scores by state and territory in 2006, 2009 and 2012



Notes: 2012 and 2009 results scaled to 2006.

This figure contains graphical comparisons of the student results (the upper and lower 5 per cent of the distribution are excluded due to the large variability associated with the extreme scores at the two ends of the scale). Shaded bands around the mean within each bar mark the 95 per cent confidence interval.

The 2012 results are similar to those of 2006 and 2009 both in terms of mean student achievement and the distribution of student scores at the national level. There is no statistically significant difference between the 2009 and 2012 results at the national level or across Australian states and territories. Similarly, the comparison between 2006 and 2012 results shows that there is no statistically significant difference for most jurisdictions and at the national level. The only exception is Western Australia, where the mean student achievement increased by 25 score points from 2006 levels and this was found to be statistically significant.

At the national level, a comparison of mean achievement between student groups showed the following results:

- For males and females, there were no significant differences in mean achievement. These results are consistent with those obtained in 2006 and 2009.

- Indigenous students had a statistically significant lower mean achievement than non-Indigenous students (see Table ES.1). These results are consistent with the 2006 and 2009 results.

- Students in both provincial areas and remote and very remote areas had a statistically significant lower mean achievement than students attending schools in metropolitan areas (see Table ES.2). This contrasts with the 2006 and 2009 results which showed that students in remote and very remote areas had a statistically significant lower mean achievement than students attending schools in both metropolitan and provincial areas.

Table ES.1 Mean scores for Indigenous and non-Indigenous students in 2012

Student group	Mean score	95 per cent CI
Indigenous	303	±15.1
Non-Indigenous	399	±4.5

Table ES.2 Mean scores of students by geographic location[2] category in 2012

MCEECDYA geographic location category	Percentage of students	Mean score
Metropolitan areas	72.9	400 (±5.2)
Provincial areas	25.3	381 (±9.5)
Remote and very remote areas	1.9	349 (±31.0)
AUST	100.0	394 (±4.4)

Note: Figures in parentheses refer to 95 per cent confidence intervals.
The percentages of students in Table ES.2 are weighted to reflect the population of Year 6 students in Australia. They are not the percentages of students in the sample[3].

Since 2009, students have also been required to complete a survey. The survey canvassed students' perceptions of and attitudes to science. It also asked students about their science learning experiences at school. Analyses showed weak to moderate correlations between survey items and achievement. Survey items related to students' 'self-concept in science' (i.e. the level of belief that students have in their science competencies) had the highest correlations with test performance. That is, a higher science self-concept was associated with a higher mark in scientific literacy. Interestingly, approximately 80 per cent of students responded that they would like to learn more science at school, indicating that a positive attitude towards this subject area exists.

2 In this report, 'geographic location' refers to whether a student attended school in a metropolitan, provincial or remote and very remote zone (Jones, 2004).
- Metropolitan zones include mainland state capital cities and major urban areas with population above 100 000, such as Hobart, Geelong, Wollongong and the Gold Coast.
- Provincial zones include provincial cities (including Darwin) and provincial areas.
- Remote and very remote zones include areas of low accessibility, such as Katherine and Coober Pedy.

3 All other statistics in this report are also weighted. For more information about the applied weights and the sampling design refer to the 2012 NAP–SL Technical Report.

Distribution of students across Proficiency Levels

One of the key objectives of the National Assessment Program is to monitor trends in scientific literacy performance over time. One convenient and informative way of describing student performance is to reference the results to Proficiency Levels. A similar process is used in several large-scale assessment programs including PISA.

Students whose results are located within a particular level of proficiency are typically able to demonstrate the understandings and skills associated with that level, and also typically possess the understandings and skills defined as applying at lower proficiency levels.

To establish the Proficiency Levels in the 2003 assessment cycle, a combination of experts' knowledge of the understandings and skills required to answer each scientific literacy item, and information from the analysis of students' responses was used. This defined five Proficiency Levels for reporting student performance from the assessment.

The standard for Year 6 scientific literacy was also established in the 2003 assessment cycle to provide parents, educators and the community with a clear picture of what students should know and be able to do by the end of Year 6.

To set the standard, an expert group comprising university science educators, curriculum officers and experienced primary teachers from all states and territories, from government, Catholic and independent schools, was brought together. The crucial scientific literacy skills and understandings needed by students for the next phase of science learning at school were discussed and debated before consensus was reached on a Proficient Standard for Year 6. The Proficient Standard was deemed to be Level 3.2 on the Proficiency Level continuum. This standard informed the development of the tests for the 2006, 2009 and 2012 assessments.

The Proficient Standard is a challenging level of performance, with students needing to demonstrate more than minimal or elementary skills to be regarded as reaching it. It is one of several achievement levels that collectively represent a continuum of learning and describe what students know and are able to do. Students who do not achieve the Proficient Standard demonstrate only partial mastery of the skills and understandings expected for Year 6. There are also students who show superior results and exceed the Proficient Standard.

Table ES.3 shows a comparison at the jurisdiction level of the proportion of students in each of the Proficiency Levels and the proportion of students performing at or above the Proficient Standard in 2012.

Table ES.3 Percentage of students in Proficiency Levels by state and territory in 2012

State/ Territory	Level 2 and below	Level 3.1	Level 3.2*	Level 3.3	Level 4 and above	At or above the Proficient Standard
ACT	4.4 (±1.7)	30.3 (±4.6)	49.4 (±4.2)	15.0 (±4.1)	0.9 (±0.7)	65.3 (±5.3)
NSW	9.2 (±2.5)	39.8 (±3.3)	40.9 (±3.8)	9.6 (±2.5)	0.4 (±0.4)	50.9 (±4.3)
NT	31.1 (±9.6)	37.9 (±7.0)	26.5 (±6.4)	4.3 (±3.0)	0.2 (±0.4)	31.0 (±7.6)
QLD	8.8 (±1.6)	41.4 (±2.9)	41.8 (±3.1)	8.0 (±1.6)	0.1 (±0.2)	49.9 (±3.3)
SA	8.8 (±1.9)	40.1 (±3.4)	43.5 (±3.6)	7.5 (±1.9)	0.1 (±0.2)	51.1 (±3.9)
TAS	9.6 (±2.3)	39.1 (±4.2)	40.2 (±4.2)	10.8 (±3.2)	0.3 (±0.5)	51.3 (±5.4)
VIC	8.3 (±2.2)	40.4 (±4.0)	43.4 (±3.8)	7.6 (±2.3)	0.2 (±0.3)	51.3 (±4.7)
WA	8.2 (±1.9)	35.5 (±3.3)	44.0 (±3.3)	12.0 (±2.4)	0.4 (±0.4)	56.4 (±4.2)
AUST	9.0 (±1.0)	39.6 (±1.6)	42.1 (±1.7)	9.0 (±1.1)	0.3 (±0.2)	51.4 (±2.0)

Notes: *The Proficient Standard has been set at Proficiency Level 3.2.
Figures in parentheses refer to 95 per cent confidence intervals.

Table ES.3 shows that the Australian Capital Territory has the highest proportion of students attaining the Proficient Standard, i.e. operating at or above Proficiency Level 3.2. The smallest proportion of such students was observed in the Northern Territory. New South Wales, Victoria, Queensland, South Australia, Western Australia and Tasmania have similar percentage distributions across Proficiency Levels and at the Proficient Standard.

In 2012, 51.4 per cent of students at the national level attained the Proficient Standard or better in scientific literacy. Table ES.4 shows a comparison at the national level of the proportion of students in each of the Proficiency Levels and the proportion of students performing at or above the Proficient Standard in 2006, 2009 and 2012[4].

4 Results from the 2003 NAP–SL assessment are not included in this Executive Summary Table. These results cannot be reliably compared with the results of the following cycles as there were important differences between the test design in the 2003 assessment and those in the 2006, 2009 and 2012 assessments. The assessments in 2006, 2009 and 2012 included wider coverage of the assessment domain and samples were more inclusive of students in remote geographic locations.

Table ES.4 Percentage distribution of students across Proficiency Levels in 2006, 2009 and 2012

AUST	Level 2 and below	Level 3.1	Level 3.2*	Level 3.3	Level 4 and above	At or above the Proficient Standard
2006	8.6 (±1.1)	37.1 (±1.7)	44.2 (±1.8)	9.6 (±1.2)	0.5 (±0.4)	54.3 (±2.1)
2009	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)	51.9 (±2.2)
2012	9.0 (±1.0)	39.6 (±1.6)	42.1 (±1.7)	9.0 (±1.1)	0.3 (±0.2)	51.4 (±2.0)

Notes: *The Proficient Standard has been set at Proficiency Level 3.2.
Figures in parentheses refer to 95 per cent confidence intervals.

Table ES.4 shows that at the national level, the difference between 2006 and 2012 in the proportion of students performing at or above the Proficient Standard is 2.9 per cent. The corresponding difference between 2009 and 2012 is 0.5 per cent. These differences were not statistically significant. Similarly, there was no significant difference in the distribution of students across Proficiency Levels at the national level across the assessments in 2006, 2009 and 2012.

In conclusion, results of the 2012 NAP–SL assessment at the national level remained the same as those observed in the previous assessment cycles, both in terms of mean student achievement and the proportion of students performing at or above the Proficient Standard in scientific literacy. Similarly, the comparison between 2006, 2009 and 2012 results shows that there are no statistically significant differences in mean student achievement and in the proportion of students performing at or above the Proficient Standard in most jurisdictions. The only exception is Western Australia, where the mean student achievement increased by 25 score points and the proportion of students performing at or above the Proficient Standard increased by 9.8 percentage points from 2006 levels. These differences were found to be statistically significant.

Chapter 1
Overview of the National Assessment

Introduction

In 1999, the state, territory and Commonwealth ministers of education agreed to the new Adelaide Declaration on National Goals for Schooling in the Twenty-First Century. The National Goals were superseded in December 2008, when the state, territory and Commonwealth ministers of education released the new Melbourne Declaration on Educational Goals for Young Australians. The new Educational Goals for Young Australians set the direction for Australian schooling for the next 10 years (Ministerial Council on Education, Employment, Training and Youth Affairs [MCEETYA] 1999 and 2008[1]) (available at www.mceecdya.edu.au).

In July 2001, MCEETYA agreed to the development of assessment instruments and key performance measures for reporting on student skills, knowledge and understandings in primary science. The National Assessment Program – Science Literacy (NAP–SL) was the first assessment program designed specifically to provide information about performance against the National Goals (now the Educational Goals). Similar sample assessment programs have since been undertaken for Civics and Citizenship, and Information and Communication Technology Literacy. Each sample assessment program is repeated every three years so that performance in these areas of study can be monitored over time.

1 subsequently the Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA) and now the Standing Council on School Education and Early Childhood (SCSEEC)

The development and implementation of the national assessment in science literacy is undertaken by the Australian Curriculum, Assessment and Reporting Authority (ACARA), the independent statutory authority responsible for the overall management and development of a national curriculum, the National Assessment Program and a national data collection and reporting program that supports 21st Century learning for all Australian students. ACARA was established under an Act of Federal Parliament on 8 December 2008 and became operational in mid 2009. ACARA receives direction from the Standing Council on School Education and Early Childhood (SCSEEC). At the direction of SCSEEC, ACARA manages the National Assessment Program.

Of the three subject areas for which sample assessments are undertaken, science is the only program that focuses entirely on primary school performance. This is because MCEECDYA agreed to use the Programme for International Student Assessment (PISA) as the measure of performance for scientific literacy among secondary students (see www.nap.edu.au).

The previous three NAP–SL assessments were conducted in 2003, 2006 and 2009. In January 2011, ACARA awarded the contract for the fourth cycle of the National Assessment Program – Science Literacy to Educational Assessment Australia (EAA). This report provides the findings of the fourth cycle of the science literacy assessment conducted in 2012.

The National Assessment Program – Science Literacy

Implementation of NAP–SL involves a large number of separate but related steps, including the development of items and tasks to assess the scientific literacy domain; the trialling of those items and tasks; the administration of the final assessment to a sample of students; and the marking, analysis and reporting of the results.

This report provides details about the school and student samples used in 2012, describes the testing process, presents the results at national, and state and territory levels, and includes comparisons with the 2006 and 2009 testing cycles. Where valid, comparisons are also made with the 2003 testing cycle.

What does NAP–SL measure?

NAP–SL measures scientific literacy.

Scientific literacy has been defined by the Organisation for Economic Co-operation and Development (OECD) – Programme for International Student Assessment (PISA) as:

> *… the capacity to use scientific knowledge, to identify questions and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity.*
>
> (OECD 1999, p. 60)

This definition has been adopted for the purpose of monitoring primary science in NAP–SL (Ball et al. 2000). The science items and instruments assess outcomes that contribute to scientific literacy, including conceptual understandings, rather than focusing solely on scientific knowledge. They also assess student competence in carrying out investigations in realistic situations.

NAP–SL relates to the ability to think scientifically in a world in which science and technology are increasingly shaping children's lives.

A Scientific Literacy Progress Map (see Appendix 1) has been developed based on this construct of scientific literacy and on an analysis of the state and territory curriculum and assessment frameworks. The Progress Map describes the development of scientific literacy across three strands of knowledge which are inclusive of Ball et al.'s concepts and processes and the elements of the OECD–PISA definition.

What aspects of scientific literacy are assessed?

As in the previous three cycles of NAP–SL, three main areas of scientific literacy were assessed in 2012:

Strand A: formulating or identifying investigable questions and hypotheses, planning investigations and collecting evidence.

Strand B: interpreting evidence and drawing conclusions from students' own or others' data, critiquing the trustworthiness of evidence and claims made by others, and communicating findings.

Strand C: using science understandings for describing and explaining natural phenomena, and for interpreting reports about phenomena.

The scientific literacy domain is detailed in Appendix 1. In addition, the items drew on four major scientific concept areas: Earth and Space; Energy and Force; Living Things; and Matter. These concept areas, found most widely in state and territory curriculum documents, were used by item developers to guide item and test development. The list of endorsed examples for each of these concept areas is in Table A1.2 of Appendix 1.

A conscious effort was made to develop assessment items that related to everyday contexts. The intention was to ensure that all Year 6 students were familiar with the materials and experiences to be used in NAP–SL and so avoid any systematic bias in the instruments being developed.

What is the national scientific literacy standard?

A standard for scientific literacy was established as part of the first cycle of the national assessment in 2003 to provide parents, educators and the community with a clear picture of the level of proficiency that students are expected to demonstrate by the end of Year 6.

To identify what students should know and be able to do by the end of Year 6, an expert group of university science educators, curriculum officers and experienced primary teachers in all states and territories, from government, Catholic and independent schools, was brought together. The members of this expert group used their classroom experience and knowledge of the science curricula in the various jurisdictions to examine the test items from the national assessment. The crucial scientific literacy skills and understandings needed by students for the next phase of science learning at school were discussed and debated before consensus was reached on a Proficient Standard for Year 6.

The Proficient Standard is a challenging level of performance, with students needing to demonstrate more than minimal or elementary skills to be regarded as reaching it. It is one of several achievement levels that collectively represent a continuum of learning and describe what students know and are able to do.

In terms of the Proficiency Levels described in Chapter 4, the standard was found to be equivalent to Level 3.2: that is, students achieving at Level 3.2 or better are considered to have a sound understanding of Year 6 science. Students at this level demonstrate considerably more skill and understanding than those performing at Level 3.1 and below.

Year 6 students who exceed the Proficient Standard (those who perform at Level 3.3 and above) demonstrate exemplary performance. Students who do not achieve the Proficient Standard demonstrate only partial mastery of the skills and understandings expected for Year 6; these students are on the way to becoming proficient.

Minimum standards like the National Minimum Standards in literacy and numeracy have not been set for scientific literacy. These minimum standards are defined as the critical level of skill and understanding without which a student will have difficulty making sufficient progress at school. They are more suited to foundational subject areas such as reading, writing and numeracy where deficiencies will have significant effects on students' future learning and functioning in society.

The Proficient Standard (equivalent to Level 3.2) is the main reference point for monitoring scientific literacy in Australian primary schools over time. Every three years, a new Year 6 national science literacy assessment is conducted to gauge whether there have been changes in student proficiency.

Information about student performance in relation to the Year 6 standard from the fourth (2012) NAP–SL assessment is reported with comparisons to 2003, 2006 and 2009 data by Proficiency Levels in Chapter 5.

Who participated in the 2012 NAP–SL assessment?

Approximately five per cent of the total Australian Year 6 student population was sampled randomly and assessed. The sample was drawn from all states and territories. Government, Catholic and independent schools participated.
Table 1.1 shows the number of schools and students in the final sample for which results were reported.

Table 1.1 Number of schools and students by state and territory in the final sample 2012

State/ Territory	Number of schools in target sample	Number and percentage of schools in final sample	Number of students in target sample	Number and percentage of students in final sample
ACT	54	54 (100%)	1305	1242 (95.2%)
NSW	92	90 (97.8%)	2246	2060 (91.7%)
NT	50	41 (82.0%)	959	710 (74.0%)
QLD	92	92 (100%)	2207	2052 (93.0%)
SA	94	94 (100%)	2082	1926 (92.5%)
TAS	64	64 (100%)	1420	1259 (88.7%)
VIC	93	90 (96.8%)	2112	1854 (87.8%)
WA	94	92 (97.9%)	2344	2133 (91.0%)
AUST	633	617 (97.5%)	14675	13236 (90.2%)

A grade-based population of students enrolled at schools was chosen. This is consistent with the other National Assessment Program sample assessments. There are differences between the states and territories in the structure and organisation of pre-primary education and the age of entry to full-time formal schooling. Information about ages of students at the time of testing is presented in Table 3.1 in Chapter 3.

Appendix 2 provides a comprehensive summary of the sample frame, with exclusions and response rates for participating schools and students by state and territory for the assessment. Further details about sampling procedures and computation of sampling weights are provided in the 2012 Technical Report (available at www.nap.edu.au).

What did NAP–SL participants have to do?

The assessment booklets comprised an objective assessment (pencil-and-paper test with multiple-choice and open-ended items), a practical assessment task and a survey. The booklets were distributed randomly so that a student in each class completed one of seven different objective assessment forms. Each class undertook one of two practical assessment tasks. The practical tasks were assigned to classes across Australia in a way that ensured that approximately equal numbers of classes attempted each of the two tasks.

The objective assessments required students to work individually to respond to approximately 40 items. The practical tasks required the students to work in groups of three. Teachers allocated students randomly to groups, using a procedure outlined in the Test Administrator's Manual. Students conducted an experiment in these groups and recorded the data they collected as a group. Students then answered a set of items independently, using their observations and the data they had collected. Only individual student responses were used in the analysis and generation of proficiency data.

Merging of the seven objective assessments onto one scale was achieved by the use of common items shared between the assessments. The practical items were then linked onto this scale using results from students doing the same objective assessment and practical task.

Students were allowed 60 minutes to complete the objective assessment and 45 minutes for the practical task.

In addition, students were asked to respond to a 34-item survey. The Student Survey sought to obtain information about students' perceptions of and attitudes to science, and their experiences of science learning within and outside school. Students were allowed 10 minutes to complete the survey. Results of the survey are summarised in Chapter 7, Student Survey.

The students' regular classroom teachers administered the NAP–SL assessment on:

- 17 October 2012 – New South Wales, Northern Territory, Queensland, South Australia, Tasmania and Victoria
- 24 October 2012 – Australian Capital Territory and Western Australia.

How are the NAP–SL results reported?

The results of NAP–SL are reported as mean scores and distributions of scores across Proficiency Levels. They are also described in terms of the understandings and skills that students demonstrated in the assessment. These understandings and skills are mapped against the scientific literacy assessment framework.

Five levels of proficiency are defined and described for scientific literacy. Further details of the Proficiency Levels, including items exemplifying these levels are contained in Chapter 4. Chapter 5 includes results in relation to the Proficiency Levels by state and territory.

Results for groups such as males and females, Indigenous students, students from different geographic locations and students from language backgrounds other than English are presented in Chapter 6.

How is this report organised?

This report provides educators and policy makers with the main findings of the 2012 NAP–SL assessment. The 2012 NAP–SL Year 6 Technical Report (available at www.nap.edu.au) provides more detailed information about the development of the assessment instruments, data collections and analyses that underpin the findings presented in this report. The 2012 NAP–SL Year 6 School Release Materials (available at www.nap.edu.au) provide sample assessment modules with accompanying marking guides.

Following this brief overview of the assessment in Chapter 1, Chapter 2 describes the development of the assessment instruments as well as the sampling and assessment administration procedures.

Chapter 3 provides a description of the scientific literacy scale. It includes results in terms of means and distributions of student performance for each state and territory as well as for the Australian population. The chapter also contains comparisons of the performance of Year 6 students over the four assessment cycles (2003, 2006, 2009 and 2012).

Chapter 4 discusses the results in terms of students' proficiency on the scientific literacy scale. The scale links the students' results to descriptions of their understandings and skills in the assessment domain. Further information about the nature and coverage of the assessment items accompanies the discussion of students' results.

Chapter 5 examines comparisons in achievement by Proficiency Levels between the tests in the four assessment cycles (2003, 2006, 2009 and 2012).

Chapter 6 provides an analysis of the results achieved by specific groups of students, including males and females, Indigenous and non-Indigenous students, and students from diverse geographic locations and language backgrounds.

Chapter 7 presents survey results about students' opinions and ideas about science and scientific literacy. The chapter also reports on the relationship between students' responses to the survey and their achievement in the assessment.

Chapter 8 provides a brief summary of the main findings of 2012 NAP–SL and the implications of those findings.

Chapter 2
The Scientific Literacy Assessment

Introduction

This chapter provides a brief description of the steps that were followed to develop the scientific literacy assessment. More detailed information about each of the steps is provided in the various publications that are referred to in this chapter. Very high standards were set for sampling, constructing assessment materials and undertaking operational procedures in order to ensure the integrity of the data.

Assessment construction and delivery

As in the previous cycles, the process of constructing and delivering the 2012 National Assessment Program – Science Literacy (NAP–SL) included the following steps:

1. clarifying the assessment domain for scientific literacy
2. constructing assessments that comprised items and tasks which defined the assessment strands operationally
3. trialling the assessments in a sample of schools
4. constructing the final assessments based on the results of the trial
5. administering the final assessments to students
6. using the measurement model and technical standards to analyse the results.

The following provides a brief description of these steps:

1. **Clarifying the assessment domain for scientific literacy**

   The conceptual framework for the assessment of scientific literacy comprises a Progress Map that describes growth in process skills and conceptual understandings, and four major scientific concept areas — broad statements of scientific understandings that students in Year 6 would be expected to demonstrate.

   The Progress Map was the key reference for test development for the 2003–2012 cycles of testing. Between 2003 and 2006 the Progress Map underwent some modification as a result of the Science Education Assessment Resources (SEAR) project. Table A1.1 in Appendix 1 includes the version of the Progress Map that has informed test development for the past three cycles.

   The four scientific concept areas that guided development for the 2003 and 2006 cycles were updated at the beginning of the 2009 cycle. This updated version of the concept areas guided test development for the 2012 cycle and is included as Table A1.2 in Appendix 1.

2. **Constructing assessments that comprised items and tasks which defined the assessment domain operationally**

   As in previous cycles, a Science Literacy Review Committee (SLRC) was established to ensure that the assessments were valid across the states and territories and to advise on critical aspects of the study. A key function of the committee was to ensure that the items comprising the assessment were fair for all students irrespective of their geographical location. In consultation with the Australian Curriculum Assessment and Reporting Authority (ACARA) and Educational Assessment Australia (EAA), the SLRC approved technical aspects of the assessment design, including the ratio of multiple-choice to open-ended items in the booklets and the percentage of items per strand.

   Test constructors developed items and tasks that enabled students at different points along the scientific literacy scale to demonstrate what they knew and could do in terms of scientific literacy. The constructors had to ensure that the tasks assessed the outcomes articulated in the assessment strands. They also had to ensure that the tasks intended to assess higher-order understandings and skills at the top of the scale were more difficult than those at the middle and bottom of the scale.

   The items were reviewed first by EAA's internal panels, then by an ACARA panel and the SLRC. Specific criteria were developed by ACARA and EAA to guide the reviews. ACARA reviewers and SLRC members were asked to judge each item against the criteria and justify their judgements. The emphasis during these reviews was on ensuring that the items and tasks reflected the understandings and skills in the assessment strands and were not biased unduly for or against particular groups of students. The feedback received was used to refine the assessment items.

   Procedures were established for recording feedback on tasks and items as the review process progressed, and associated documentation was prepared.

All practical tasks were piloted as part of the review process and the feedback from these pilot studies was taken into account in the review process.

3. **Trialling the assessments in a sample of schools**

   Once the items and tasks had been written and reviewed, they were trialled with a convenience sample of 1057 students in 30 schools selected from the government, Catholic and independent sectors in the Australian Capital Territory, New South Wales, South Australia, Queensland and Western Australia. The main purposes of the trial were to ensure that the item pool developed targeted the population appropriately and to select items that displayed excellent psychometric properties.

4. **Constructing the final assessments based on the results of the trial**

   The trial results were analysed to determine the degree to which the items and tasks measured the scientific literacy domain. The characteristics of each item were also evaluated to determine whether the scoring categories were appropriate (in the case of polytomous items) and whether the item should be included in the final item pool. The SLRC reviewed the data from the trial testing, gauged the validity of the assessment items and approved the final item pool for the 2012 assessment.

5. **Administering the final assessments to students**

   The final assessments were administered to a stratified random sample of students in October 2012. The final sample contained a total of 13 236 students at 617 schools. Information about the achieved sample is shown in Appendix 2.

6. **Using the measurement model and technical standards to analyse the results**

   Item Response Modelling was used to analyse the results from the final sample of students who participated in NAP–SL. These statistical models are used in all state and territory testing programs and in major international testing programs such as the Programme for International Student Assessment (PISA).

   Details of the application of Item Response Modelling, including the Rasch Model, can be found in the 2012 Technical Report for NAP–SL (available at www.nap.edu.au).

In Chapter 4, additional meaning and depth are added to the summary statistics by using examples of test items to reference the data to descriptions of the understandings and skills students were able to demonstrate.

The assessment booklets

In 2012, NAP–SL involved the use of seven assessment booklets. A cluster rotation design similar to that used in other sample-based international assessments was implemented. In the rotation design, assessment booklets are assembled so that each booklet is linked through common clusters to other booklets. In this way a broader range of assessment items can be completed by students and linked to other items using Modern Test Theory.

To achieve the rotation design for NAP–SL, the items were first written in contextual units. Each unit contained one or more items that were developed around a single theme or stimulus. Clusters were then constructed by grouping three to five units together. Each cluster contained approximately 13 items.

From there, booklets were compiled by arranging three clusters in every booklet following a Balanced Incomplete Block rotation design, which reduces the possibility that an item's position in an assessment booklet has an impact on its difficulty and discrimination.

Each booklet contained an objective (pencil-and-paper) test and two practical tasks. Participating students had to complete the objective section of their booklet and one of the two practical tasks. Each booklet had approximately 40 items in the objective section. The practical task required students to undertake an activity in groups of three to collect and record data from that activity. Students then responded individually to ten or eleven items, depending on the practical task they had completed.

Coverage of scientific literacy

The scientific literacy domain comprises three strands (see Appendix 1). These strands specify processes and concepts, rather than traditional subject boundaries such as physics, chemistry or biology. The strands are considered to be more relevant to students at primary schools and, according to PISA, '... to all people in their lives beyond school than the more traditional subject areas ...' (Lokan, Greenwood & Cresswell, 2001, p. 97).

Strand A involves experimental design and data gathering. More specifically, it involves skills such as formulating or identifying investigable questions and hypotheses, planning investigations and collecting evidence.

Strand B involves interpreting experimental data and requires skills such as interpreting evidence and drawing conclusions from students' own or others' data, critiquing the trustworthiness of evidence and claims made by others, and communicating findings.

Strand C involves using scientific understandings for describing and explaining natural phenomena, and for interpreting reports about phenomena.

The distribution of items across each strand is shown in Table 2.1. There were 112 items distributed across the seven objective pencil-and-paper tests and two practical tasks. In addition to those items written specifically for the 2012 cycle, 36 items from the 2003, 2006 and 2009 cycles which had been held secure were incorporated into the item pool to enable trend analyses to be undertaken.

Table 2.1 shows that 59 items assessed the process strands (17 for Strand A and 42 for Strand B) and 53 assessed the conceptual strand (Strand C).

The scientific literacy domain also comprises four major concept areas: Earth and Space; Energy and Force; Living Things; and Matter. Each major concept area includes a set of concepts — broad statements of scientific understandings that Year 6 students would be expected to demonstrate — found most widely in the various state and territory curriculum documents. Table 2.1 shows the distribution of items across each concept area.

Table 2.1 Distribution of assessment items across the assessment strands and major concept areas

Domain	Item type and number of items			
	Multiple choice	Short answer	Extended response	Total
Distribution of items by strand				
Strand A	8	-	9	17
Strand B	17	6	19	42
Strand C	27	5	21	53
Total	52	11	49	112
Distribution of items by major science concept area				
Earth and Space	13	2	5	20
Energy and Force	10	5	21	36
Living Things	17	1	11	29
Matter	12	3	12	27
Total	52	11	49	112

Types of assessment items

The strands and major concepts of scientific literacy were assessed through multiple-choice and open-ended items. The open-ended items required students to construct their own responses. These were categorised into those that required one or more words (short-answer items) and those that required more substantive responses of one to three sentences (extended-response items). The multiple-choice and short-answer items in the booklets had only single correct answers. The scores allocated to these items were zero or one. The scores allocated to extended-response items varied between zero and two.

Of the 112 items, 52 were classified as being multiple-choice, 11 as short-answer and 49 as extended-response (Table 2.1).

The sampling procedures

As in 2009, the sample design for NAP–SL was a two-stage stratified cluster sample. Stratification involves ordering and grouping schools according to state, sector, size and school location. This helps ensure that all desired school types are represented in the sample.

Stage 1 consisted of selecting schools that had Year 6 students. Within this process the list of schools was explicitly stratified and schools were then selected with probabilities proportional to the estimated Year 6 enrolments relative to their stratum.

Stage 2 involved the random selection of an intact Year 6 class from the sampled schools selected in Stage 1.

In 2012 (as in 2009), as many schools as possible were included in the defined population as possible. Essentially this meant that there were no school-level exclusions from the supplied sampling frame prior to sample selection. If a small school (with fewer than three Year 6 students) was selected, then this school could administer the objective assessment only.

The number of students sampled in each jurisdiction was determined with the following considerations in mind:

- results for each jurisdiction should be of similar precision. While this was an ultimate goal, it was recognised that reduced sample sizes would be needed for the smaller jurisdictions (Australian Capital Territory, Northern Territory and Tasmania)
- the nationwide achieved sample was to be approximately equal to 12,000 students who were to be located within approximately 600 schools throughout Australia.

Further information about the characteristics of the sample, including details of students who were granted exemptions or excluded from the sample and the procedures used to determine the standard errors of estimates, is provided in Appendix 2 of this Report and in the 2012 Technical Report (available at www.nap.edu.au).

Assessment administration procedures

Students' regular class teachers administered NAP–SL, so as to minimise disruption to the normal class environment.

Standardised administration procedures were developed and published in a Test Administrator's Manual. In all schools in which students were to complete the NAP–SL assessment, teachers and school administrators were provided with the manual. Detailed instructions were also given in relation to the participation or exclusion of students with disabilities and students from language backgrounds other than English.

Teachers were able to review the Test Administrator's Manual before the assessment date and raise questions with the coordinators of NAP–SL in their jurisdiction. EAA provided a help-desk with a toll-free telephone number and an email address to ensure all queries were dealt with promptly.

A quality-monitoring program was established to gauge the extent to which class teachers followed the specified administration procedures. This involved trained invigilators observing the administration of the 2012 assessment in a random sample of classes in 32 of the 617 schools involved. The invigilators reported conformity with the administration procedures (for further details, refer to the 2012 Technical Report).

Marking of responses to open-ended items

Over half of the items were open-ended and required marking by trained markers.

Marking Guides were prepared by EAA and refined during the trialling process. The marking team included experienced teacher-markers employed by EAA. Many markers had marked NAP–SL assessments in previous cycles.

The markers participated in a one-and-a-half day training session led by the Test Development Manager. The session involved formal presentations followed by hands-on practice with pre-marked sample student answer booklets. Presentations included leading markers through an overview of each cluster or practical task and discussing the marking criteria and illustrative answers for correct and incorrect student responses exemplified in the marking guides. In the hands-on practice, markers practised marking with a pre-marked sample of items and discussed the scores assigned to each item to help clarify distinctions between score levels. At the end of the session, all markers were asked to mark the same set of student answer booklets. The scores were compared to those agreed to by expert scorers (the Project Director, the Test Development Manager and the group leaders). Trainers discussed with markers agreements and disagreements between their scores and the scores given by expert scorers. Additional practice was provided to markers for items where consistency and accuracy were low.

Markers were monitored constantly for reliability by having samples of their student answer booklets check-marked by group leaders. In cases where there were differences between markers and group leaders, the scoring was reconciled jointly in consultation with the Test Development Manager. In addition, once a day all markers were asked to mark the same set of student answer booklets. The scores were compared to the scores agreed to by expert scorers and differences were discussed and reconciled.

In addition, approximately five per cent of the 2009 NAP–SL trend item responses were also marked by the 2012 markers to ensure the reliability of marking. These procedures, coupled with the intensive training at the beginning of the marking exercise, ensured that markers applied the scoring criteria consistently and accurately.

Data entry procedures

The multiple-choice responses and teacher-marked scores were data processed. A validation of the data processing was performed that ensured accuracy in data capture.

Scanning software was used to capture images of all the student responses. These have been indexed and provided to ACARA for future reference.

School reports

Schools that participated in NAP–SL were provided with feedback about the performance of their students on the assessment prior to the close of the 2012 school year. The reports showed the results for each student on an item-by-item basis with comparative data showing the percentage of the school and the national sample of students responding correctly to the item. In the case of items that had more than one mark available for the response to the item, the percentage of students achieving the maximum score on the item was provided.

NAP–SL School Release Materials

Some assessment items have been released from the 2012 NAP–SL assessment to enable teachers to administer assessments under similar conditions and gauge their own students' proficiency in relation to the national standards. The School Release Materials comprise an objective assessment containing 39 items, made up of multiple-choice, short-answer and extended-response questions, as well as a practical assessment task. The School Release Materials will be made available at www.nap.edu.au.

The remaining 2012 assessment items have been secured for the purpose of equating the next NAP–SL assessment (which is to be undertaken in 2015). This, together with the previous assessments, will allow longitudinal data on student performance to be obtained.

Chapter 3
Student Performance in Scientific Literacy for 2012

Introduction

In this chapter, summary statistics for the 2012 National Assessment Program – Science Literacy (NAP–SL) are shown in terms of student mean scores and distributions of scores by state and territory. In addition, an overview of the methodology used to construct the scientific literacy scale for reporting the results of NAP–SL is provided. This chapter also contains comparisons of performance of Year 6 students over the 2006, 2009 and 2012 assessment cycles.

Scientific literacy scale

A scientific literacy scale was constructed in 2003, using the Rasch measurement model. The Rasch analysis produced information about the relative difficulty of the assessment items, as well as information about students' abilities. These data were located on a continuum to form the scientific literacy scale, and a national mean was set at 400 with a standard deviation of 100.

While the first cycle of NAP–SL was conducted in 2003, after the second cycle in 2006 it was decided to use the results of the 2006 assessment to reconstruct the scientific literacy scale. The reasons for reconstructing the scale included:

- The 2006 test design was more robust than the 2003 test design.
- There were considerably more items in 2006 than in 2003, resulting in a better coverage of the assessment domain in 2006.
- The 2006 items were generally more discriminating than the 2003 items.
- The 2006 sampling was more comprehensive, as remote schools were also included in the sample (see the 2006 Technical Report for more information).

The Rasch measurement model that included item position and a set of relevant student characteristics (e.g. gender, jurisdiction and school location) as parameters were used to estimate item difficulties and student abilities. The plausible values methodology was utilised to obtain a precise estimate of student abilities (for detailed information see the 2012 Technical Report). These results were then mathematically transformed to construct the scientific literacy scale that has a mean of 400 and a standard deviation of 100. In the remainder of this report, all references to the scientific literacy scale are to the 2006 reconstructed scale.

Establishing Proficiency Levels

One of the main objectives of NAP–SL is to monitor trends in scientific literacy performance over time. One convenient and informative way of doing so is to reference students' results to the Proficiency Levels. Typically, students whose results are located within a particular Proficiency Level are able to demonstrate the understandings and skills associated with that level and possess the understandings and skills of lower Proficiency Levels. NAP–SL covers a range of five Proficiency Levels: Level 2, Level 3.1, Level 3.2, Level 3.3 and Level 4. In 2006, Proficiency Levels were assigned corresponding to cut-points on the scientific literacy scale. The Proficient Standard in scientific literacy was set at the boundary between Level 3.1 and Level 3.2. This means that students who obtain a score equal to or above the Level 3.2 cut-point of 393 are deemed to have attained the Proficient Standard in scientific literacy. This cut-point is used for each assessment cycle.

An overview of 2012 results relative to the distribution of student scores in Proficiency Levels, as well as information about the proportion of students who attained the Proficient Standard, are presented in Chapter 5 and Chapter 6 of this report.

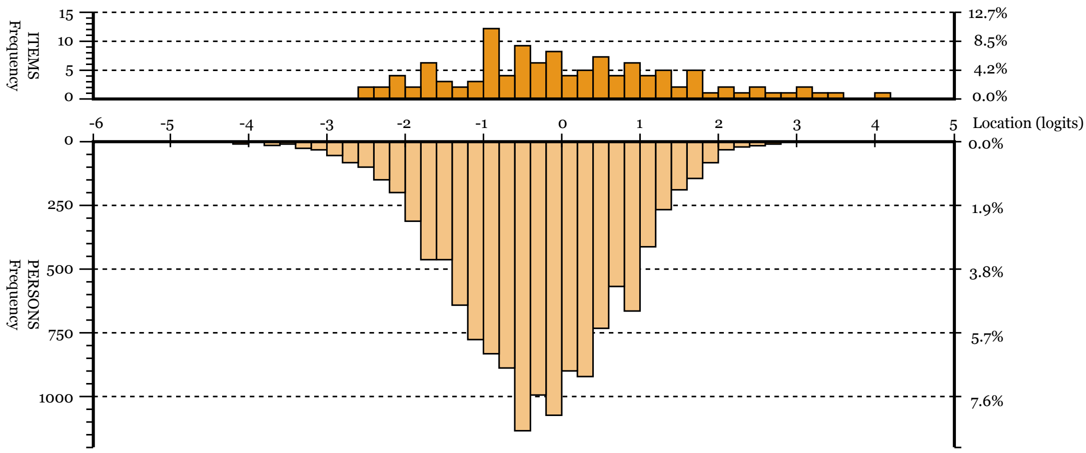
Analysing the 2012 results

The measurement model for analysing student responses in 2012 was the same as that of the previous cycles. The common item equating methodology was used to place the 2012 results on the scientific literacy scale.

In 2012, care was taken to increase the number of link items that could be used to equate the 2012 results to the scientific literacy scale. A total of 36 link items were included in the 2012 assessment. This included six link items from the 2003 assessment, 14 link items from the 2006 assessment and 16 link items from the 2009 assessment. These link items covered a range of scientific literacy strands, concept areas and item difficulties. This set of 36 potential link items was evaluated and used to equate the 2012 item and student parameters to the scientific literacy scale (for detailed information about the link items see the 2012 Technical Report).

Before presenting data for the application of the Rasch measurement model, it is important to ensure that the test has appropriately targeted the student population. As can be seen from Figure 3.1, the 2012 assessment achieved a good spread of item difficulties and was appropriately matched to the Year 6 cohort. This demonstrates that the items targeted the population well and were able to discriminate between achievements at the highest level while still catering for less able students.

Figure 3.1 2012 NAP–SL: Item-Person map



Achievement by state and territory in 2012

Age of students

Table 3.1 displays the average age of students at the time of testing.

Table 3.1 Distribution of ages of students in the sample by state and territory

State/Territory	Average age at time of testing
ACT	12 yrs 0 mths
NSW	12 yrs 0 mths
NT	11 yrs 10 mths
QLD	11 yrs 5 mths
SA	12 yrs 0 mths
TAS	12 yrs 3 mths
VIC	12 yrs 3 mths
WA	11 yrs 9 mths

It can be seen that the average age of students varies considerably between states and territories with Queensland having the youngest students on average.

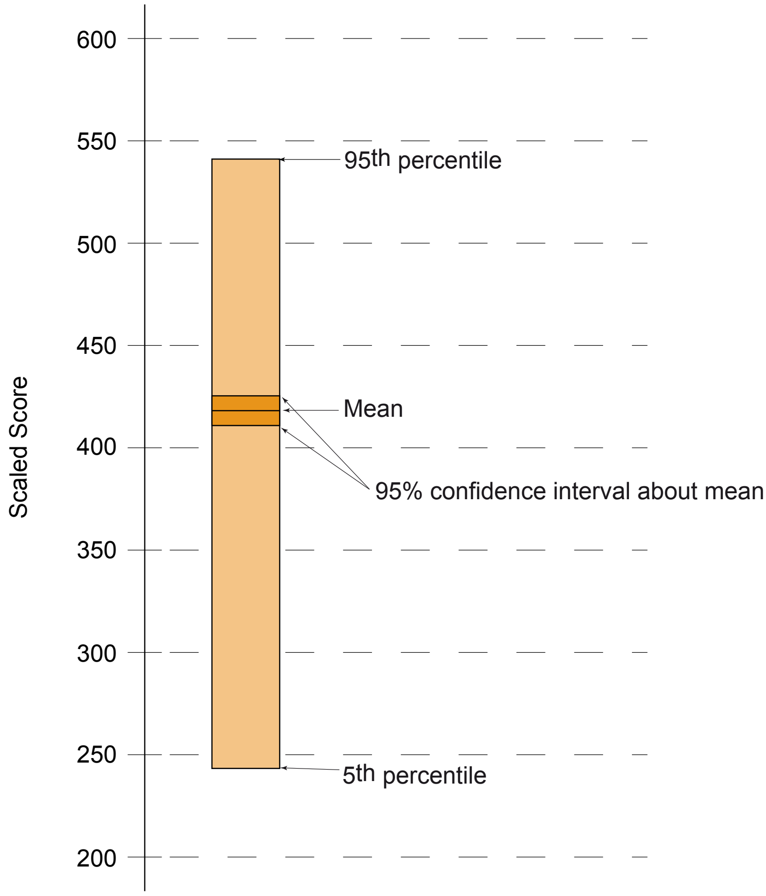
Reading the bar charts

Figure 3.2 is an example of a bar chart used to display the scaled mean scores and distributions for states and territories. The vertical bar shows the range of student performance.

The highest point on the bar is the 95th percentile, which is the point above which the highest-scoring five per cent of the students are located.

The lowest point on the vertical bar is the 5th percentile, which is the point below which the lowest-scoring five per cent of students are located.

Figure 3.2 Sample bar chart



Located in the middle region of the bar is a darker gold band that contains a thin horizontal black line. This black line denotes the mean score, while the darker regions on either side represent a confidence interval which gives an indication, through the width of the band, of the level of accuracy with which the mean was measured.

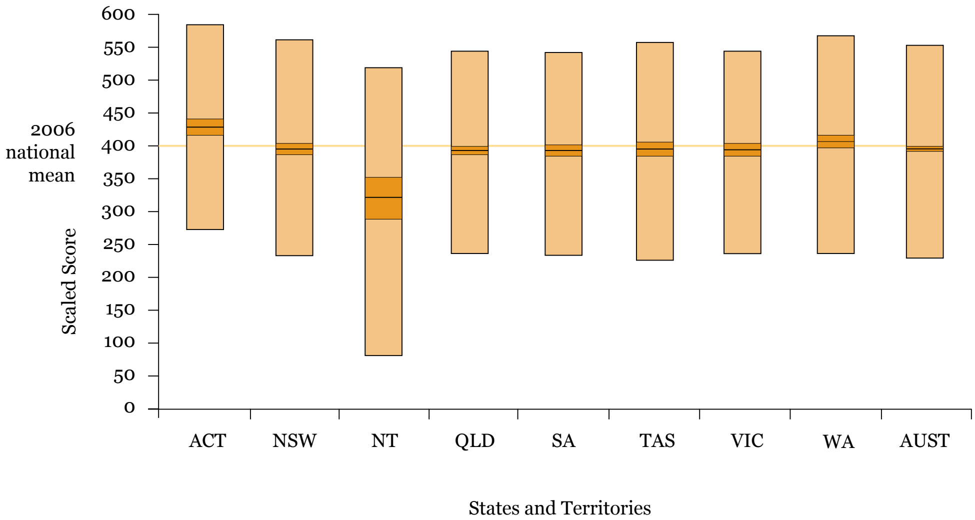
Given that NAP–SL is a sample-based assessment, the reported means are estimates of a true population mean (the mean that would be measured if the complete population of Year 6 students in Australia could be assessed). A confidence interval provides the range that contains the value of the true population mean.

Confidence intervals in this report were constructed with an estimate of statistical precision. This estimate of precision is such that the probability that a confidence interval does not contain the true population mean equals the probability of obtaining five incorrect measurements out of 100 random draws. Such a confidence interval is referred to as the 95 per cent confidence interval.

Mean and range of student scores in 2012

Figure 3.3 shows student performance in scientific literacy for each state and territory in 2012 against the 2006 mean. The bars show the spread of scores for each state and territory that were achieved by the middle 90 per cent of the population. Shaded bands within each bar mark the confidence interval around the corresponding mean. Any interpretation of results needs to be made by considering the relative precision (as indicated by the size of the confidence intervals) of the reported estimates of the student mean achievement.

Figure 3.3 2012 distribution of Year 6 student performance by state and territory



	ACT	NSW	NT	QLD	SA	TAS	VIC	WA	AUST
Mean	429	395	319	392	392	395	393	406	394
95% CI	±13.2	±9.9	±31.1	±6.4	±7.9	±12.3	±9.7	±9.5	±4.4

It can be seen that the Northern Territory had the widest spread of scores achieved by the middle 90 per cent of students and the largest confidence interval around the mean score. All other states and territories have relatively similar widths of score range and confidence interval. This suggests a more heterogeneous sample in the Northern Territory compared with other jurisdictions[1].

Additional information about the range of student scores in 2012 is provided by listing the scaled scores corresponding to the standard range of percentile values by each state and territory in Table 3.2. The table shows that the Northern Territory has lower percentile scores than all other jurisdictions, whereas the Australian Capital Territory has the highest percentile scores.

1 This is likely due to the following features of the Northern Territory:
- it has no metropolitan schools
- 25 per cent of the Year 6 student population is Indigenous
- 33 per cent of the Year 6 student population lives in remote or very remote areas.

Table 3.2 Distribution of percentile scores by state and territory in 2012

State/ Territory	Mean score	95 per cent confidence interval	Percentile						
			5th	10th	25th	50th	75th	90th	95th
ACT	429	±13.2	269	305	365	432	495	547	580
NSW	395	±9.9	231	267	325	394	463	523	560
NT	319	±31.1	78	128	233	331	413	480	517
QLD	392	±6.4	236	270	327	392	458	514	544
SA	392	±7.9	234	269	330	394	458	510	542
TAS	395	±12.3	226	265	326	396	466	528	559
VIC	393	±9.7	237	273	332	395	457	511	544
WA	406	±9.5	236	274	339	409	478	534	567
AUST	394	±4.4	232	269	329	396	461	519	553

Comparisons of means by state and territory in 2012

Tables 3.3a and 3.3b contain results of a series of pair-wise comparisons between means for states and territories to determine if the jurisdictional differences are statistically significant. The Bonferroni adjustment to statistical significance testing is conducted in order to account for the possibility that a difference can be deemed to be statistically significant by chance when multiple comparisons are conducted using the same data. The Bonferroni adjustment increases the strictness of the criterion for establishing statistical significance relative to a pair-wise comparison, hence making it harder to claim that a difference is statistically significant. By reading across the lines in Tables 3.3a and 3.3b it is possible to draw a comparison between any two jurisdictions. The results in Table 3.3a do not include the Bonferroni adjustment, while the results in Table 3.3b incorporate the Bonferroni adjustment. Comparisons that are statistically significant are shown by an upward or downward symbol.

Table 3.3a Multiple comparisons of scientific literacy results by state and territory for 2012 without the Bonferroni adjustment

	Mean score	95% CI	ACT	NSW	NT	QLD	SA	TAS	VIC	WA
ACT	429	±13.2		▲	▲	▲	▲	▲	▲	▲
NSW	395	±9.9	▼		▲	●	●	●	●	●
NT	319	±31.1	▼	▼		▼	▼	▼	▼	▼
QLD	392	±6.4	▼	●	▲		●	●	●	▼
SA	392	±7.9	▼	●	▲	●		●	●	▼
TAS	395	±12.3	▼	●	▲	●	●		●	●
VIC	393	±9.7	▼	●	▲	●	●	●		●
WA	406	±9.5	▼	●	▲	▲	▲	●	●	

▲	Mean performance that is statistically significantly higher than in comparison state/territory
●	No statistically significant difference from comparison state/territory
▼	Mean performance that is statistically significantly lower than in comparison state/territory

Table 3.3b Multiple comparisons of scientific literacy results by state and territory for 2012 with the Bonferroni adjustment

	Mean score	95% CI	ACT	NSW	NT	QLD	SA	TAS	VIC	WA
ACT	429	±13.2		▲	▲	▲	▲	▲	▲	●
NSW	395	±9.9	▼		▲	●	●	●	●	●
NT	319	±31.1	▼	▼		▼	▼	▼	▼	▼
QLD	392	±6.4	▼	●	▲		●	●	●	●
SA	392	±7.9	▼	●	▲	●		●	●	●
TAS	395	±12.3	▼	●	▲	●	●		●	●
VIC	393	±9.7	▼	●	▲	●	●	●		●
WA	406	±9.5	●	●	▲	●	●	●	●	

▲	Mean performance that is statistically significantly higher than in comparison state/territory
●	No statistically significant difference from comparison state/territory
▼	Mean performance that is statistically significantly lower than in comparison state/territory

It can be seen in Table 3.3a that when the test of statistical significance did not include the Bonferroni adjustment, the mean score for the Australian Capital Territory was significantly higher than that for all other states and territories. The differences in mean achievement between students from Western Australia and students from the Northern Territory, Queensland and South Australia were also statistically significant when the test was conducted without the Bonferroni adjustment.

However, when the Bonferroni adjustment was implemented (Table 3.3b), the mean score for students from the Australian Capital Territory was no longer significantly different from the mean score for students from Western Australia. Similarly, the mean score for students from Western Australia was no longer significantly different to the mean scores for Queensland and South Australia.

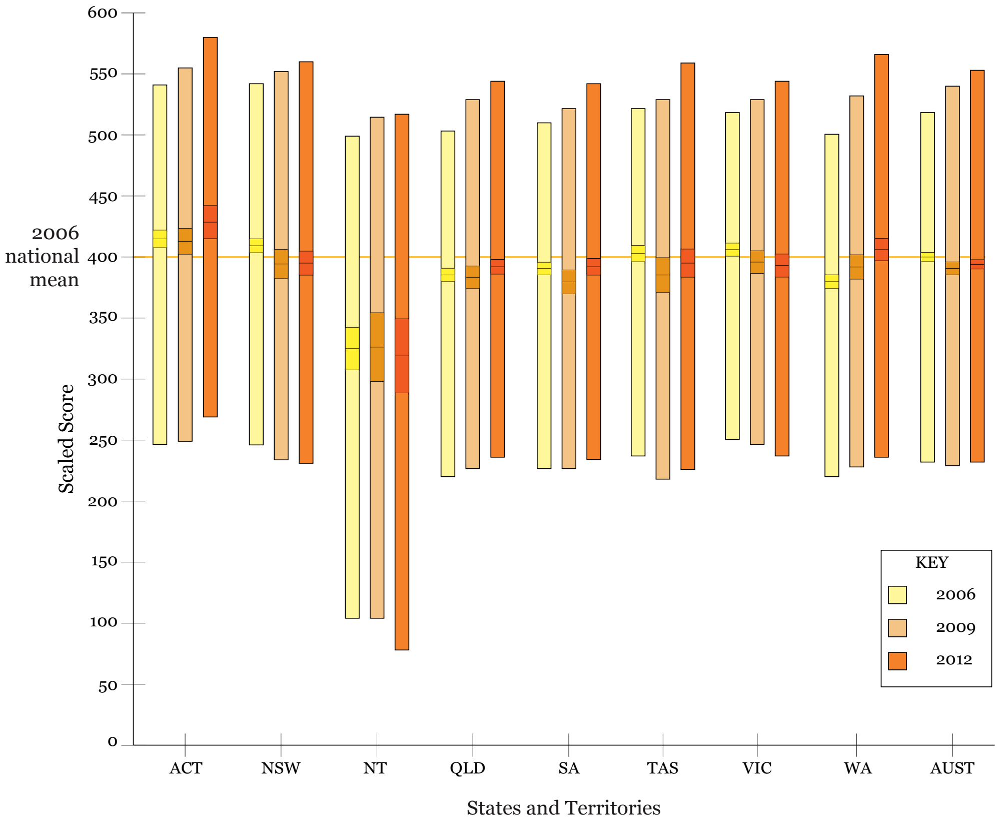
Students from the Northern Territory achieved a substantially lower mean score than students in all other states and territories with all comparisons producing a statistically significant difference, with and without the Bonferroni adjustment.

Comparisons of student results in 2006, 2009 and 2012

The 2012 NAP–SL was the fourth time the science domain had been assessed in the National Assessment Programs, with the first assessment being carried out in 2003 (Primary Science Assessment Program). Given that the 2003 assessment differed from the 2006 assessment in terms of the item booklet design, the sampling plan and the number of items, the decision was made to use the 2006 results to construct the scientific literacy scale. For this reason, tests of statistical difference were conducted only between the 2006 and 2012, and between the 2009 and 2012 assessment cycles. To test whether the 2012 results differ from those of previous assessment cycles, link errors were added to the standard error estimate in a simple pair-wise test of statistical significance (for further details regarding the link error, refer to the 2012 Technical Report).

Figure 3.4 contains graphical comparisons of the student results in the middle 90 per cent of the distribution. Shaded bands around the mean within each bar mark the 95 per cent confidence interval.

Figure 3.4 Comparison of distributions of student scores by state and territory in 2006, 2009 and 2012



As depicted in Figure 3.4, overall the 2012 means at the national level and at the state and territory level were about the same as those of 2006. The exception was Western Australia which showed an increase of 25 points in 2012 compared to the mean student achievement in 2006.

Table 3.4 shows the 2006, 2009 and 2012 mean scores by state and territory and indicates whether the differences in means between 2009 and 2012, and between 2006 and 2012 are statistically significant.

Table 3.4 Comparison of 2006, 2009 and 2012 jurisdiction mean scores

State/ Territory	Mean score			Change from 2009 to 2012	Statistically significant	Change from 2006 to 2012	Statistically significant
	2006	2009	2012				
ACT	418 (±14.3)	415 (±10.6)	429 (±13.2)	14	NO	11	NO
NSW	411 (±12.5)	396 (±12.1)	395 (±9.9)	-1	NO	-16	NO
NT	325 (±33.7)	326 (±28.6)	319 (±31.1)	-7	NO	-6	NO
QLD	387 (±8.6)	385 (±8.9)	392 (±6.4)	7	NO	5	NO
SA	392 (±10.0)	380 (±10.4)	392 (±7.9)	12	NO	0	NO
TAS	406 (±12.1)	386 (±13.5)	395 (±12.3)	9	NO	-11	NO
VIC	408 (±10.2)	398 (±9.2)	393 (±9.7)	-5	NO	-15	NO
WA	381 (±10.0)	393 (±9.6)	406 (±9.5)	13	NO	25	YES
AUST	400 (±5.4)	392 (±5.1)	394 (±4.4)	2	NO	-6	NO

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
Mean scores have been rounded.

As can be seen in Table 3.4 there is no statistically significant difference between 2009 and 2012 results at the national level or across Australian states or territories. Similarly, the comparison between 2006 and 2012 results shows that there is no statistically significant difference at the national level. The only exception at the jurisdiction level is Western Australia for which the increase in student mean achievement from 2006 to 2012 is statistically significant.

Ranking of jurisdictions by mean scores

Table 3.5 shows a jurisdiction-by-jurisdiction comparison of the mean scores in rank order for 2006, 2009 and 2012.

Table 3.5 State and territory mean score rankings in 2006, 2009 and 2012

Rank by jurisdiction mean score	2006		2009		2012	
	State/ Territory	Mean score	State/ Territory	Mean score	State/ Territory	Mean score
1	ACT	418 (±14.3)	ACT	415 (±10.6)	ACT	429 (±13.2)
2	NSW	411 (±12.5)	VIC	398 (±9.2)	WA	406 (±9.5)
3	VIC	408 (±10.2)	NSW	396 (±12.1)	TAS	395 (±12.3)
4	TAS	406 (±12.1)	WA	393 (±9.6)	NSW	395 (±9.9)
5	SA	392 (±10.0)	TAS	386 (±13.5)	VIC	393 (±9.7)
6	QLD	387 (±8.6)	QLD	385 (±8.9)	SA	392 (±7.9)
7	WA	381 (±10.0)	SA	380 (±10.4)	QLD	392 (±6.4)
8	NT	325 (±33.7)	NT	326 (±28.6)	NT	319 (±31.1)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
Mean scores have been rounded.
Ranking is based on mean scores before rounding.

It can be seen from Table 3.5 that the mean score for students from Western Australia changed from seventh rank in 2006 to fourth rank in 2009 and to second rank in 2012. This means that in Western Australia students' performance in science literacy has improved consistently from 2006 to 2012. New South Wales and Victoria were among the top three states in 2006 and 2009. In 2012 however, they moved down two places from 2006 levels to the fourth and fifth rank respectively. Compared with the 2006 rankings, South Australia and Queensland moved down one place to the sixth and seventh rank, while Tasmania moved up one place to the third position. The Australian Capital Territory and the Northern Territory did not change positions from the 2006 to the 2012 assessment. Nevertheless, given that the differences in mean achievement in 2012 were not statistically significant between Western Australia, Tasmania, New South Wales, Victoria, South Australia and Queensland, the change in ranking order for these states should be regarded as indicative only.

Trends in mean achievement in scientific literacy

An overview of the trends in scientific literacy at the national level for 2003, 2006, 2009 and 2012 is provided below in Table 3.6. However, as tests of statistical significance between 2003 and 2012 results were deemed not to be sound owing to the reasons detailed at the beginning of this chapter, the 2003 results are indicative only.

Table 3.6 Trends in mean scores in scientific literacy in 2003, 2006, 2009 and 2012

AUST	Mean score
2003	409 (±3.7)
2006	400 (±5.4)
2009	392 (±5.1)
2012	394 (±4.4)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
2003 results are indicative only.

Summary

In summary, the 2012 results are similar to those of 2006 and 2009 both in terms of student mean achievement and the distribution of student scores. The analysis shows that there are no statistically significant differences between the 2009 and 2012 results at jurisdictional levels and at the national level. The analysis also shows that there are no statistically significant differences between the 2006 and 2012 results for most jurisdictions and at the national level. The only exception is Western Australia, where the mean student achievement obtained in 2012 increased from 2006 levels and this difference was found to be statistically significant.

Chapter 4
Interpreting the Scientific Literacy Results

Introduction

Chapter 3 showed students' score distributions on the scientific literacy scale. The results can also be referenced directly to the assessment domain, by the items comprising the tests, to reveal the understandings and skills demonstrated by students.

For the purposes of this report the scientific literacy scale has been partitioned into levels called 'Proficiency Levels'.

This chapter discusses the establishment of the Proficiency Levels and the cut-off scores for each of the levels, and provides examples of items which illustrate the skills and knowledge required at each level.

Establishing Proficiency Levels

One of the main objectives of the National Assessment Program – Science Literacy (NAP–SL) is to monitor trends in scientific literacy performance over time. One convenient and informative way of doing so is to reference the results to the Proficiency Levels.

Typically, students whose results are located within a particular Proficiency Level are able to demonstrate the understandings and skills associated with that level as well as the understandings and skills of lower Proficiency Levels.

Initially, three Proficiency Levels, corresponding with Levels 2, 3 and 4 of the assessment domain, were identified. However, as 90 per cent of students' scores fell within Level 3 in the 2003 assessment, three further Proficiency Levels within Level 3 were created, providing five levels for reporting student performance in the assessment.

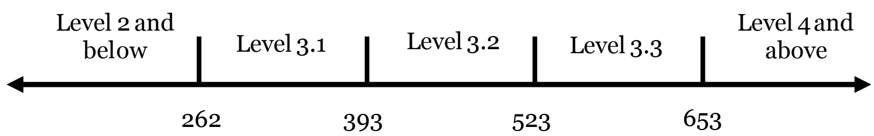
The cut-off points, which denote the boundaries between the Proficiency Levels, were established using a combination of experts' knowledge of the skills required to answer each scientific literacy item and information from the analysis of student responses.

The difficulty range spanned by each Proficiency Level is such that students whose scores are at the top of a level have a 65 per cent chance of answering the hardest items in that level correctly and an 87 per cent chance of answering the easiest items correctly. On average these students would be expected to answer about 76 per cent of the items in that level correctly.

Students who are at the bottom of a level have a 65 per cent chance of answering the easiest items in the level correctly and a 35 per cent chance of success on the hardest items. On average these students would be expected to answer about 50 per cent of the items in that level correctly.

The cut-off scores for each level are shown in Figure 4.1. The same cut-off scores have been used to determine the Proficiency Levels for the past four assessment cycles.

Figure 4.1 Cut-off scores



A score greater than 653 locates students in Proficiency Level 4 and above.

Scores in the range of 262 to 653 relate to Proficiency Level 3 on the assessment framework.

Figures 4.2 and 4.3 show comparisons, at the national level, of the percentage of students in each of the Proficiency Levels in 2006 and 2012, and in 2009 and 2012 respectively.

Figure 4.2 Distribution of students in Proficiency Levels for 2006 and 2012

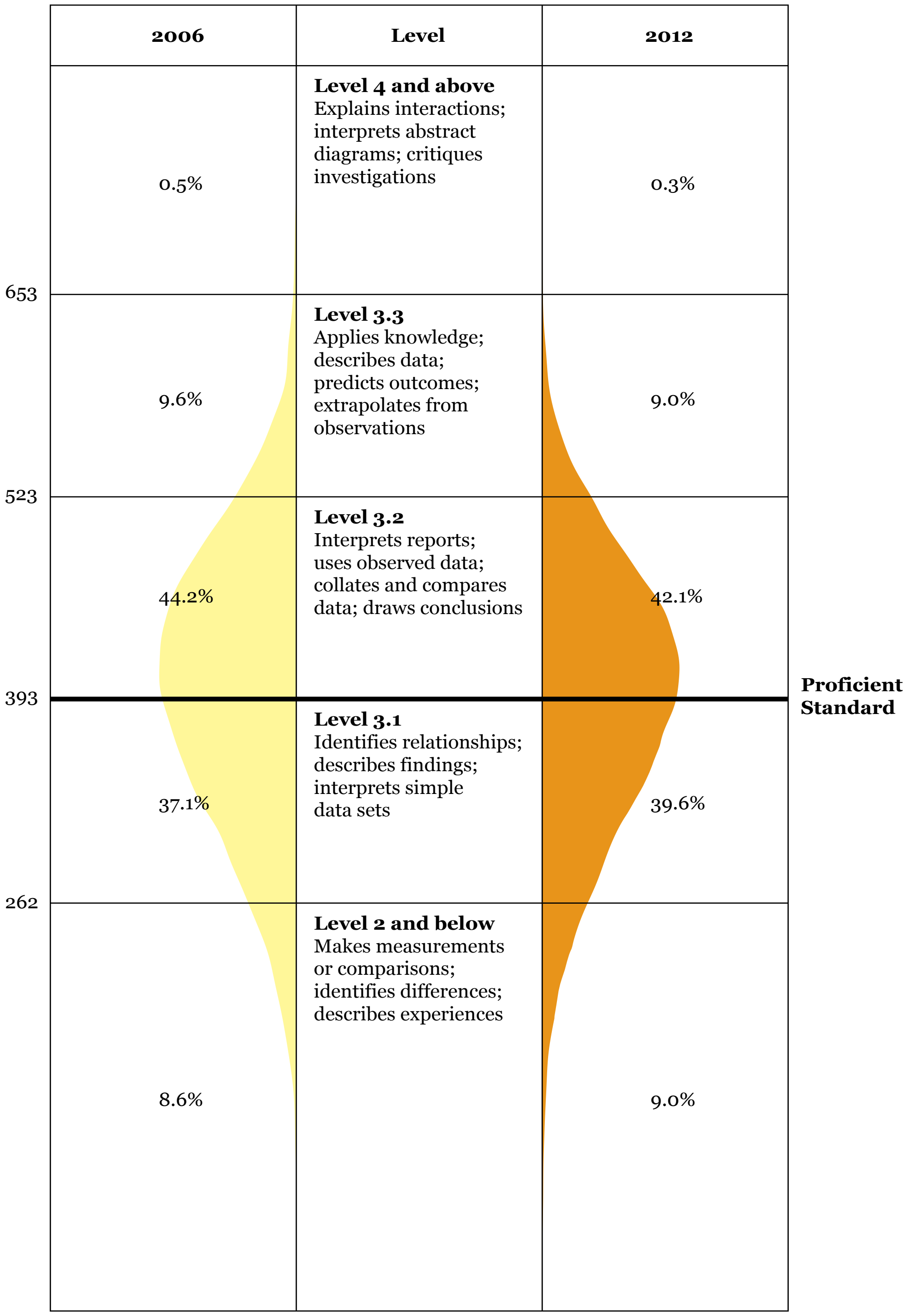
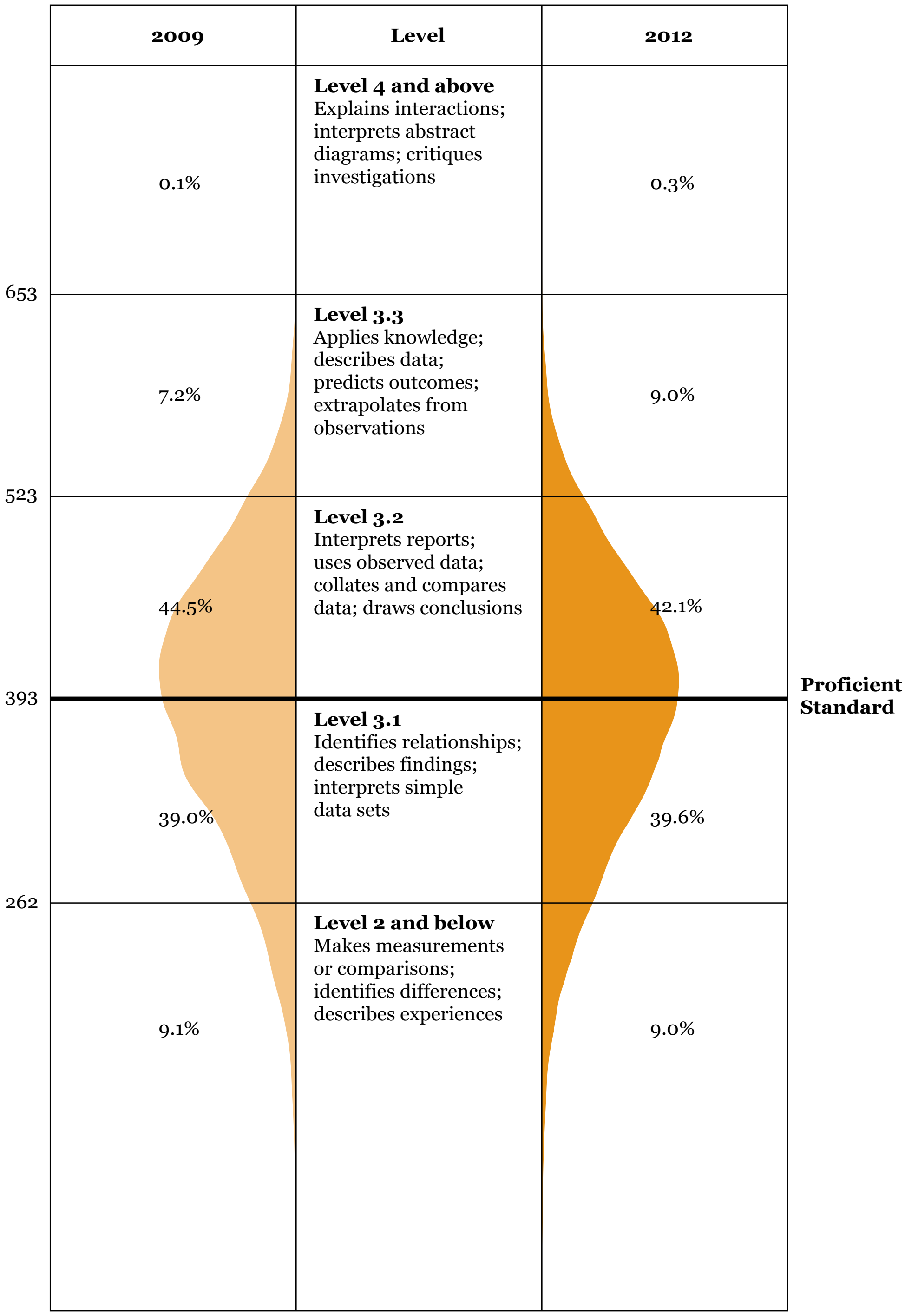


Figure 4.3 Distribution of students in Proficiency Levels for 2009 and 2012



Overall the distribution of student scores in 2006, 2009 and 2012 is very similar.

Between 2006 and 2012, at the national level, the difference between the proportion of students performing at or above the Proficient Standard is 2.9 per cent, which is not a statistically significant difference. The corresponding difference between 2009 and 2012 is 0.5 per cent, which is also not statistically significant. There is also no significant difference in the distribution of the percentages of students across corresponding Proficiency Levels in 2006, 2009 and 2012.

Describing Proficiency Levels

Appendix 3 provides the descriptions of the understandings and skills required of students at each Proficiency Level. The descriptions come from the scientific literacy assessment domain presented in Appendix 1, but Level 3 has been divided into sub-levels 3.1, 3.2 and 3.3. Table A3.1 in Appendix 3 also includes descriptors for sample items from the 2012 assessment at each Proficiency Level.

Sample items from the 2012 assessment which illustrate the skill expectations of each Proficiency Level follow.

Sample items illustrating Proficiency Levels

The following sections provide sample items that illustrate the types of understandings and skills that students at a particular Proficiency Level are likely to display.

At each Proficiency Level, a wide range of items that varied in context, format and difficulty were used to give students the best opportunity to provide evidence of what they knew and could do in relation to scientific literacy.

Only a small number of items have been released in this report. These items, as well as a further set, have been included in the School Release Materials (available at www.nap.edu.au); others have been retained as secure trend items for scaling purposes in future national science assessment cycles.

Sample items illustrating performance at Proficiency Level 4 and above

Question 4 in the item set 'Evaporating liquids' (Figure 4.4) illustrates performance at Level 4 and above. This extended-response item assesses Strand A, and relates to the major scientific concept area Matter. It assesses students' ability to explain why it is important to conduct fair tests in science.

Students were provided with a scenario in which a student places the same amount of water and lemonade in two different containers to investigate if water or lemonade evaporates faster. The question required students to explain that it is important to conduct fair tests to be able to compare the results meaningfully and/or draw valid conclusions.

Students who can complete items of this level of scientific literacy can also be expected to identify the variable to be changed, the variable to be measured and, in addition, identify at least one variable to be controlled within the context of formulating their own scientific questions for investigation.

This item is located at 848 on the scientific literacy scale.

Figure 4.4 Item illustrating performance at Proficiency Level 4 and above

Q4 In science, it is important to conduct fair tests. Why is this necessary?

____

____

____

Figure 4.5 contains a second example illustrative of Level 4 and above. This item assesses Strand B for the concept area Living Things.

In the Practical Task 'Reaction time', students were asked to perform an experiment to investigate how long it takes to react to and then catch a falling ruler. Students were randomly assigned to groups of three to carry out the experiment. Alternating roles, one student was asked to practise dropping a ruler while another tried to catch it with their dominant hand ("the hand they write with") and a third measured the distance the ruler fell. For the experiment, students were asked to collect data on the distance the ruler fell before it was caught. They had to use their dominant hand first, and then their non-dominant hand.

Question 6 in 'Reaction time' required students to identify a suggestion that would result in an improvement to the experimental method and justify their choice. Students who can complete items at this level of scientific literacy can be expected to make general suggestions for improving an investigation.

This item is located at 811 on the scientific literacy scale.

Figure 4.5 Item illustrating performance at Proficiency Level 4 and above

Q6 Three students were discussing this experiment. They had some more suggestions for improving the experiment.

Matt: 'People who always write with their right hand should have been put into the same group.'

Kim: 'We should have practised with our non-dominant hand too.'

Dana: 'We should have used our dominant hand for all of the tests.'

Which student made a suggestion that would improve the experiment?

◯ Matt ◯ Kim ◯ Dana

Give a reason for your choice.

____

____

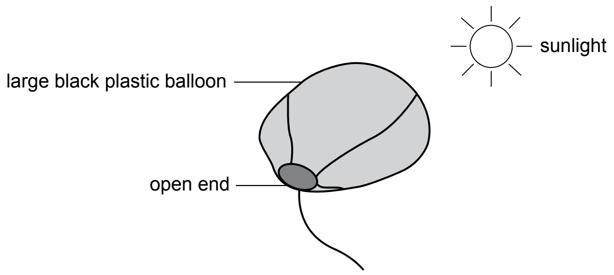
____

Question 2 in 'Solar energy' assesses Strand C and the concept area Energy and Force (Figure 4.6). Students were initially presented with information about an experiment designed to investigate the effect of sunlight on objects. The stimulus presented students with second-hand data in the form of a table showing the change in temperature of boxes of different material and outside colour. In Question 2, students had to use the information presented in the stimulus and their own knowledge to explain in detail why a solar balloon rises.

This is an example of a polytomous item, that is the item had a possible score of 0, 1 or 2. At score 2, students were required to provide a response which demonstrated the ability to explain the rising of the solar balloon in terms of an abstract science concept, namely that the balloon absorbs heat and that the air inside the balloon expands making the air inside the balloon less dense. A response which achieved a score of 1 was one that provided an explanation based on observable properties or a cause and effect relationship (e.g. 'the balloon absorbs heat and hot air rises').

Figure 4.6 Stimulus and item illustrating performance at Proficiency Level 4 and above

Q2 Fiona constructed a large balloon using black plastic. She used a fan to fill the balloon with air and left it in the sunlight.



After some time, the balloon rose in the air without Fiona having to do anything more to it.

Explain in detail why the balloon rose.

____

____

____

____

Figure 4.7 shows another item which is also illustrative of Level 4 and above. The item assesses Strand C and the concept area Earth and Space. Students were presented with information about an experiment designed to measure the loss of mass of rock pieces placed in a container with water after the container was shaken for a certain amount of time. Question 3 required students to relate the experiment to real world phenomena. Students needed to demonstrate an understanding that the experiment shows that rocks can be broken down by collisions/grinding against each other or to provide examples of how this occurs in the environment (e.g. rivers rushing down mountains).

This item is located at 734 on the scientific literacy scale.

Figure 4.7 Stimulus and item illustrating performance at Proficiency Level 4 and above

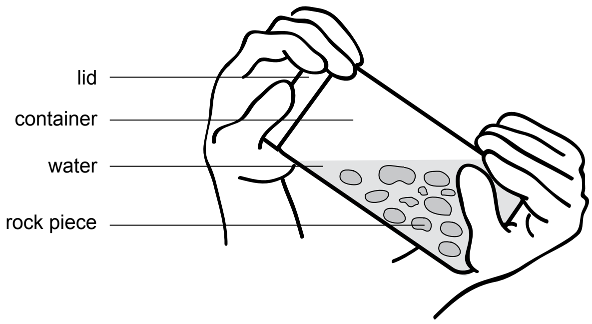
Changing rocks

Rocks can be broken down by water, ice, wind and the actions of plants and animals.

Class 6B wanted to find out whether some rocks break down more easily than others.

They conducted the following experiment:

1. They took 100 g of sandstone pieces and 100 g of quartzite pieces. The sandstone and quartzite pieces were approximately the same size.
2. They placed the sandstone and quartzite pieces in separate containers, each half full of water.
3. They covered each container and shook each one for 5 minutes.



4. They removed the small pieces that were chipped off by draining each container.
5. They dried and weighed the remaining sandstone and quartzite pieces.
6. They repeated steps 2-5 three times.

The table shows the students' results.

Table: Mass of rock pieces at the end of each trial

Trial	Mass of sandstone pieces (g)	Mass of quartzite pieces (g)
1	98.5	99.7
2	96.9	99.4
3	95.3	99.3
4	92.7	99.0

Q3 In what way is the experiment conducted by Class 6B similar to how rocks break down in the environment?

____

____

____

Sample items illustrating performance at Proficiency Level 3.3

Question 1 in ‘Making jelly’ (Figure 4.8) assesses Strand A and is illustrative of Level 3.3. This item set draws on the concept area Matter.

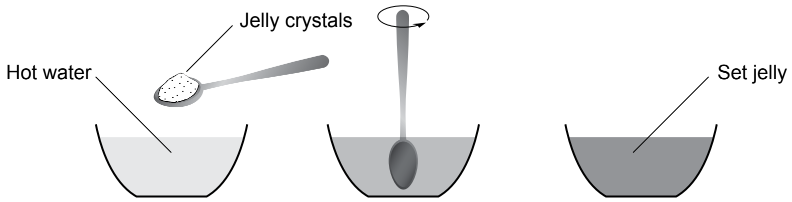
Students were presented with information about an experiment on dissolving jelly crystals in hot water and the data collected in the experiment. Question 1 required students to formulate the question investigated in the experiment. Students needed to provide a scientific question which identified the names of the variables changed and measured (number of tablespoons of jelly crystals and time taken for jelly to set).

This question is located at 542 on the scientific literacy scale. Students who can complete items requiring this level of scientific literacy would be able to demonstrate an awareness of the principles of conducting a fair experiment recognising variables to be changed and measured.

Figure 4.8 Stimulus and item illustrating performance at Proficiency Level 3.3

Making jelly

Jelly is made by dissolving jelly crystals in hot water. After the crystals have dissolved, the mixture is cooled to allow the jelly to set. The jelly is ready to eat after it has set.



Bob made jelly using five bowls.

He did the following:

- He filled each bowl with two cups of hot water.
- He added jelly crystals to each bowl.
- He stirred the mixtures.
- He covered the bowls and placed them in the fridge.

Bob measured how long it took for each bowl of jelly to set. He recorded his results in the table below.

Bowl	Number of tablespoons of jelly crystals	Time taken to set (minutes)
1	2	210
2	4	185
3	6	Not recorded
4	8	115
5	10	90

Q1 What question was Bob trying to answer?

____

____

Question 3 in ‘Food and energy’ (Figure 4.9) assesses Strand B. It is located at 541 on the scientific literacy scale. This item required students to make a valid suggestion for improving the method in an experiment about the effect of marks on the taste of apples (e.g. ‘Tim should compare more apples’ or ‘Tim should compare apples of a different type’).

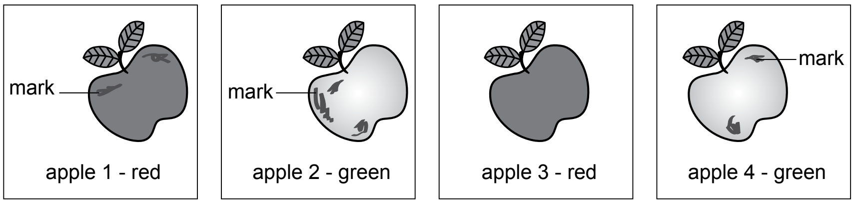
Question 2 in this item set is illustrative of Level 3.2 and is discussed on page 42.

Figure 4.9 Stimulus and items illustrating performance at Proficiency Levels 3.2 and 3.3

Supermarkets throw away fruit and vegetables that do not look nice.

Tim conducted a test to find out whether marks on apples affect the taste of the apples.

Tim had four apples.



Q2 Which apples should he have used for his test?

- ◯ apple 1 and apple 2
- ◯ apple 1 and apple 3
- ◯ apple 1 and apple 4
- ◯ apple 2 and apple 3

Q3 What could Tim do to improve his test?

____

____

Question 1 in 'Changing rocks' (Figure 4.10, with stimulus in Figure 4.7) also assesses Strand B and relates to the concept area Earth and Space. Students were required to draw a conclusion that accurately summarised the pattern in the data presented in the table (e.g. 'quartzite is more resistant to breaking down than sandstone').

This item is located at 525 on the scientific literacy scale. Students at this level of scientific literacy would be able to identify and summarise patterns in science data in the form of a rule. They would also be expected to recognise the need for improvement to the method used in an experiment.

Figure 4.10 Item illustrating performance at Proficiency Level 3.3

Q1 Write a conclusion that summarises the results from this experiment.

____
____
____

Another example at Level 3.3 is Question 4 from 'Light and shadows' (Figure 4.11). It assesses Strand C and the concept area Energy and Force. Students were presented with a brief scenario on two students measuring the length of the shadow of a flag pole at different times of the day. Students were then asked to explain why the length of the shadow of the flag pole changes during the day.

Figure 4.11 Item illustrating performance at Proficiency Level 3.3

Q4 Why does the length of the shadow of the flag pole change during the day?

____
____
____

This item is located at 542 on the scientific literacy scale. At a more general level, students who complete items requiring the scientific literacy skills and understandings at this level could be expected to describe the relationship between individual events that were experienced or reported, to generalise and apply an inferred rule by predicting future events, and to apply knowledge of a relationship to explain a reported phenomenon.

Sample items illustrating performance at Proficiency Level 3.2

Question 2 in 'Food and energy' (see Figure 4.9) at Proficiency Level 3.2 assesses Strand A. It is a multiple-choice item and is located at 486 on the scientific literacy scale. The item required students to demonstrate that they understand the need for fair testing by identifying the experimental design that exemplifies a fair test. In the design in which apples 1 and 3 are compared, the variable 'presence/absence of marks' changes while the variable 'colour of the apple' is kept the same. Knowledge and use of the terms 'variable' and 'controlled variable' are not required to respond correctly to the item.

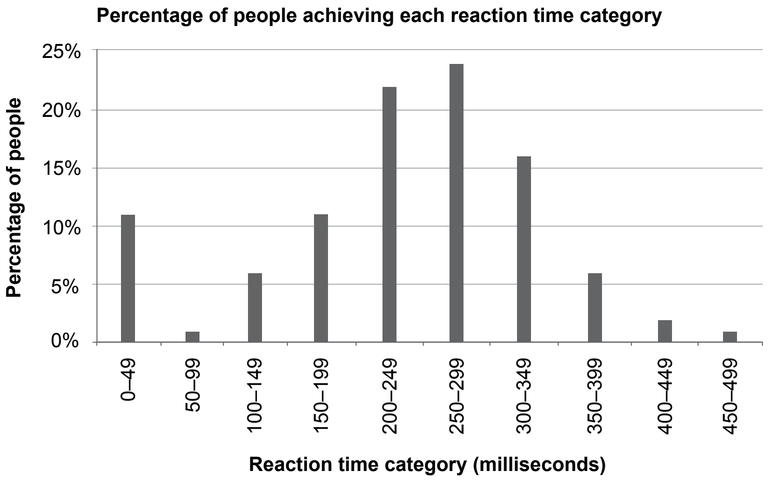
Question 9 from the Practical Task 'Reaction time' assesses Strand B and is located at 472 on the scientific literacy scale (Figure 4.12). Students were provided with a graph on the percentage of people who achieved certain reaction time categories on a computer. To respond correctly, students needed to interpret the graph and provide a justification for disagreeing with a given statement which incorrectly summarised the data presented in the graph.

Students at this level of scientific literacy would be also expected to be able to construct bar graphs from collected or given data.

Figure 4.12 Stimulus and items illustrating performance at Proficiency Levels 3.1 and 3.2

There are computer programs that test reaction time. One of these programs asks you to click on the mouse as soon as an object on the screen changes colour. The computer measures your reaction time. You can enter your results and the program then combines them with the results entered by other people.

This graph shows the reaction times entered by a large number of people.



Q8 How many categories had fewer than 5% of the people in them?

◯ 0 ◯ 1 ◯ 2 ◯ 3

Q9 A student looked at the graph and said: 'The most common reaction time was between 200 and 249 milliseconds.' Do you agree with the student's statement about the graph?

◯ Yes ◯ No

Give a reason for your answer.

____

____

Question 8 in this item set is illustrative of Level 3.1 and is discussed on page 47.

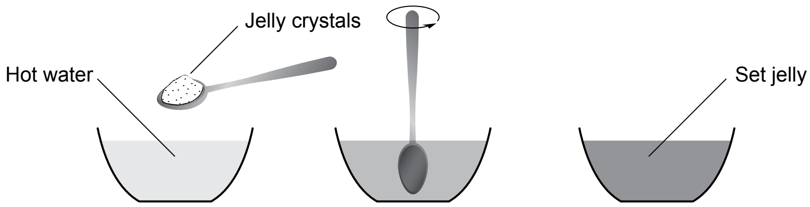
Another example of a Strand B item at Level 3.2 is Question 3 (Figure 4.13) from 'Making jelly'. The stimulus for this item set is also presented on page 39 (see Figure 4.8).

Question 3 is a multiple-choice question and is located at 442 on the scientific literacy scale. Students were asked to identify an appropriate suggestion about how to improve an investigation in the context of dissolving different amounts of jelly crystals in water.

Figure 4.13 Stimulus and item illustrating performance at Proficiency Level 3.2

Making jelly

Jelly is made by dissolving jelly crystals in hot water. After the crystals have dissolved, the mixture is cooled to allow the jelly to set. The jelly is ready to eat after it has set.



Bob made jelly using five bowls.

He did the following:

- He filled each bowl with two cups of hot water.
- He added jelly crystals to each bowl.
- He stirred the mixtures.
- He covered the bowls and placed them in the fridge.

Bob measured how long it took for each bowl of jelly to set. He recorded his results in the table below.

Bowl	Number of tablespoons of jelly crystals	Time taken to set (minutes)
1	2	210
2	4	185
3	6	Not recorded
4	8	115
5	10	90

Q3 What is one thing that Bob could have done to improve his investigation?

Bob could have

- ◯ used cold water to make the jelly crystals dissolve faster.
- ◯ put the jelly mixtures into different fridges to set.
- ◯ used different types of jelly crystals in each bowl.
- ◯ repeated the experiment to check his results.

It should be noted that this chapter includes three examples of items that assess student skills related to making improvements in experiments (see Figures 4.5, 4.9 and 4.13). These examples assess skills at different levels of sophistication, namely identifying the correct suggestion about how to improve a complex experiment and justifying the choice made in a complex experiment, formulating students' own suggestions for making improvements in a simple experiment, and identifying the correct suggestion about making an improvement in a simple experiment. These questions performed at Levels 4, 3.3 and 3.2 respectively.

Question 3 in 'Recycling' (Figure 4.14) assesses Strand C at Level 3.2. It draws from the concept area Living Things and is located at 509 on the scientific literacy scale. Students were asked to explain why decomposers, such as worms, are important in composting. Correct responses indicated that decomposers break down waste into compost or that decomposers feed on waste material.

Figure 4.14 Item illustrating performance at Proficiency Level 3.2

Q3 Why are decomposers (e.g. worms) important in composting?

____

____

____

Question 1 in 'Seed dispersal' also assesses Strand C and the concept area Living Things (Figure 4.15). It is a multiple-choice question and is located at 482 on the scientific literacy scale. Students were asked to identify a seed feature that helps seed dispersal (in this instance, hooks).

Figure 4.15 Item illustrating performance at Proficiency Level 3.2

Seed dispersal

Seed dispersal is the transport of a plant's seeds from the plant to another location by wind, water or animals.

Q1 Some seeds are dispersed on the fur of animals.

Which of the following seeds are most likely to be dispersed this way?

- ○ seeds with hooks
- ○ seeds that can float
- ○ seeds with a bright colour
- ○ seeds with a smooth surface

A third example which assesses Strand C at Level 3.2 is shown in Figure 4.16. It is an extended response item that assesses the concept area Matter. It is located at 449 on the scientific literacy scale. This item from the set 'Evaporating liquids' (see also page 34) required students to use their knowledge about factors that affect how quickly liquids evaporate to account for increased evaporation in the experiment presented.

Figure 4.16 Stimulus and item illustrating performance at Proficiency Level 3.2

Luca changed his method so that his experiment was a fair test. He measured the volumes of water and lemonade at 9 am every morning for five days. His results are shown in the table.

Table: Volume of water and lemonade on five days

Day	Volume of water (mL)	Volume of lemonade (mL)
1	100	100
2	96	94
3	92	88
4	80	73
5	76	65

Q7 Luca's results show that more liquid evaporated on Day 3 than any other day. Give one reason why this might have happened.

____

____

At a more general level, students demonstrating achievement on items such as these could be expected to describe relationships between individual events, including cause and effect relationships, either from direct or indirect experience. They can also predict outcomes by generalising and applying rules.

Sample items illustrating performance at Proficiency Level 3.1

Question 8 from the Practical Task 'Reaction time' (see Figure 4.12) is illustrative of Level 3.1. It assesses Strand B and is located at 365 on the scientific literacy scale. Students were required to interpret a column graph to identify the number of categories that matched a specified criterion.

Question 1 in 'Food and energy' (Figure 4.17) assesses Strand B. It relates to the concept area Earth and Space and is located at 290 on the scientific literacy scale. It required students to use information available in a table to rank types of food from the one that needs the smallest amount of water for production to the one that needs the largest. In general students who can answer questions requiring this level of scientific literacy could be expected to make simple measurements, display data as a table and construct simple column (bar) graphs when given the variables for each axis.

Figure 4.17 Stimulus and item illustrating performance at Proficiency Level 3.1

Q1 The table shows how much water is needed to produce different types of food.

Table: Amount of water needed to produce different types of food

Type of food	Amount of water needed to produce 1kg of food (litres)
beef	100 000
chicken	3 900
potatoes	500
white rice	1 550

Rank the types of food from the one that needs the **smallest** amount of water (1) to the one that needs the **largest** amount of water (4).

Write the numbers 1, 2, 3 and 4 in the following boxes to show your ranking.

☐ beef
☐ chicken
☐ potatoes
☐ white rice

Question 4 (Figure 4.18) from the Practical Task 'Reaction time' assesses Strand C and relates to the concept area Living Things. It is located at 278 on the scientific literacy scale.

This item required students to apply simple knowledge about the nervous system to identify the relationships between events that take place when a person tries to catch a ruler. Students at this level could be expected to describe cause and effect relationships between individual events in a familiar context.

Figure 4.18 Item illustrating performance at Proficiency Level 3.1

Q4 Which of these sequences best shows what happens when a person catches a dropped ruler?

- ◯ eyes detect ruler movement ➔ hand muscles work ➔ brain receives signal ➔ brain sends signal
- ◯ brain receives signal ➔ eyes detect ruler movement ➔ hand muscles work ➔ brain sends signal
- ◯ eyes detect ruler movement ➔ brain receives signal ➔ brain sends signal ➔ hand muscles work
- ◯ brain sends signal ➔ brain receives signal ➔ eyes detect ruler movement ➔ hand muscles work

Another item which also addressed Strand C at this level is Question 1 from the item set 'Light and shadows' (Figure 4.19) which draws on the Energy and Force concept area. It is located at 315 on the scientific literacy scale.

This item required students to identify the relationship between blocking the path of light and the formation of a shadow. At a more general level, students who responded correctly to this item or items requiring the same level of scientific literacy could be expected to apply known rules to explain specific instances related to personal experience.

Figure 4.19 Stimulus and item illustrating performance at Proficiency Level 3.1

Light and shadows

The photograph shows a boy and his shadow.

Q1 Shadows form when

- ◯ a light is turned off.
- ◯ light reflects off an object.
- ◯ light shines through an object.
- ◯ the path of light is blocked by an object.

Sample items illustrating performance at Proficiency Level 2 and below

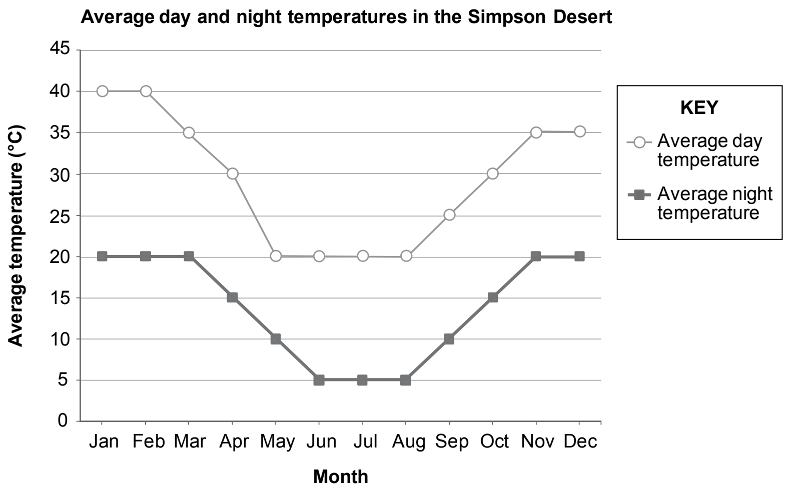
Question 1 in 'Life in the desert' (Figure 4.20) assesses Strand B and relates to the concept area Earth and Space. It is a short-answer response item located at 226 on the scientific literacy scale. Students needed to locate a piece of data in a supplied graph displaying day and night temperatures in the Simpson Desert.

Figure 4.20 Stimulus and item illustrating performance at Proficiency Level 2 and below

Life in the desert

The Simpson Desert is located in Central Australia. It is a dry environment with long periods of no rainfall.

The graph shows the average day and night temperatures in the Simpson Desert for each month of the year.



Q1 What is the highest average day temperature in the Simpson Desert?

__________°C

Question 2 in 'Recycling' (Figure 4.21) is another example of an item at Proficiency Level 2 or below. It assesses Strand C and the concept area Earth and Space. It is located at 197 on the scientific literacy scale. This multiple-choice item required students to use their real-world knowledge or experience to identify which object would be added to compost.

Figure 4.21 Stimulus and item illustrating performance at Proficiency Level 2 and below

Q2 Jason prepared chicken vegetable soup.

Soup Ingredients

1 chicken wrapped in plastic
12 cups of water
2 potatoes
2 onions
2 sticks of celery
parsley
1 can of tomatoes

Compost heap

Jason had a compost heap at his house. Which waste could Jason add to his compost?

- ◯ empty tomato can
- ◯ peel from potatoes
- ◯ plastic wrapper
- ◯ nylon onion bag

At a more general level, students who could respond correctly to items requiring this level of scientific literacy could be expected to describe changes to, differences between or properties of objects or events based on first-hand concrete experiences. They can also compare aspects of data in simple supplied tables of results.

Table 4.1 shows the percentage correct for the illustrative sample items found in this chapter. Table 4.2 provides results on the illustrative sample items by state and territory. Further details about the performance of the items are provided in the 2012 Technical Report (available at www.nap.edu.au).

Table 4.1 Summary of results for sample items 2012

Page	Figure	Unit	Question	% correct	Level	Strand	Scaled score
34	4.4	Evaporating liquids	4	3.7	≥ 4	A	848
35	4.5	Reaction time	6	5.4	≥ 4	B	811
36	4.6	Solar energy	2	9.1	≥ 4	C	782
37	4.7	Changing rocks	3	8.9	≥ 4	C	734
41	4.11	Light and shadows	4	33.6	3.3	C	542
39	4.8	Making jelly	1	34.6	3.3	A	542
40	4.9	Food and energy	3	33.9	3.3	B	541
41	4.10	Changing rocks	1	38.3	3.3	B	525
45	4.14	Recycling	3	39.8	3.2	C	509
40	4.9	Food and energy	2	44.5	3.2	A	486
45	4.15	Seed dispersal	1	45.9	3.2	C	482
43	4.12	Reaction time	9	50.0	3.2	B	472
46	4.16	Evaporating liquids	7	53.2	3.2	C	449
44	4.13	Making jelly	3	55.1	3.2	B	442
43	4.12	Reaction time	8	71.3	3.1	C	365
49	4.19	Light and shadows	1	75.4	3.1	C	315
47	4.17	Food and energy	1	81.1	3.1	B	290
48	4.18	Reaction time	4	82.7	3.1	B	278
50	4.20	Life in the desert	1	86.6	≤ 2	B	226
51	4.21	Recycling	2	87.3	≤ 2	C	197

Table 4.2 Performance of students from each state and territory on sample items 2012

Page	Figure	Unit	Question	Level	Strand	Scaled score	Percentage correct								
							AUST	ACT	NSW	NT	QLD	SA	TAS	VIC	WA
34	4.4	Evaporating liquids	4	≥ 4	A	848	3.7	4.4	3.1	1.7	3.7	3.7	4.2	2.6	5.0
35	4.5	Reaction time	6	≥ 4	B	811	5.4	5.2	5.6	1.9	5.3	4.6	7.1	4.5	6.9
36	4.6	Solar energy	2	≥ 4	C	782	9.1	11.5	9.3	6.3	8.7	6.7	9.9	9.7	9.6
37	4.7	Changing rocks	3	≥ 4	C	734	8.9	10.6	8.7	7.4	9.4	6.9	9.8	8.5	9.6
41	4.11	Light and shadows	4	3.3	C	542	33.6	39.8	32.7	27.1	30.2	32.9	33.4	32.1	38.0
39	4.8	Making jelly	1	3.3	A	542	34.6	44.9	31.8	26.7	31.0	32.5	37.8	35.6	36.3
40	4.9	Food and energy	3	3.3	B	541	33.9	38.9	34.7	24.8	30.0	29.6	37.1	35.8	36.9
41	4.10	Changing rocks	1	3.3	B	525	38.3	48.8	37.4	27.6	35.3	36.9	37.7	37.9	41.2
45	4.14	Recycling	3	3.2	C	509	39.8	51.5	39.3	26.4	37.2	39.9	39.7	39.0	41.2
40	4.9	Food and energy	2	3.2	A	486	44.5	48.6	42.6	44.2	43.7	42.9	47.1	40.9	48.0
45	4.15	Seed dispersal	1	3.2	C	482	45.9	58.9	44.3	36.5	42.3	47.3	47.1	43.4	46.4
43	4.12	Reaction time	9	3.2	B	472	50.0	53.1	54.5	37.4	47.9	45.2	48.4	54.6	52.1
46	4.16	Evaporating liquids	7	3.2	C	449	53.2	59.0	53.0	37.7	56.3	50.0	47.6	52.0	59.2
44	4.13	Making jelly	3	3.2	B	442	55.1	59.3	53.3	48.6	56.4	57.1	55.0	51.1	56.9
43	4.12	Reaction time	8	3.1	C	365	71.3	77.8	72.1	61.9	68.9	69.4	69.6	71.1	75.0
49	4.19	Light and shadows	1	3.1	C	315	75.4	77.8	70.4	74.2	75.3	77.7	75.8	74.2	77.9
47	4.17	Food and energy	1	3.1	B	290	81.1	84.1	82.8	64.8	77.8	81.8	80.3	84.3	82.9
48	4.18	Reaction time	4	3.1	B	278	82.7	86.0	84.8	76.1	81.1	80.1	84.7	82.3	84.0
50	4.20	Life in the desert	1	≤ 2	B	226	86.6	89.1	88.9	77.2	86.7	87.8	85.8	86.7	84.8
51	4.21	Recycling	2	≤ 2	C	197	87.3	90.8	86.5	76.9	84.2	89.9	91.2	88.1	86.7

Chapter 5
Distribution of students within Proficiency Levels for 2012 with comparisons to previous cycles

Introduction

Student achievement in scientific literacy is reported against three broad levels of achievement, with Level 3 being further segmented into three sub-levels represented by 3.1, 3.2 and 3.3. The Proficient Standard in scientific literacy is situated at the boundary between Level 3.1 and 3.2.

Student performance by Proficiency Level

The 2012 distribution of students across Proficiency Levels is shown in Table 5.1.

The National Assessment Program – Science Literacy (NAP–SL) assessment was constructed with the expectation that most Year 6 students would demonstrate the understandings and skills described at Proficiency Level 3. Table 5.1 shows that, at the national level, nine per cent of students did not reach Proficiency Level 3.

However, in the Northern Territory 31.1 per cent of students did not demonstrate scientific literacy corresponding to Proficiency Level 3.

Table 5.1 Percentage of students in Proficiency Levels by state and territory in 2012

State/ Territory	Level 2 and below	Level 3.1	Level 3.2	Level 3.3	Level 4 and above
ACT	4.4 (±1.7)	30.3 (±4.6)	49.4 (±4.2)	15.0 (±4.1)	0.9 (±0.7)
NSW	9.2 (±2.5)	39.8 (±3.3)	40.9 (±3.8)	9.6 (±2.5)	0.4 (±0.4)
NT	31.1 (±9.6)	37.9 (±7.0)	26.5 (±6.4)	4.3 (±3.0)	0.2 (±0.4)
QLD	8.8 (±1.6)	41.4 (±2.9)	41.8 (±3.1)	8.0 (±1.6)	0.1 (±0.2)
SA	8.8 (±1.9)	40.1 (±3.4)	43.5 (±3.6)	7.5 (±1.9)	0.1 (±0.2)
TAS	9.6 (±2.3)	39.1 (±4.2)	40.2 (±4.2)	10.8 (±3.2)	0.3 (±0.5)
VIC	8.3 (±2.2)	40.4 (±4.0)	43.4 (±3.8)	7.6 (±2.3)	0.2 (±0.3)
WA	8.2 (±1.9)	35.5 (±3.3)	44.0 (±3.3)	12.0 (±2.4)	0.4 (±0.4)
AUST	9.0 (±1.0)	39.6 (±1.6)	42.1 (±1.7)	9.0 (±1.1)	0.3 (±0.2)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

Table 5.1 also shows that 51.4 per cent of students demonstrated skills and understandings that placed them at or above Proficiency Level 3.2 (i.e. at or above the Proficient Standard). Results further indicate that, at both the national and jurisdiction levels, the proportion of students performing at Proficiency Level 4 and above accounts for only 0.1 to 0.9 per cent of the student population.

The proportion of students who demonstrated scientific literacy skills and understandings at or above the Proficient Standard is presented in Table 5.2, with jurisdictions listed in rank order according to the percentage of students operating at or above the Proficient Standard. Table 5.2 also contains the corresponding results and ranking for the 2006 and 2009 assessments.

In 2012, 51.4 per cent of students were found to be performing at or above the Proficient Standard (i.e. Proficiency Level 3.2 or above) at the national level. This compares to the 54.3 and 51.9 per cent of students found to be performing at or above the Proficient Standard at the national level in 2006 and 2009 respectively. At the national level the difference between 2012 and 2006 in the percentage of students achieving at or above the Proficient Standard is 2.9 per cent, which is not statistically significant. It needs to be noted that no significant difference was found between the 2012 and 2009 assessments.

Table 5.2 shows that Western Australia improved its ranking from seventh in 2006 to second in 2012. The difference in the percentages of students in Western Australia achieving at or above the Proficient Standard in 2006 and 2012 is statistically significant. Tasmania improved its ranking from fourth to third.

However, the difference between the 2006 and 2012 results for Tasmania is not statistically significant. The differences between the 2006 and 2012 results for Victoria, New South Wales and Queensland, where the 2012 ranking is lower compared to that of 2006, are also not statistically significant; nor are those for the Australian Capital Territory, South Australia and Northern Territory, whose rankings did not change in 2012 from those found in 2006.

Table 5.2 Jurisdictions by percentage of students at or above the Proficient Standard in rank order for 2006, 2009 and 2012

Rank by jurisdiction	2006		2009		2012	
	State/ Territory	At or above the Proficient Standard	State/ Territory	At or above the Proficient Standard	State/ Territory	At or above the Proficient Standard
1	ACT	62.0 (±5.6)	ACT	61.2 (±4.8)	ACT	65.3 (±5.3)
2	VIC	58.3 (±5.0)	VIC	54.6 (±4.6)	WA	56.4 (±4.2)
3	NSW	57.4 (±4.3)	WA	53.3 (±4.5)	TAS	51.3 (±5.4)
4	TAS	57.4 (±5.5)	NSW	53.0 (±5.0)	VIC	51.3 (±4.7)
5	SA	51.6 (±4.7)	TAS	49.8 (±6.0)	SA	51.1 (±3.9)
6	QLD	49.2 (±3.8)	QLD	48.8 (±3.8)	NSW	50.9 (±4.3)
7	WA	46.6 (±4.7)	SA	46.5 (±5.0)	QLD	49.9 (±3.3)
8	NT	38.4 (±6.5)	NT	33.6 (±7.5)	NT	31.0 (±7.6)
	AUST	54.3 (±2.1)	AUST	51.9 (±2.2)	AUST	51.4 (±2.0)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

The trend in distribution of students at the national level across Proficiency Levels in all four NAP–SL test cycles is presented in Table 5.3.

Table 5.3 Percentage distribution of students across Proficiency Levels in 2003, 2006, 2009 and 2012

AUST	Level 2 and below	Level 3.1	Level 3.2	Level 3.3	Level 4 and above
2003	4.1 (±0.7)	36.5 (±1.7)	52.2 (±1.7)	7.1 (±0.9)	0.1 (±0.1)
2006	8.6 (±1.1)	37.1 (±1.7)	44.2 (±1.8)	9.6 (±1.2)	0.5 (±0.4)
2009	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)
2012	9.0 (±1.0)	39.6 (±1.6)	42.1 (±1.7)	9.0 (±1.1)	0.3 (±0.2)

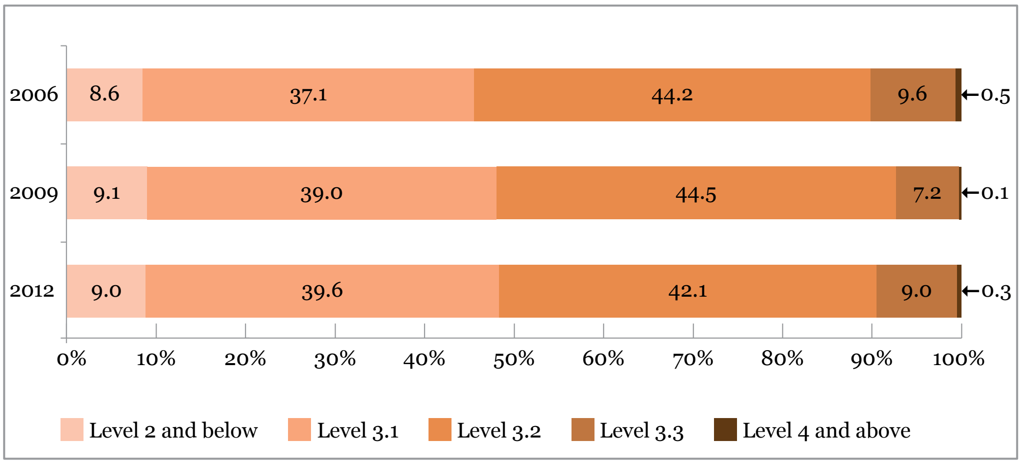
Notes: Figures in parentheses refer to 95 per cent confidence intervals.
2003 results are indicative only.

Table 5.3 shows that the distribution of students across Proficiency Levels at the national level remained relatively stable across the assessments in 2006, 2009 and 2012.

This table demonstrates that in 2012, 51.4 per cent of students were proficient at Level 3.2 and above. For 2006 and 2009, these percentages were 54.3 and 51.9 respectively. These differences are not statistically significant. Comparisons between 2003 figures with 2006, 2009 and 2012 figures should be interpreted with caution. As noted in Chapter 3 of this report, there were important differences between the test design in the 2003 assessment and those in the 2006, 2009 and 2012 assessments. The assessments in 2006, 2009 and 2012 included wider coverage of the assessment domain and samples were more inclusive of students in remote geographic locations.

Figure 5.1 illustrates the percentage of students across Proficiency Levels in the 2006, 2009 and 2012 assessments.

Figure 5.1 Percentage of students across Proficiency Levels in the 2006, 2009 and 2012 assessments



This figure shows that while the percentage of students who performed at or above Proficiency Level 3.2 in 2009 and 2012 is very similar (51.9 and 51.4 in 2009 and 2012 respectively), the distribution of students within levels 3.2 and 3.3 is slightly different. In the 2009 assessment, the percentage of students who performed at levels 3.2 and 3.3 was 44.5 and 7.2 per cent, respectively. In 2012, however, 42.1 and 9 per cent of students performed at levels 3.2 and 3.3, respectively. This means that the movement from Proficiency Level 3.2 to Proficiency Level 3.3 in the 2012 assessment was slightly higher. However, this movement was not statistically significant.

Chapter 6
Sub-group results and comparisons by mean and Proficiency Levels

Introduction

In this chapter, the differences in achievement in terms of mean scores and the distribution of results for male and female students, Indigenous and non-Indigenous students and students from diverse geographic locations and language backgrounds are considered across the states and territories. This chapter also contains the same information, where applicable, from the 2003, 2006 and 2009 assessments in order to allow trends in results for the National Assessment Program – Science Literacy (NAP–SL) to be investigated. However, as in previous cycles, differences in achievement cannot be reported based on Parent Occupation or Parent Education, since insufficient data were provided by schools to enable any meaningful analysis. For 2012, an analysis comparing the performance of students from English-speaking backgrounds and those from language backgrounds other than English has been made because these data were adequately collected.

Gender results by mean

Mean scores for male and female students across jurisdictions are presented in Table 6.1. This table also provides information about the proportion of male and female students in the sample.

Table 6.1 Mean scores for male and female students by state and territory in 2012

State/Territory	Percentage of males in sample	Mean score	
		Male	Female
ACT	49.8	432 (±13.7)	425 (±18.0)
NSW	51.1	395 (±13.0)	394 (±8.7)
NT	46.6	316 (±35.8)	322 (±30.6)
QLD	50.0	391 (±8.1)	392 (±7.6)
SA	50.4	391 (±8.8)	393 (±10.7)
TAS	50.0	395 (±13.2)	396 (±15.5)
VIC	52.3	392 (±11.5)	395 (±10.6)
WA	51.8	405 (±11.5)	407 (±11.0)
AUST	50.6	394 (±5.6)	395 (±4.4)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

It can be seen from Table 6.1 that at the national level the mean score for females is one point higher than that for males. However, this difference is not statistically significant nor are any other gender comparisons outlined in Table 6.1. The table indicates that males in the Australian Capital Territory were the highest-performing group overall.

Gender results by Proficiency Levels

Table 6.2 shows the distribution of results across the Proficiency Levels for males and females and demonstrates that there were no significant differences in performance. The table also shows the percentage of students at or above the Proficient Standard in scientific literacy.

Table 6.2 Percentage distribution of male and female students across Proficiency Levels by state and territory in 2012

State / Territory	Gender	Level 2 and below	Level 3.1	Level 3.2	Level 3.3	Level 4 and above	At or above the Proficient Standard
ACT	Male	4.5 (±2.4)	28.7 (±4.9)	50.0 (±5.6)	15.7 (±4.4)	1.1 (±1.2)	66.8 (±5.5)
	Female	4.3 (±2.1)	31.8 (±7.0)	48.8 (±6.1)	14.4 (±6.1)	0.7 (±0.8)	63.9 (±7.7)
NSW	Male	10.4 (±3.3)	37.9 (±4.6)	40.9 (±5.1)	10.2 (±3.4)	0.6 (±0.7)	51.7 (±5.9)
	Female	8.1 (±2.6)	41.8 (±3.7)	40.9 (±4.1)	9.1 (±2.5)	0.3 (±0.4)	50.2 (±4.1)
NT	Male	32.2 (±10.8)	34.9 (±7.5)	28.7 (±8.2)	4.2 (±3.2)	0.0 (±0.0)	32.9 (±9.2)
	Female	29.6 (±9.7)	40.7 (±8.7)	24.8 (±6.9)	4.5 (±3.8)	0.4 (±0.8)	29.7 (±8.4)
QLD	Male	9.9 (±2.3)	39.9 (±4.1)	41.3 (±4.4)	8.8 (±2.2)	0.1 (±0.3)	50.2 (±4.4)
	Female	7.6 (±2.3)	42.8 (±3.7)	42.2 (±3.5)	7.2 (±2.4)	0.1 (±0.3)	49.6 (±3.8)
SA	Male	9.6 (±2.5)	39.6 (±4.3)	42.5 (±4.7)	8.2 (±2.5)	0.1 (±0.4)	50.8 (±4.9)
	Female	7.9 (±2.8)	40.7 (±5.2)	44.5 (±5.2)	6.9 (±2.4)	0.1 (±0.2)	51.4 (±5.6)
TAS	Male	9.0 (±3.1)	40.3 (±5.1)	39.2 (±5.4)	11.2 (±3.6)	0.3 (±0.7)	50.7 (±6.2)
	Female	10.3 (±3.4)	37.9 (±5.6)	41.3 (±5.7)	10.3 (±4.4)	0.3 (±0.7)	51.9 (±6.6)
VIC	Male	9.4 (±2.7)	39.7 (±5.0)	42.9 (±4.6)	7.8 (±2.9)	0.2 (±0.4)	50.9 (±5.7)
	Female	7.1 (±2.6)	41.2 (±4.9)	44.1 (±4.6)	7.4 (±2.7)	0.3 (±0.4)	51.7 (±5.3)
WA	Male	8.6 (±2.3)	34.6 (±3.9)	44.7 (±4.7)	11.6 (±3.3)	0.5 (±0.5)	56.8 (±5.0)
	Female	7.7 (±2.7)	36.4 (±5.2)	43.2 (±4.7)	12.4 (±3.1)	0.2 (±0.4)	55.9 (±5.3)
AUST	Male	9.9 (±1.4)	38.4 (±2.2)	41.9 (±2.3)	9.4 (±1.5)	0.4 (±0.2)	51.7 (±2.6)
	Female	8.0 (±1.3)	41.0 (±1.9)	42.3 (±1.9)	8.6 (±1.3)	0.2 (±0.2)	51.1 (±2.2)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

Trend analysis by gender

Table 6.3 shows the mean scores achieved by male and female students at the national level, as observed in the 2003, 2006, 2009 and 2012 NAP–SL assessments.

Table 6.3 Mean scores for male and female students in 2003, 2006, 2009 and 2012

AUST	Percentage of males in sample	Mean score	
		Male	Female
2003	51.1	412 (±4.7)	405 (±4.0)
2006	50.8	402 (±6.4)	398 (±5.1)
2009	50.5	393 (±6.0)	391 (±5.2)
2012	50.6	394 (±5.6)	395 (±4.4)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
2003 results are indicative only.

In conducting statistical testing of the differences between the 2006, 2009 and 2012 results, variability in the data caused by equating the 2012 results to the 2006 scale was taken into account (for more detailed information see the 2012 Technical Report).

As can be seen from Table 6.3, in 2012 females achieved a slightly higher mean than males. However, this difference is not statistically significant. In previous cycles, males achieved a slightly higher mean than females. These differences were not statistically significant.

Table 6.4 shows the distribution of performance across all Proficiency Levels for males and females.

Table 6.4 Percentage distribution across Proficiency Levels of male and female students in 2003, 2006, 2009 and 2012

AUST	Level 2 and below		Level 3.1		Level 3.2		Level 3.3		Level 4 and above	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
2003	4.1 (±1.3)	4.2 (±0.8)	35.3 (±2.2)	37.6 (±2.1)	52.5 (±2.0)	52.0 (±2.1)	8.1 (±1.4)	6.1 (±1.0)	0.0 (±0.2)	0.1 (±0.1)
2006	9.0 (±1.4)	8.2 (±1.3)	36.1 (±2.2)	38.2 (±2.1)	43.6 (±2.2)	44.8 (±2.1)	10.6 (±1.7)	8.5 (±1.3)	0.7 (±0.6)	0.3 (±0.3)
2009	9.7 (±1.5)	8.3 (±1.3)	38.0 (±2.1)	40.2 (±2.4)	44.2 (±2.3)	45.1 (±2.2)	8.0 (±0.8)	6.5 (±0.6)	0.1 (±0.1)	0.1 (±0.1)
2012	9.9 (±1.4)	8.0 (±1.3)	38.4 (±2.2)	41.0 (±1.9)	41.9 (±2.3)	42.3 (±1.9)	9.4 (±1.5)	8.6 (±1.3)	0.4 (±0.2)	0.2 (±0.2)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
2003 results are indicative only.

As can be seen from Table 6.4, the pattern of results remained relatively stable for the percentage of male and female students achieving various Proficiency Levels across the 2006, 2009 and 2012 assessments.

Table 6.5 shows the distribution of the percentage of males and females at or above the Proficient Standard in 2006, 2009 and 2012.

Table 6.5 Percentage of male and female students at or above the Proficient Standard in 2006, 2009 and 2012

AUST	At or above the Proficient Standard	
	Male	Female
2006	54.9 (±2.5)	53.7 (±2.3)
2009	52.3 (±2.6)	51.7 (±2.6)
2012	51.7 (±2.6)	51.1 (±2.2)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
The percentage of males and females achieving at or above the Proficient Standard is not available for 2003 due to changes in scaling that occurred between 2003 and 2006.

Indigenous students

The results for Indigenous students relative to non-Indigenous students are shown in Table 6.6.

Table 6.6 Mean scores for Indigenous and non-Indigenous students in 2003, 2006, 2009 and 2012

AUST	Mean score	
	Indigenous	Non-Indigenous
2003	350 (±11.3)	412 (±3.7)
2006	311 (±29.4)	402 (±5.8)
2009	297 (±16.0)	397 (±5.0)
2012	303 (±15.1)	399 (±4.5)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
2003 results are indicative only.

Table A2.6 in Appendix 2 provides a breakdown of the number of students in the achieved sample by Indigenous status across jurisdictions. Table A2.7 in Appendix 2 provides a breakdown by Indigenous status across geographic locations.

In 2012, the mean score for Indigenous students was 303, indicating that they did not perform as well as non-Indigenous students with a mean score of 399. This difference is statistically significant, as it also was in 2006 and 2009.

The differences between the 2006, 2009 and 2012 mean scores are not statistically significant for either Indigenous or non-Indigenous students.

Table 6.7 contains a summary of differences in distribution across Proficiency Levels between Indigenous and non-Indigenous students in 2003, 2006, 2009 and 2012.

Table 6.7 Percentage distribution of Indigenous and non-Indigenous students across Proficiency Levels in 2003, 2006, 2009 and 2012

AUST	Level 2 and below		Level 3.1		Level 3.2		Level 3.3		Level 4 and above	
	Indigenous	Non-Indigenous	Indigenous	Non-Indigenous	Indigenous	Non-Indigenous	Indigenous	Non-Indigenous	Indigenous	Non-Indigenous
2003	15.9 (±1.3)	3.6 (±0.6)	51.6 (±6.3)	35.7 (±1.7)	30.9 (±6.7)	53.3 (±1.7)	1.7 (±2.0)	7.4 (±0.9)	0.0 (±0.0)	0.1 (±0.1)
2006	31.4 (±8.1)	8.1 (±1.1)	43.1 (±7.5)	37.3 (±1.8)	22.3 (±7.4)	44.3 (±1.9)	3.1 (±3.9)	9.8 (±1.4)	0.1 (±0.4)	0.6 (±0.5)
2009	38.6 (±6.8)	7.4 (±0.1)	41.8 (±5.8)	38.7 (±1.9)	18.6 (±6.0)	46.1 (±1.8)	1.0 (±1.5)	7.6 (±1.2)	0.0 (±0.0)	0.1 (±0.1)
2012	33.4 (±6.3)	7.9 (±1.0)	46.5 (±7.3)	39.3 (±1.6)	19.0 (±5.7)	43.1 (±1.8)	1.1 (±1.4)	9.4 (±1.2)	0.0 (±0.0)	0.3 (±0.2)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
2003 results are indicative only.

Table 6.7 shows that in 2012, approximately 33 per cent of Indigenous students were working at Level 2 and below, whereas less than eight per cent of non-Indigenous students were working at that level. The percentage of Indigenous students performing at Level 2 and below is slightly lower than in 2009 and about the same as in 2006. However, the differences are not statistically significant.

Table 6.8 shows the percentage of Indigenous and non-Indigenous students at or above the Proficient Standard in 2006, 2009 and 2012.

Table 6.8 Percentage of Indigenous and non-Indigenous students achieving at or above the Proficient Standard in 2006, 2009 and 2012

AUST	At or above the Proficient Standard	
	Indigenous	Non-Indigenous
2006	25.5 (±10.0)	54.7 (±2.2)
2009	19.6 (±6.0)	53.9 (±2.3)
2012	20.1 (±5.8)	52.8 (±2.0)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
The percentage of Indigenous and non-Indigenous students achieving at or above the Proficient Standard is not available for 2003 due to changes in scaling that occurred between 2003 and 2006.

Table 6.8 also shows that in 2012, 20.1 per cent of Indigenous students performed at or above the Proficient Standard. This represents a decrease of approximately five and a half percentage points compared to 2006. This difference is not statistically significant. The percentages of Indigenous students performing at or above the Proficient Standard in 2009 and 2012 are very similar.

Geographic location of schools

Table 6.9 shows the mean scaled scores in 2009 and 2012 for students attending schools in different geographic locations. The table shows that differences between the performance of students living in metropolitan areas and provincial areas were statistically significant. Students living in remote and very remote areas showed significantly lower performances in scientific literacy than students attending schools in metropolitan areas. However, differences between the performance of students living in remote and very remote areas and provincial areas were not statistically significant.

In 2012, students attending schools in metropolitan areas achieved the highest mean scaled scores. Similar results were found in the 2009 assessment. However, the results cannot be compared directly with those obtained in 2003 and 2006 as categories for the geographic locations changed between 2006 and 2009.

Table 6.9 Mean scores of students by school geographic location in 2009 and 2012

MCEECDYA geographic location category	2009		2012	
	Percentage of students	Mean score	Percentage of students	Mean score
Metropolitan areas	72.3	395 (±6.2)	72.9	400 (±5.2)
Provincial areas	24.7	389 (±7.9)	25.3	381 (±9.5)
Remote and very remote areas	3.0	336 (±23.6)	1.9	349 (±31.0)
AUST	100.0	392 (±5.1)	100.0	394 (±4.4)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
The percentage of students in geographic location regions are weighted to reflect the population percentages. They are not the percentages of students in the sample.

The percentages of students in this table are weighted to reflect the population of Year 6 students in Australia. They are not the percentages of students in the sample. For more information about the applied weights and the sampling design refer to the 2012 Technical Report.

Table 6.10 shows the percentage of students across Proficiency Levels by the geographic location of the sampled schools in 2009 and 2012. It can be seen that the percentage of students across Proficiency Levels in the 2009 and 2012 assessments is similar.

Table 6.10 Percentage distribution across Proficiency Levels by school geographic location in 2009 and 2012

MCEECDYA geographic location category	Level 2 and below		Level 3.1		Level 3.2		Level 3.3		Level 4 and above	
	2009	2012	2009	2012	2009	2012	2009	2012	2009	2012
Metropolitan areas	8.4 (±1.5)	7.8 (±1.2)	38.1 (±2.0)	39.0 (±2.0)	45.5 (±2.1)	42.9 (±1.9)	7.8 (±1.4)	10.0 (±1.5)	0.1 (±0.1)	0.4 (±0.2)
Provincial areas	8.6 (±1.7)	11.3 (±2.8)	41.9 (±3.4)	41.7 (±3.4)	43.5 (±3.3)	40.5 (±3.8)	6.0 (±1.5)	6.5 (±1.6)	0.1 (±0.1)	0.1 (±0.1)
Remote and very remote areas	28.2 (±8.8)	23.2 (±9.5)	37.9 (±8.4)	35.1 (±7.4)	29.6 (±7.2)	35.5 (±9.2)	4.1 (±3.7)	6.2 (±3.5)	0.2 (±0.5)	0.1 (±0.2)
AUST	9.1 (±1.2)	9.0 (±1.0)	39.0 (±1.7)	39.6 (±1.6)	44.5 (±1.8)	42.1 (±1.7)	7.2 (±1.1)	9.0 (±1.1)	0.1 (±0.1)	0.3 (±0.2)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

Table 6.11 shows the percentage of students achieving at or above the Proficient Standard in 2009 and 2012 by geographic location.

Table 6.11 Percentage of students achieving at or above the Proficient Standard in 2009 and 2012 by geographic location

AUST	At or above the Proficient Standard		
	Metropolitan areas	Provincial areas	Remote and very remote areas
2009	53.4 (±2.6)	49.5 (±4.1)	33.9 (±8.2)
2012	53.2 (±2.3)	47.0 (±4.4)	41.7 (±9.2)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

Table 6.11 shows that 41.7 per cent of students who attended schools located in remote and very remote areas were at or above the Proficient Standard in scientific literacy in 2012, while in 2009 only 33.9 per cent of students from the same geographic location were at or above the Proficient Standard. However, this difference is not statistically significant.

Language background

In 2006, an online system for collecting demographic information about students participating in NAP–SL was implemented. However, the system did not deliver accurate and complete information, therefore only 2003, 2009 and 2012 data are presented here.

In 2012, data were collected to understand the language background of students in Year 6. Results for students from a language background other than English (LBOTE) and students from an English speaking background (ESB) were compared. Table 6.12 provides a comparison of results between 2003, 2009 and 2012. It should be noted that a student's language background does not indicate the student's proficiency in English.

Table 6.12 Comparison of mean scores by student language background in 2003, 2009 and 2012

AUST	Mean score	
	LBOTE	ESB
2003	374 (±10.7)	405 (±4.5)
2009	384 (±13.0)	396 (±4.7)
2012	389 (±13.7)	397 (±4.5)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
2003 results are indicative only.

In 2012, LBOTE students, with a mean score of 389, did not perform as well as ESB students, with a mean score of 397. However, this difference is not statistically significant. A similar trend was observed in the 2009 assessment.

The distribution of students across the Proficiency Levels by language background is given in Table 6.13.

Table 6.13 Percentage distribution across Proficiency Levels by student language background in 2003, 2009 and 2012

AUST	Level 2 and below		Level 3.1		Level 3.2		Level 3.3		Level 4 and above	
	LBOTE	ESB	LBOTE	ESB	LBOTE	ESB	LBOTE	ESB	LBOTE	ESB
2003	7.4 (±2.4)	3.5 (±0.6)	43.1 (±4.0)	35.3 (±1.8)	44.7 (±4.1)	53.5 (±1.7)	4.9 (±1.9)	7.5 (±0.9)	0.0 (±0.0)	0.1 (±0.1)
2009	12.4 (±3.2)	7.7 (±0.1)	38.7 (±3.9)	38.9 (±1.9)	40.0 (±3.5)	46.3 (±1.9)	8.8 (±13.2)	7.0 (±1.1)	0.2 (±0.3)	0.1 (±0.1)
2012	11.0 (±2.5)	8.2 (±1.1)	41.4 (±4.4)	39.2 (±1.6)	36.7 (±3.7)	43.6 (±1.8)	10.2 (±4.1)	8.9 (±1.1)	0.7 (±0.6)	0.2 (±0.1)

Notes: Figures in parentheses refer to 95 per cent confidence intervals.
2003 results are indicative only.

Table 6.13 shows that 11 per cent of LBOTE students were working at Level 2 and below, whereas only 8.2 per cent of ESB students were working at the same level. Looking at the other end of the scale, 10.9 per cent of LBOTE students achieved Proficiency Level 3.3 or above, compared to 9.1 per cent of ESB students. A similar trend was observed in the 2009 assessment.

Table 6.14 shows the percentage of students achieving at or above the Proficient Standard by language background in 2009 and 2012. It shows that the differences between the percentages of LBOTE and ESB students performing at or above the Proficient Standard were not statistically significant.

Table 6.14 Percentage of students achieving at or above the Proficient Standard in 2009 and 2012 by student language background

AUST	At or above the Proficient Standard	
	LBOTE	ESB
2009	48.9 (±4.9)	53.4 (±2.3)
2012	47.6 (±5.4)	52.6 (±2.1)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

Due to scaling changes between 2003 and 2006 and insufficient data being provided in 2006, only 2009 information is available for comparison with the 2012 achievement data.

Chapter 7
Student Survey

Introduction

In 2009, a Student Survey about students' attitudes to and interests in science and science experiences in school was introduced into the National Assessment Program – Science Literacy (NAP–SL). This addition to the testing program was continued in 2012. The survey was conducted following completion of the practical task.

A survey instrument consisting of 42 items was developed in consultation with the Science Literacy Review Committee (SLRC) and trialled. The instrument was guided by the items and results of the 2009 survey as well as by recommendations from the SLRC. Following analysis of the responses from the trial and feedback from the SLRC, 34 items were selected for inclusion in the final survey form. Approximately half of the items from the 2009 Student Survey were included in the final form. The survey required Year 6 students to provide responses which varied from simple Yes/No responses to Likert scale and frequency rating scale formats.

Thematically, the Student Survey was divided into eight categories:

1. Interest in science
2. Self-concept in science
3. Value of science
4. Perceptions of science
5. Science-related activities outside school
6. Science-related activities at school
7. Science teaching and investigations
8. Science topics studied at school

The survey results provide important insights into Australian Year 6 students' perceptions and attitudes toward science as they are based on a large sample of students. The sample comprised the number of students in the achieved sample (13 236 students) who also completed the Student Survey.

In this chapter, student responses to survey items are presented in frequency distribution tables. In addition, the chapter summarises the results of the statistical analysis that was conducted for the purpose of examining the relationship between students' responses to specific Student Survey items and their achievements in the 2012 NAP–SL.

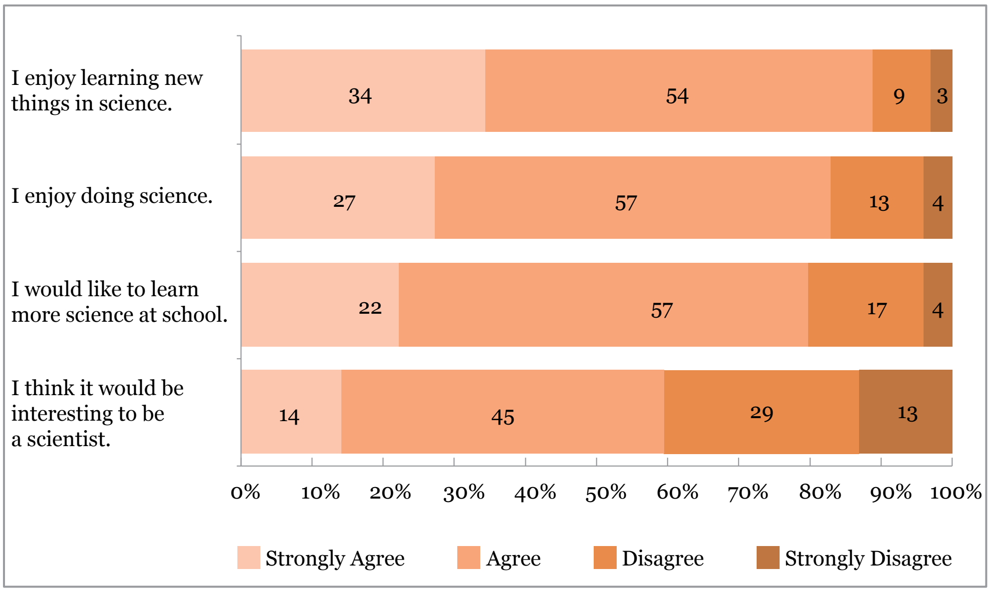
Distribution of students' responses to the Student Survey

As noted in the previous section, the Student Survey items can be divided into eight categories. A description of each category is provided below, followed by the results of the Student Survey presented as percentage frequencies. The percentages provided in each figure are derived from the responses received from all students for a particular response category. The response categories are defined underneath each figure.

In all the following figures, percentages have been rounded and may not add up to 100.

Interest in science. This section included four statements which sought to elicit whether students had an interest in learning and doing science as well as their interest in being a scientist.

Figure 7.1 Students' interest in science

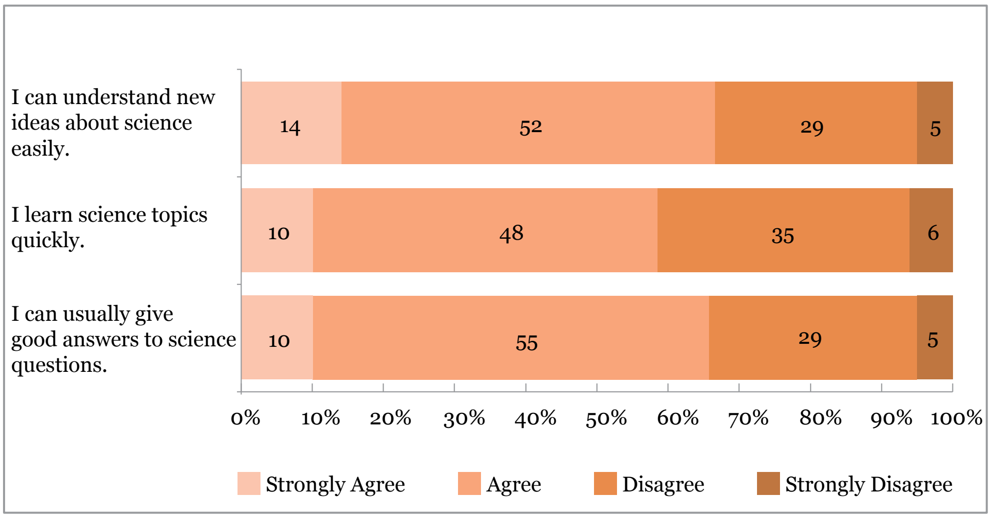


Note: Percentages may not add up to 100 per cent due to rounding.

Figure 7.1 demonstrates that very high percentages of student agreement were found for these statements. A range of 45 to 57 per cent was found for the 'agree' category and a range of 14 to 34 per cent was found for the 'strongly agree' category. While the great majority of students appear to be interested in learning about science and doing science, fewer students think it would be interesting to be a scientist. The percentage of students who agreed or strongly agreed with the statement 'I think it would be interesting to be a scientist' (59 per cent) in 2012 is lower than the corresponding percentage of students (67 per cent) in 2009. This difference is statistically significant. However, the percentage of students who agreed or strongly agreed with the statement 'I would like to learn more science at school' (79 per cent) in 2012 is higher than the corresponding percentage of students (74 per cent) in 2009. This difference is also statistically significant. Further information about the comparison of student responses in the 2009 and 2012 Student Surveys can be obtained in Appendix 5.

Self-concept in science. This section included three statements which sought to elicit whether students believe in their own science competencies.

Figure 7.2 Students' self-concept in science

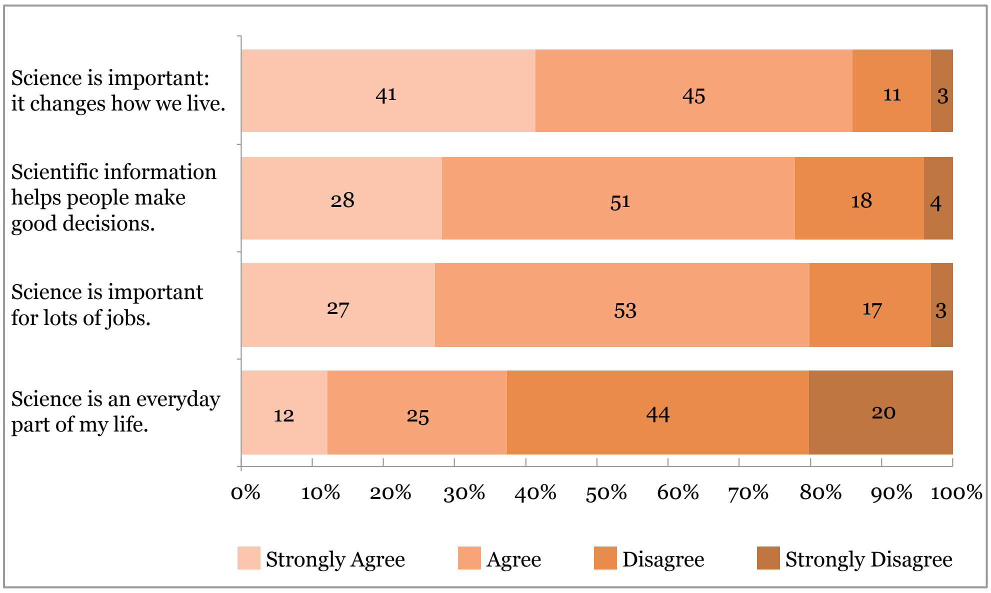


Note: Percentages may not add up to 100 per cent due to rounding.

Figure 7.2 demonstrates that high percentages of student agreement were found for these statements. A range of 48 to 55 per cent was found for the 'agree' category and a range of 10 to 14 per cent was found for the 'strongly agree' category. This indicates that a large proportion of students appear to be confident in learning science, reporting that they can understand and learn science ideas easily and quickly.

Value of science. This section included four statements which sought to elicit student perceptions of the importance of science to society and themselves.

Figure 7.3 Students' perceived value of science

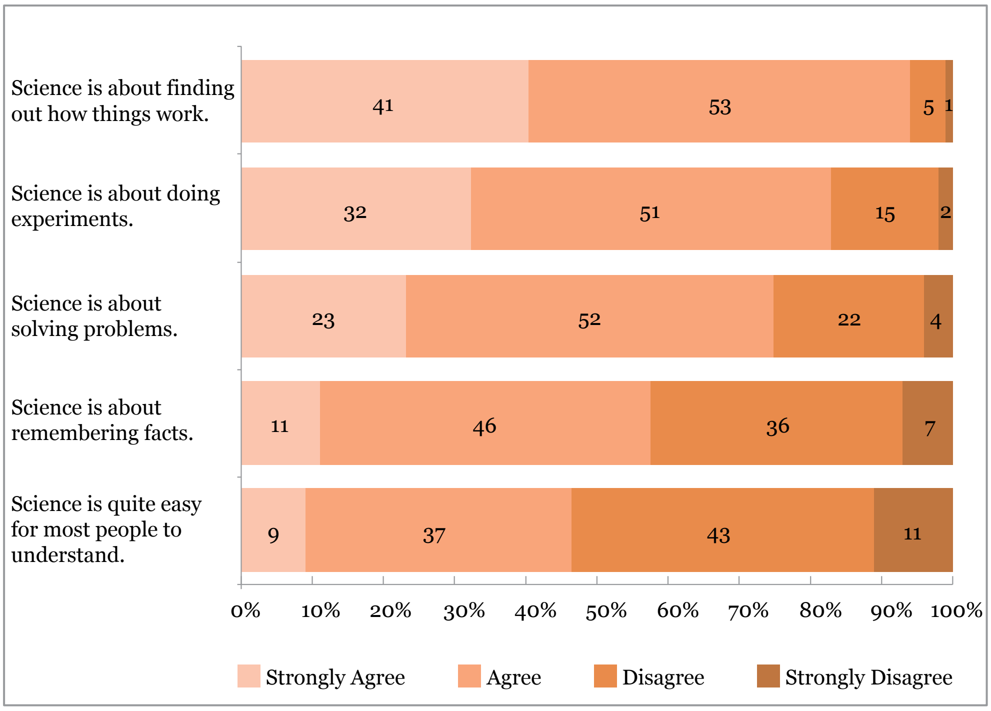


Note: Percentages may not add up to 100 per cent due to rounding.

Very high percentages of student agreement were found for the first three statements that relate to the importance of science to society. A range of 45 to 53 per cent was found for the 'agree' category and a range of 27 to 41 per cent was found for the 'strongly agree' category. However, the pattern of responses was not the same for the statement 'Science is an everyday part of my life'. This indicates that a large proportion of students show a general appreciation for science but do not necessarily relate this to their own lives. This is consistent with the findings of the 2009 NAP–SL survey. It should be noted, however, that the percentage of students who agreed or strongly agreed with the statement 'Science is an everyday part of my life' in 2009 (53 per cent) was higher than the corresponding percentage of students (37 per cent) in 2012. This difference is statistically significant.

Perceptions of science. This section included five statements which sought to elicit what students considered to be 'science' and whether science is easy for people to understand.

Figure 7.4 Students' perceptions of science

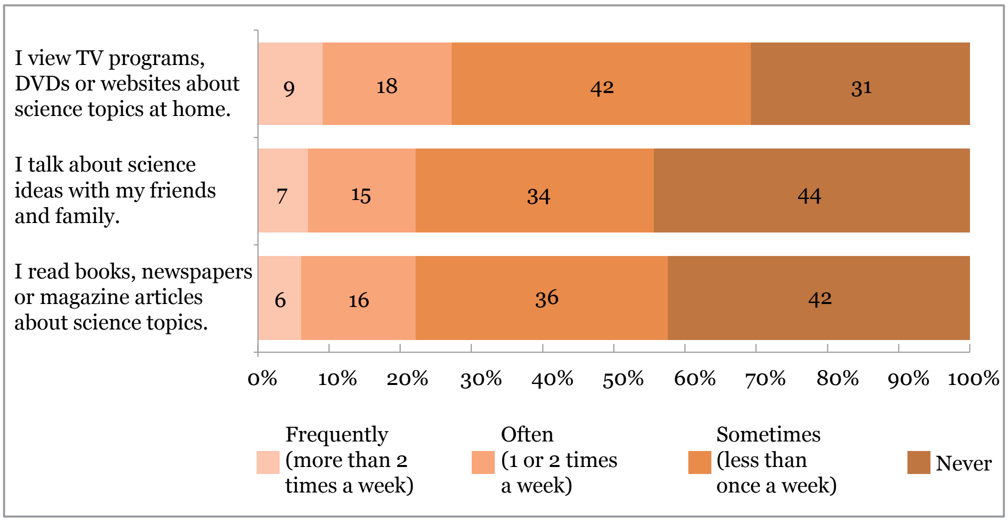


Note: Percentages may not add up to 100 per cent due to rounding.

Figure 7.4 indicates that most students agreed that science is about finding out how things work (94 per cent), doing experiments (83 per cent) and solving problems (75 per cent). Fewer students agreed that science is about remembering facts (57 per cent), a smaller percentage than in 2009 when 73 per cent of students agreed with this statement. This difference is statistically significant.

Science-related activities outside school. The three statements in this section of the Student Survey sought to gather information about the frequency with which students watch television programs or DVDs about science topics at home, read books and newspaper or magazine articles about science topics, or talk about science ideas with friends and family.

Figure 7.5 Science-related activities outside school



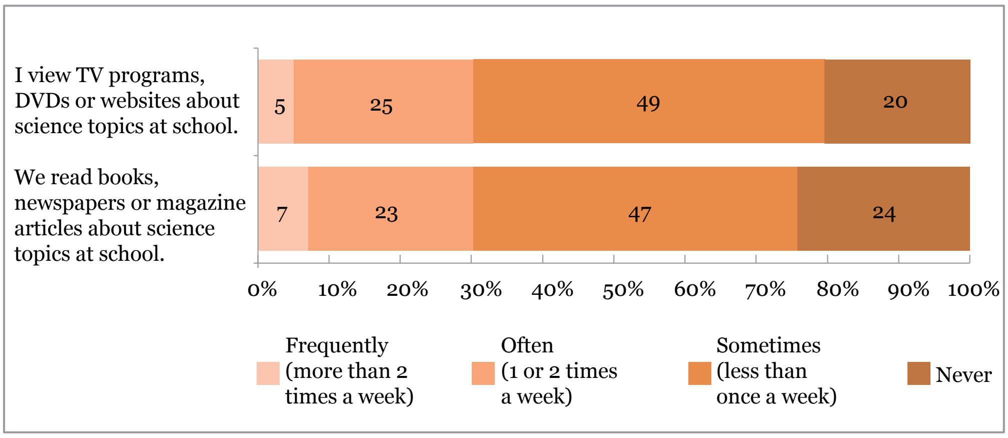
Note: Percentages may not add up to 100 per cent due to rounding.

Responses to this section indicate some of the ways that students learn about science beyond the school and reflect students' engagement with science in their personal lives. Figure 7.5 shows that 69 per cent of the students responded that they viewed science programs or science websites at home, with 27 per cent of the students doing so 'often' or 'frequently'.

In relation to a print medium, 58 per cent of the students indicated that they read about science topics, with 22 per cent doing so 'often' or 'frequently'. A similar percentage of students indicated that they talked about science ideas with their friends and family.

Science-related activities at school. This section sought to gather information about a) the frequency of use of audio-visual materials, website and print materials to enhance science teaching in the classroom; and b) student participation in science-related activities at school (e.g. school excursions and talks by visitors on science). The results of this section are shown in Figures 7.6a and 7.6b.

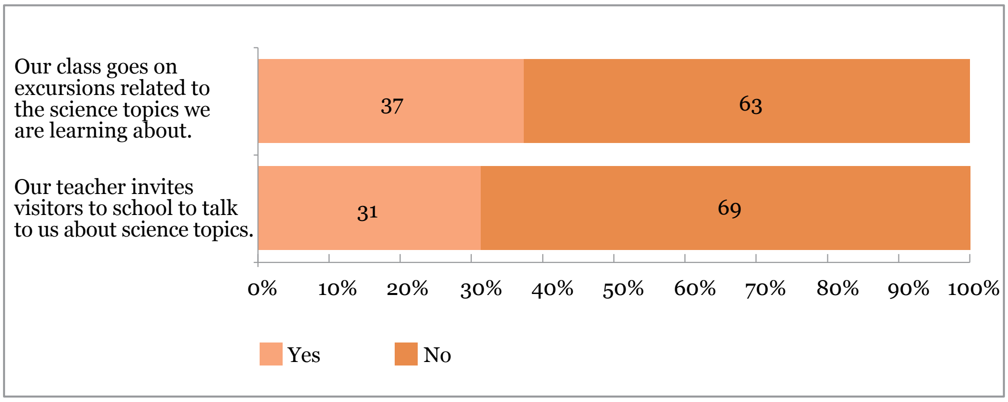
Figure 7.6a Science-related activities at school (materials)



Note: Percentages may not add up to 100 per cent due to rounding.

Figure 7.6a shows that students provided a broadly similar pattern of responses to the two questions about the use of materials in science teaching at school. Students provided similar responses to the questions about excursions and talks about science topics at school (Figure 7.6b).

Figure 7.6b Science-related activities at school (excursions and talks)



Note: Percentages may not add up to 100 per cent due to rounding.

Science teaching and investigations. This section sought to gather information about three aspects of students' experiences at school:

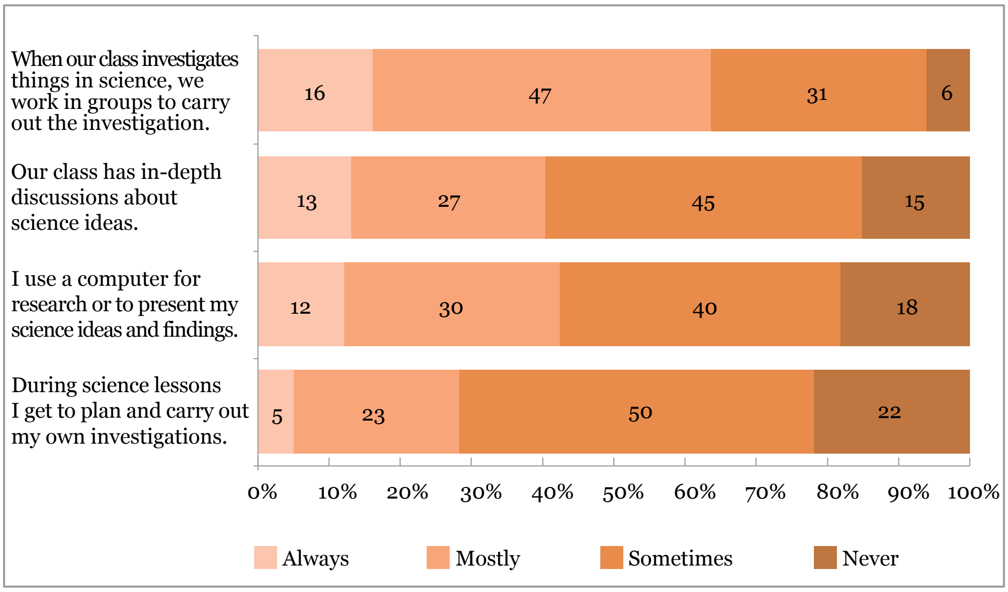
a. how science is taught in the classroom

b. their science teacher

c. the frequency of science lessons.

The results of this section of the Student Survey are shown in Figures 7.7a to 7.7c.

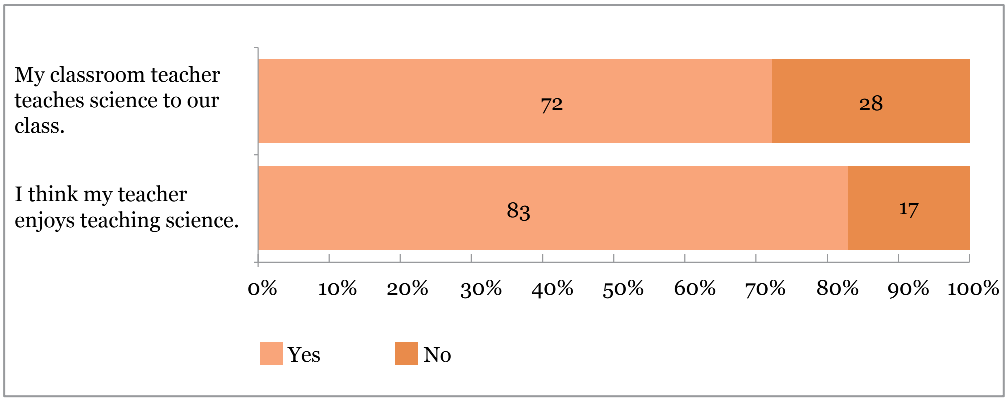
Figure 7.7a Science teaching and investigation (how science is taught in the classroom)



Note: Percentages may not add up to 100 per cent due to rounding.

Figure 7.7a shows that while 28 per cent of students reported that they 'always' or 'mostly' carried out their own self-directed investigations in science, 63 per cent reported that they 'always' or 'mostly' worked in groups to carry out investigations. These percentages are similar to the ones reported in the 2009 Student Survey. In the 2012 Student Survey, approximately 40 per cent of the students reported that they 'always' or 'mostly' had in-depth discussions about science ideas and that they used a computer in science.

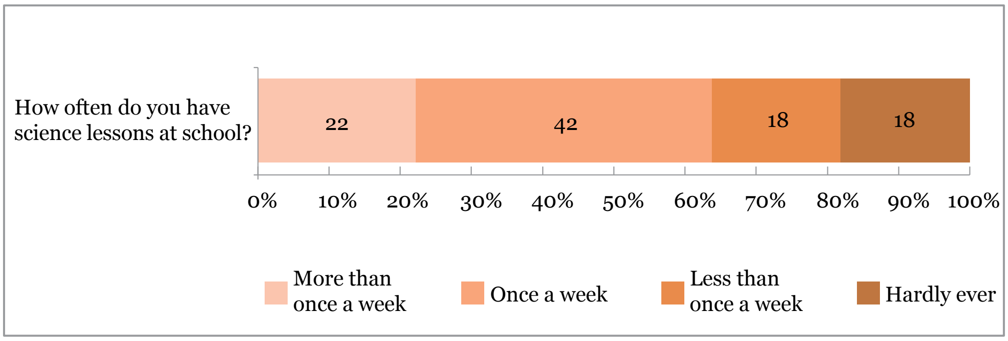
Figure 7.7b Science teaching and investigation (science teacher)



Note: Percentages may not add up to 100 per cent due to rounding.

Figure 7.7b shows that the regular classroom teacher is the person who teaches science for 72 per cent of the students and that 83 per cent of the students reported that they believe their teacher enjoys teaching science.

Figure 7.7c Science teaching and investigation (frequency of science lessons)



Note: Percentages may not add up to 100 per cent due to rounding.

Figure 7.7c indicates that the majority of students reported having at least one science lesson each week with 22 per cent indicating that they have a science lesson more than once a week.

Science topics studied at school. In this section, students were provided with prompts to assist them in recognising the broad science concept areas that may have been covered at school.

For the concept area Earth and Space, students were prompted with the following examples: 'weather, soil, rocks, gravity, using Earth's resources, the planets, Sun and Moon'.

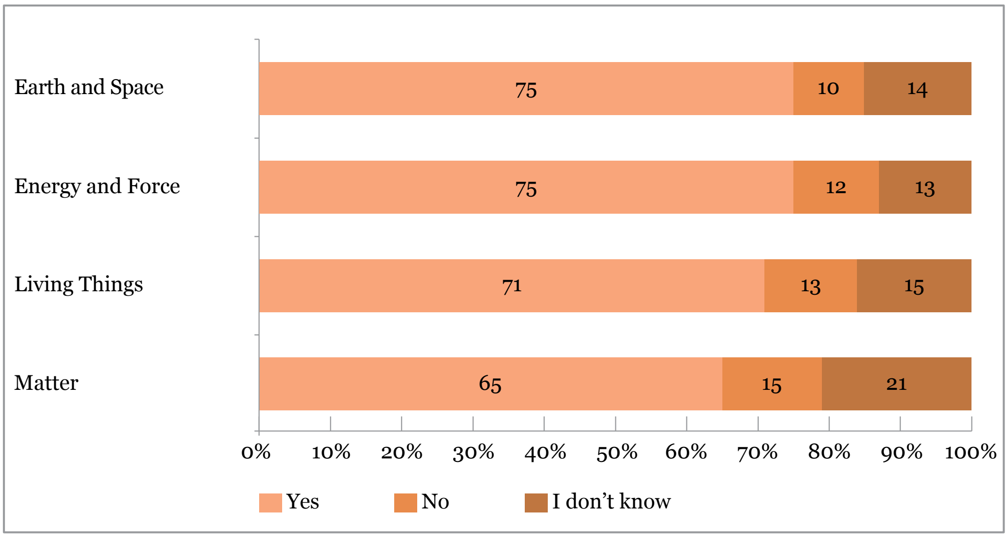
For the concept area Energy and Force, students were prompted with the following examples: 'how toys and other machines work, electricity, heat, light, sound, magnets'.

For the concept area Living Things, students received the following prompts: 'living and non-living things, how animals and plants survive in their environment, life cycles, interdependence'.

For the concept area Matter, the following prompts were supplied: 'the different properties of materials such as plastics and metals, the different uses of materials, changes to materials (solids, liquids and gases)'.

Examples were based on advice from the SLRC regarding the most commonly studied science concept areas in primary schools.

Figure 7.8 Science concept areas studied at school



Note: Percentages may not add up to 100 per cent due to rounding.

As Figure 7.8 shows, the most commonly studied science concept areas are Earth and Space, Energy and Force, and Living Things. Matter is somewhat less commonly studied.

Relationship between Student Survey responses and scientific literacy

To examine whether there are systematic patterns between students' responses to the Student Survey, the responses were subjected to factor analysis. Factor analysis is a statistical technique for identifying latent factors, or latent commonalities, by the intercorrelations between survey items.

The factor analysis model applied to analyse data collected from the survey included an orthogonal rotation, which results in the extraction of latent factors that do not correlate. Factor analysis showed that there are eight factors in the Student Survey that correspond broadly to the categories used to design the survey. These eight factors explain 51.1 per cent of the variability in student responses. Items that form the eight latent factors are presented in Table 7.1.

Table 7.1 Latent factors and items correlating with factors

Latent factors	Items correlating with latent factors	Percentage of explained variance	Cumulative percentage of explained variance
First	We read books, newspapers or magazine articles about science topics at school. I read books, newspapers or magazine articles about science topics. I view TV programs, DVDs or websites about science topics at home. I view TV programs, DVDs or websites about science topics at school. I talk about science ideas with my friends and family. I use a computer for research or to present my science ideas and findings.	7.8	7.8
Second	I enjoy doing science. I would like to learn more science at school. I enjoy learning new things in science. I think it would be interesting to be a scientist.	7.6	15.4
Third	Science is important: it changes how we live. Science is important for lots of jobs. Scientific information helps people make good decisions. Science is an everyday part of my life.	7.5	22.9
Fourth	I learn science topics quickly. I can understand new ideas about science easily. I can usually give good answers to science questions.	6.8	29.7
Fifth	I think my teacher enjoys teaching science. My classroom teacher teaches science to our class. How often do you have science lessons at school? Our class has in-depth discussions about science ideas. When our class investigates things in science, we work in groups to carry out the investigation. During science lessons I get to plan and carry out my own investigations.	6.1	35.8
Sixth	Science is about doing experiments. Science is about remembering facts. Science is finding out about how things work. Science is about solving problems. Science is quite easy for most people to understand.	6.0	41.8
Seventh	Living Things - for example, living and non-living things, how animals and plants survive in their environment, life cycles, interdependence. Earth and Space - for example, weather, soil, rocks, gravity, using Earth's resources, the planets, Sun and Moon. Matter - for example, the different properties of materials such as plastics and metals, the different uses of materials, changes to materials (solids, liquids and gases). Energy and Force - for example, how toys and other machines work, electricity, heat, light, sound, magnets.	5.2	47.0
Eighth	Our class goes on excursions related to the science topics we are learning about. Our teacher invites visitors to school to talk to us about science topics.	4.1	51.1

Items in Table 7.1 are listed in the order of magnitude of their correlation with the respective latent factor, while the latent factors are listed in order of their contribution to the amount of explained variance in the Student Survey.

It should be noted that the Student Survey was designed as a multifaceted instrument. Therefore, it is expected that no single factor will explain a substantial proportion of the variance in students' responses to the survey items.

In order to investigate the relationships between information collected in the Student Survey and student achievement in NAP–SL, a regression analysis was conducted. The regression analysis used items from the survey as independent variables, and student achievement, measured by plausible values, as dependent variables. In order to account for the stratified structure in the response data, the regression analysis was conducted using the students' sampling weights and the jackknife procedure was also used to calculate standard errors for the regression coefficients.

The regression analysis showed that only 23 per cent of the variability in students' scores in scientific literacy could be predicted based on their responses to the items in the survey. Furthermore, only three items demonstrated a correlation of meaningful magnitude, with a regression coefficient of absolute value above 0.2. These items were:

> Item 6: 'I can usually give good answers to science questions.' The regression analysis (b=0.26, SE=0.003) showed that an increase in scientific literacy achievement is associated with an increase in students' agreement with this statement.
>
> Item 16: 'Science is quite easy for most people to understand'. The regression analysis (b=-0.24, SE=0.002) showed that an increase in scientific literacy achievement is associated with a decrease in students' agreement with this statement.
>
> Item 5: 'I learn science topics quickly'. The regression analysis (b=0.23, SE=0.003) showed that an increase in scientific literacy achievement is associated with an increase in students' agreement with this statement.

Such a result has rendered uninformative any further use of regression analysis in explaining student achievement in NAP–SL. However, in order to provide an illustrative overview of the relationship between students' responses to survey items and their achievement in scientific literacy, the distribution of students across Proficiency Levels and response categories for a selected set of questions from the survey is provided below.

Items 6, 16 and 5 are included in this set because they showed a significant regression coefficient with students' scores in scientific literacy. Item 30 is also presented because it shows the relationship between student achievement in scientific literacy and the frequency of science teaching that students in the sample reported they typically receive. Tables 7.2 to 7.5 show the distribution of students' responses across Proficiency Levels and response categories for these items. (NB: Some cells contain no responses and return zero or small estimates.)

Table 7.2 Student Responses to Item 6 by Proficiency Levels

Item 6: I can usually give good answers to science questions.	Level 2 and below	Level 3.1	Level 3.2	Level 3.3	Level 4 and above
Strongly agree	6.0 (±1.2)	26.9 (±2.0)	47.4 (±2.2)	19.0 (±1.8)	0.8 (±0.4)
Agree	5.6 (±0.5)	35.3 (±1.0)	48.0 (±1.1)	10.8 (±0.8)	0.3 (±0.1)
Disagree	12.9 (±1.0)	49.1 (±1.4)	33.9 (±1.4)	4.1 (±0.5)	0.1 (±0.1)
Strongly disagree	20.8 (±2.5)	55.7 (±3.2)	22.1 (±2.8)	1.4 (±0.8)	0.0 (±0.0)
AUST	9.0 (±1.0)	39.6 (±1.6)	42.1 (±1.7)	9.0 (±1.1)	0.3 (±0.2)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

As can be seen in Table 7.2, the proportion of students who expressed agreement with the statement 'I can usually give good answers to science questions' and whose achievement was at or above the Proficient Standard (i.e. achieving at Level 3.2 and above) is higher than the proportion of students in the other two categories.

Table 7.3 Student Responses to Item 16 by Proficiency Levels

Item 16: Science is quite easy for most people to understand.	Level 2 and below	Level 3.1	Level 3.2	Level 3.3	Level 4 and above
Strongly agree	17.9 (±1.9)	48.1 (±2.2)	29.9 (±1.8)	4.1 (±0.8)	0.1 (±0.1)
Agree	9.1 (±0.8)	42.8 (±1.3)	40.4 (±1.3)	7.6 (±0.7)	0.2 (±0.1)
Disagree	5.7 (±0.6)	34.7 (±1.1)	47.5 (±1.3)	11.7 (±0.9)	0.4 (±0.2)
Strongly disagree	9.8 (±1.3)	40.2 (±2.4)	40.6 (±2.3)	9.2 (±1.2)	0.2 (±0.2)
AUST	9.0 (±1.0)	39.6 (±1.6)	42.1 (±1.7)	9.0 (±1.1)	0.3 (±0.2)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

As can be seen in Table 7.3, the proportion of students at or above the Proficient Standard was lower for students who agreed or strongly agreed with the statement 'Science is quite easy for most people to understand' compared to those who disagreed or strongly disagreed.

Table 7.4 Student Responses to Item 5 by Proficiency Levels

Item 5: I learn science topics quickly.	Level 2 and below	Level 3.1	Level 3.2	Level 3.3	Level 4 and above
Strongly agree	7.1 (±1.1)	26.5 (±1.9)	48.3 (±2.1)	17.4 (±1.6)	0.8 (±0.5)
Agree	5.4 (±0.6)	34.7 (±1.1)	47.6 (±1.2)	11.9 (±0.9)	0.4 (±0.1)
Disagree	11.6 (±0.9)	47.9 (±1.2)	36.3 (±1.3)	4.1 (±0.5)	0.1 (±0.0)
Strongly disagree	17.9 (±1.9)	52.8 (±2.4)	26.5 (±2.3)	2.8 (±0.9)	0.0 (±0.0)
AUST	9.0 (±1.0)	39.6 (±1.6)	42.1 (±1.7)	9.0 (±1.1)	0.3 (±0.2)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

As can be seen in Table 7.4, the proportion of students who expressed agreement with the statement 'I learn science topics quickly' and whose achievement was at or above the Proficient Standard (i.e. achieving at Level 3.2 and above) is higher than the proportion of students in the other two categories.

Table 7.5 Student Responses to Item 30 by Proficiency Levels

Item 30: How often do you have science lessons at school?	Level 2 and below	Level 3.1	Level 3.2	Level 3.3	Level 4 and above
More than once a week	7.9 (±0.9)	38.4 (±1.6)	43.6 (±1.5)	9.8 (±1.1)	0.3 (±0.2)
Once a week	7.8 (±0.7)	38.0 (±1.2)	43.6 (±1.2)	10.4 (±0.8)	0.3 (±0.1)
Less than once a week	8.3 (±1.1)	38.1 (±2.0)	43.7 (±1.6)	9.6 (±1.4)	0.3 (±0.2)
Hardly ever	11.2 (±1.6)	45.2 (±1.9)	37.6 (±2.0)	5.9 (±0.8)	0.2 (±0.1)
AUST	9.0 (±1.0)	39.6 (±1.6)	42.1 (±1.7)	9.0 (±1.1)	0.3 (±0.2)

Note: Figures in parentheses refer to 95 per cent confidence intervals.

Item 30 asked students to report on the frequency of science lessons in their classroom. Table 7.5 indicates that the reported frequency of science lessons has only a small impact on student achievement in scientific literacy as measured by NAP–SL. Students who reported 'hardly ever' having a science lesson, had slightly lower achievements than the other students. Results should be interpreted with caution as students may not always be aware that what they are learning is 'science' particularly when much primary teaching is done in integrated cross-curriculum units.

Conclusion

The Student Survey provided interesting insights into Australian students' attitudes toward science. Items from the category 'Self-concept in Science' have the highest correlations with test performance. That is, a higher science self-concept is associated with a higher mark in scientific literacy. However, in examining the relationship between students' attitudes and test performance, it is difficult to draw many reliable conclusions from the survey data collected, as the relationship between survey items to achievement was weak to moderate. This could be explained by several factors. It may be that some students, regardless of their level of achievement, provided answers that they thought would please their teachers. It may also be that some students did not fully understand what was being asked by each survey item and therefore chose a response at random or in a set fashion. Such a result does indicate the need for more detailed investigation into the level of student engagement with questionnaires accompanying the main assessments in National Assessment Programs.

Chapter 8
Conclusion

The 2012 cycle of the National Assessment Program – Science Literacy (NAP–SL) provided the opportunity to report on the progress of Year 6 students in scientific literacy over a six-year period using the same assessment framework as well as sampling and test design methodologies. One of the main objectives of NAP–SL is to measure trends over time. To this end the scientific literacy scale was initially established in 2003. However in 2006, a more robust test design was implemented, resulting in a sample that was more inclusive of remote schools and the items providing better discrimination of students. Consequently, the 2006 results were utilised to establish a new baseline scientific literacy scale.

The Assessment Domain for scientific literacy and science concept areas have remained stable since the 2006 assessment cycle. The number of test booklets and the number of assessment items was increased in the 2006 assessment. In 2003, students were assessed using two test booklets. In 2006 and the following cycles, seven test booklets were implemented allowing for a cluster rotation design to be used. This allowed for clusters of items to be presented to students at varying points of the test booklet, thereby minimising any effect on performance due to an item's position within a test booklet.

Student achievement in scientific literacy from 2006 to 2012

In 2012, 51.4 per cent of students at the national level attained the Proficient Standard or better in scientific literacy. In 2009, the percentage was 51.9 per cent and in 2006 it was 54.3 per cent. The differences in the percentage of students attaining the Proficient Standard in the three assessment cycles are not statistically significant. Chapter 3 of this report contains detailed information about the performance of students nationally and at state or territory level in NAP–SL. Student results are reported against five Proficiency Levels (Level 2, Level 3.1, Level 3.2, Level 3.3 and Level 4) with Level 3.2 being described as the Proficient Standard. The distribution of students across the Proficiency Levels at the national level has remained relatively stable across the three cycles.

Factors associated with achieving scientific literacy

As outlined in Chapter 6 of this report, student background characteristics are related to achievement of scientific literacy. Background data were collected related to gender, Indigenous status, language background and geographical location.

At the national level, girls slightly outperformed boys, although the difference was not statistically significant. In the two previous cycles, boys slightly outperformed girls but the differences were not statistically significant.

Nationally, non-Indigenous students achieved significantly higher levels of scientific literacy than Indigenous students, as was also the case in 2009 and 2006. This finding is similar to that of other National Assessment Programs and indicates that strategies need to be found to address the gap in achievement between the two groups.

Students living in metropolitan areas achieved the highest mean score in scientific literacy. Their results were significantly different from those living in provincial areas and those living in remote and very remote areas. Similar findings are evident in the National Assessment Programs in Literacy, Numeracy, Civics and Citizenship, and Information and Communication Technology Literacy.

Students from English-speaking backgrounds also achieved slightly higher means nationally than students from language backgrounds other than English. However, the difference was not statistically significant. A similar trend was observed in the 2009 assessment.

Student Survey

As discussed in Chapter 7, a Student Survey was administered as part of the 2012 assessment. Analyses found only weak to moderate correlations between student performance in scientific literacy and particular survey responses. However, the survey provided interesting insights into students' perceptions of and attitudes to science and their experiences with science learning at and outside school. Approximately 80 per cent of students responded that they would like to learn more science at school indicating that a positive attitude towards this subject area exists. The responses to the Student Survey will guide further survey development for future cycles and can provide impetus for discussion at school and jurisdictional levels regarding students' perceptions of, attitudes towards and experiences of science in their lives.

References

Australian Government Department of Education, Science and Training (2004) Science Education Assessment Resource <http://cms.curriculum.edu.au/sear/>

Ball, S. et al. (2000) Options for the assessment and reporting of primary students in the key learning area of science to be used for the reporting of nationally comparable outcomes of schooling within the context of the National Goals for Schooling in the Twenty-First Century: National Education Performance Monitoring Taskforce.

Biggs, J.B. & Collis, K.F. (1982) Evaluating the quality of learning: The SOLO taxonomy. New York: Academic Press.

Goodrum, D., Hackling, M.W. & Rennie, L. (2001) The status and quality of teaching and learning of science in Australian schools. Canberra: Department of Education, Training and Youth Affairs.

Jones, R. (2004) Geolocation Questions and Coding Index: A technical report submitted to the Performance Measurement and Reporting Taskforce of the Ministerial Council on Education, Employment, Training and Youth Affairs. MCEETYA.

Laugksch, R.C. (2000) Scientific literacy: A conceptual overview. *Science Education*, 84(1), 71–94.

Lokan, J.L., Greenwood, L. & Cresswell, J. (2001) The PISA 2000 survey of students' reading, mathematical and scientific literacy skills. Melbourne: ACER.

OECD Programme for International Student Assessment (1999) Measuring student knowledge and skills: A new framework for assessment. Paris: OECD.

Appendix 1
National Year 6 Primary Science Assessment Domain

Assessment strands: Scientific literacy

The national review of the status and quality of teaching and learning of science in Australian schools (Goodrum, Hackling & Rennie 2001) argued that the broad purpose of science in the compulsory years of schooling is to develop scientific literacy for all students.

Scientific literacy is a high priority for all citizens, helping them to:

- be interested in and understand the world around them
- engage in discourses of and about science
- be sceptical and questioning of claims made by others about scientific matters
- be able to identify questions, investigate and draw evidence-based conclusions
- make informed decisions about the environment and their own health and wellbeing.

Scientific literacy is important because it contributes to the economic and social wellbeing of the nation and improved decision-making at public and personal levels (Laugksch 2000).

The Programme for International Student Assessment (PISA) focuses on aspects of preparedness for adult life in terms of functional knowledge and skills that allow citizens to participate actively in society. It is argued that scientifically literate people are 'able to use scientific knowledge and processes not just to understand the natural world but also to participate in decisions that affect it' (OECD 1999, p. 13).

The OECD–PISA defined scientific literacy as:

> *... the capacity to use scientific knowledge, to identify questions (investigate)*[1] *and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity.*

(OECD 1999, p. 60)

This definition has been adopted for the National Assessment Program – Science Literacy (NAP–SL) in accordance with the Ball et al. 2000 report recommendation.

1 Because of the constraints of large-scale testing, PISA was not able to include performance tasks such as conducting investigations. Consequently, its definition of scientific literacy omitted reference to investigating. The word 'investigate' was inserted into the definition for the purposes of NAP–SL, as the sample testing methodology allowed for assessments of students' ability to conduct investigations.

Scientific literacy: Progress Map

A scientific literacy Progress Map was developed based on the construct of scientific literacy and an analysis of state and territory curriculum and assessment frameworks. The Progress Map describes the development of scientific literacy across three strands of knowledge which are inclusive of Ball et al.'s concepts and processes and the elements of the OECD–PISA definition.

The five elements of scientific literacy, including concepts and processes used in PISA 2000 (OECD 1999), include:

1. demonstrating understanding of scientific concepts
2. recognising scientifically investigable questions
3. identifying evidence needed in a scientific investigation
4. drawing or evaluating conclusions
5. communicating valid conclusions.

These elements have been clustered into three more holistic strands which are described below. The second and third elements and conducting investigations to collect data are encompassed in Strand A; the fourth and fifth elements and conducting investigations to collect and interpret data are included in Strand B; and the first element is included in Strand C.

Strand A: Formulating or identifying investigable questions and hypotheses, planning investigations and collecting evidence.

This process strand includes posing questions or hypotheses for investigation or recognising scientifically investigable questions; planning investigations by identifying variables and devising procedures where variables are controlled; gathering evidence through measurement and observation; and making records of data in the form of descriptions, drawings, tables and graphs using a range of information and communication technologies.

Strand B: Interpreting evidence and drawing conclusions from students' own or others' data, critiquing the trustworthiness of evidence and claims made by others, and communicating findings.

This process strand includes identifying, describing and explaining the patterns and relationships between variables in scientific data; drawing conclusions that are evidence-based and related to the questions or hypotheses posed; critiquing the trustworthiness of evidence and claims made by others; and communicating findings using a range of scientific genres and information and communications technologies.

Strand C: Using science understandings for describing and explaining natural phenomena, and for interpreting reports about phenomena.

This conceptual strand includes demonstrating conceptual understandings by being able to describe, explain and make sense of natural phenomena; understand and interpret reports (e.g. TV documentaries, newspaper or magazine articles or conversations) related to scientific matters; and make decisions about scientific matters in students' own lives which may involve some consideration of social, environmental and economic costs and benefits.

Scientific literacy has been described here in three strands to facilitate the interpretation of student responses to assessment tasks. However, authentic tasks should require students to apply concepts and processes together to address problems set in real-world contexts. These tasks may involve ethical decision-making about scientific matters in students' own lives and some consideration of social, environmental and economic costs and benefits.

The scientific literacy Progress Map (see Table A1.1) describes progression in six levels from 1 to 6 in terms of three aspects:

- increasing complexity, from explanations that involve one aspect to several aspects, through to relationships between aspects of a phenomenon
- progression from explanations that refer to and are limited to directly experienced phenomena (concrete) to explanations that go beyond what can be observed directly and involve abstract scientific concepts (abstract)
- progression from descriptions of 'what' happened in terms of objects and events, to explanations of 'how' it happened in terms of processes, to explanations of 'why' it happened in terms of science concepts.

Strand C has been abstracted and makes no reference to particular science concepts or contexts. As the progression in this strand is based on increasing complexity and abstraction, links have been made to the Structure of Observed Learning Outcomes (SOLO) taxonomy (Biggs & Collis 1982).

The taxonomy was written to describe levels of student responses to assessment tasks. The basic SOLO categories include:

prestructural	no logical response
unistructural	refers to only one aspect
multistructural	refers to several independent aspects
relational	can generalise (describe relationships between aspects) within the given or experienced context
extended abstract	can generalise to situations not experienced.

The three main categories of unistructural, multistructural and relational can also be applied, as cycles of learning, to the four modes of representation:

sensorimotor	the world is understood and represented through motor activity
iconic	the world is represented as internal images
concrete	writing and other symbols are used to represent and describe the experienced world
formal	the world is represented and explained using abstract conceptual systems.

The conceptual strand, Strand C, of the Progress Map therefore makes links to the SOLO categories of concrete unistructural (level 1), concrete multistructural (level 2), concrete relational (level 3), abstract unistructural (level 4), abstract multistructural (level 5) and abstract relational (level 6).

The SOLO levels of performance should not be confused with Piagetian stages of cognitive development. Biggs and Collis (1982, p. 22) explain that the relationship between Piagetian stages and SOLO levels 'is exactly analogous to that between ability and attainment' and that level of performance depends on quality of instruction, motivation to perform, prior knowledge and familiarity with the context. Consequently, performance for a given individual is highly variable and often sub-optimal.

NAP–SL focuses on levels 2, 3 and 4 of the scientific literacy Progress Map, the levels of scientific literacy attained by students in Year 6.

The agreed Proficiency Levels serve to further elaborate the Progress Map. Level 3 is described as 3.1, 3.2, and 3.3. A Proficient Standard is a challenging level of performance, with students needing to demonstrate more than minimal or elementary skills.

Table A1.1 Scientific Literacy Progress Map

Level	Strands of scientific literacy		
	Strand A Formulating or identifying investigable questions and hypotheses, planning investigations and collecting evidence. Process strand: experimental design and data gathering.	**Strand B** Interpreting evidence and drawing conclusions from students' own or others' data, critiquing the trustworthiness of evidence and claims made by others, and communicating findings. Process strand: interpreting experimental data.	**Strand C** Using understandings for describing and explaining natural phenomena, and for interpreting reports about phenomena. Conceptual strand: applies conceptual understanding.
6	Uses scientific knowledge to formulate questions, hypotheses and predictions and to identify the variables to be changed, measured and controlled. Trials and modifies techniques to enhance reliability of data collection.	Selects graph type and scales that display the data effectively. Conclusions are consistent with the data, explain the patterns and relationships in terms of scientific concepts and principles, and relate to the question, hypothesis or prediction. Critiques the trustworthiness of reported data (e.g. adequate control of variables, sample or consistency of measurements, assumptions made in formulating the methodology), and consistency between data and claims.	Explains complex interactions, systems or relationships using several abstract scientific concepts or principles and the relationships between them. SOLO taxonomy: Abstract relational
5	Formulates scientific questions or hypotheses for testing and plans experiments in which most variables are controlled. Selects equipment that is appropriate and trials measurement procedure to improve techniques and ensure safety. When provided with an experimental design involving multiple independent variables, can identify the questions being investigated.	Conclusions explain the patterns in the data using science concepts, and are consistent with the data. Makes specific suggestions for improving/extending the existing methodology (e.g. controlling an additional variable, changing an aspect of measurement technique). Interprets/compares data from two or more sources. Critiques reports of investigations noting any major flaw in design or inconsistencies in data.	Explains phenomena, or interprets reports about phenomena, using several abstract scientific concepts. SOLO taxonomy: Abstract multistructural
4	Formulates scientific questions, identifies the variable to be changed, the variable to be measured and in addition identifies at least one variable to be controlled. Uses repeated trials or replicates. Collects and records data involving two or more variables.	Calculates averages from repeat trials or replicates, plots line graphs where appropriate. Interprets data from line graph or bar graph. Conclusions summarise and explain the patterns in the science data. Able to make general suggestions for improving an investigation (e.g. make more measurements).	Explains interactions, processes or effects that have been experienced or reported, in terms of a non-observable property or abstract science concept. SOLO taxonomy: Abstract unistructural

Level	Strands of scientific literacy		
	Strand A Formulating or identifying investigable questions and hypotheses, planning investigations and collecting evidence. Process strand: experimental design and data gathering.	**Strand B** Interpreting evidence and drawing conclusions from students' own or others' data, critiquing the trustworthiness of evidence and claims made by others, and communicating findings. Process strand: interpreting experimental data.	**Strand C** Using understandings for describing and explaining natural phenomena, and for interpreting reports about phenomena. Conceptual strand: applies conceptual understanding.
3	Formulates simple scientific questions for testing and makes predictions. Demonstrates awareness of the need for fair testing and appreciates scientific meaning of 'fair testing'. Identifies variable to be changed and/or measured but does not indicate variables to be controlled. Makes simple standard measurements. Records data as tables, diagrams or descriptions.	Displays data as tables or constructs bar graphs when given the variables for each axis. Identifies and summarises patterns in science data in the form of a rule. Recognises the need for improvement to the method. Applies the rule by extrapolating and predicting.	Describes the relationships between individual events (including cause and effect relationships) that have been experienced or reported. Can generalise and apply the rule by predicting future events. SOLO taxonomy: Concrete relational
2	Given a question in a familiar context, identifies that one variable/factor is to be changed (but does not necessarily use the term 'variable' to describe the changed variable). Demonstrates intuitive level of awareness of fair testing. Observes and describes or makes non-standard measurements and limited records of data.	Makes comparisons between objects or events observed. Compares aspects of data in a simple supplied table of results. Can complete simple tables and bar graphs given table column headings or prepared graph axes.	Describes changes to, differences between or properties of objects or events that have been experienced or reported. SOLO taxonomy: Concrete multistructural
1	Responds to the teacher's questions and suggestions, manipulates materials and observes what happens.	Shares observations; tells, acts out or draws what happened. Focuses on one aspect of the data.	Describes (or recognises) one aspect or property of an individual object or event that has been experienced or reported. SOLO taxonomy: Concrete unistructural

Major scientific concepts in NAP–SL

A table of the major scientific concepts found most widely in the various state and territory curriculum documents has been developed to accompany the scientific literacy Progress Map (see Table A1.2).

These major concepts are broad statements of scientific understandings that Year 6 students would be expected to demonstrate. They provided item writers with a specific context in which to assess scientific literacy. An illustrative list of examples for each of the major concepts provides elaboration of these broad conceptual statements and, in conjunction with the scientific literacy Progress Map which describes the typical developmental stages for scientific literacy, was used as a guide for the development of assessment items.

It should be noted that, because the NAP–SL test instruments are constructed within the constraints of test length, it is not feasible to include all the listed concepts in instruments constructed for a single testing cycle.

Table A1.2 Major scientific concepts in NAP–SL

Major scientific concepts	Examples
Earth and Space	
Earth, sky and people: Our lives depend on air, water and materials from the ground; the ways we live depend on landscape, weather and climate.	Features of weather, soil and sky, and effects on me. People use resources from the Earth; need to use them wisely. Sustainability.
The changing Earth: The Earth is composed of materials that are altered by forces within and upon its surface.	Changes in weather, weather data, seasons, soil landscape and sky (e.g. Moon phases, weathering and erosion, movement of the Sun and shadows, bush fires, land clearing). Climate change.
Our place in space: The Earth and life on Earth are part of an immense system called the universe.	Rotation of the Earth and night/day, spatial relationships between Sun, Earth and Moon. Planets of our solar system and their characteristics. Space exploration and new developments.
Energy and Force	
Energy and us: Energy is vital to our existence and our quality of life as individuals and as a society.	Uses of energy, patterns of energy use and variations with time of day and season. Energy sources, renewable and non-renewable.
Transferring energy: Interaction and change involve energy transfers; control of energy transfer enables particular changes to be achieved.	Sources, transfers, carriers and receivers of energy, energy and change. Types of energy, energy of motion – toys and other simple machines – light, sound.
Energy sources and receivers: Observed change in an object or system is indicated by the form and amount of energy transferred to or from it.	Forces as pushes and pulls, magnetic attraction and repulsion.
Living Things	
Living together: Organisms in a particular environment are interdependent.	Living vs non-living. Plant vs animal and major groups. Dependence on the environment: Survival needs – food, space and shelter. Interactions between organisms and interdependence (e.g. simple food chains).
Structure and function: Living things can be understood in terms of functional units and systems.	Major structures and systems and their functions. Healthy lifestyle, diet and exercise.
Biodiversity, change and continuity: Life on Earth has a history of change and disruption, yet continues generation to generation.	Change over lifetime, reproduction and lifecycles. Adaptation to physical environment.
Matter	
Materials and their uses: The properties of materials determine their uses; properties can be modified.	Materials have different properties and uses. Processing materials to make useful things produces waste, use of alternative materials to better care for the environment. Waste reduction – recycling. Nanotechnology.
Structure and properties: The substructure of materials determines their behaviour and properties.	The properties of materials can be explained in terms of their visible substructure, such as fibres. Materials can change their state and properties.
Reactions and change: Patterns of interaction of materials enable us to understand and control those interactions.	Solids, liquids and gases.

Appendix 2
Sampling

Sampling results

The target population for NAP–SL consisted of all students enrolled in Year 6 in Australian schools in 2012.

The nationwide sample aimed to be approximately 12 000 students located within approximately 600 schools throughout Australia. The 2012 sample design was closely aligned to those of the 2006 and 2009 assessments.

Target sample sizes across the jurisdictions were determined so that the precisions of estimates were as similar across jurisdictions as possible.

The sample design for NAP–SL was a two-stage stratified cluster sample. Stratification involves ordering and grouping schools according to state, sector, size and school location. This helps ensure adequate coverage of all desired school types in the sample.

Stage 1 consisted of selecting schools that had Year 6 students. In this stage, schools were selected with probabilities proportional to the estimated Year 6 enrolments. Within this process the list of schools was explicitly stratified by state, sector and school size.

Stage 2 involved the random selection of an intact Year 6 class from the sampled schools selected in Stage 1.

No school-level exclusions from the supplied sampling frame were made prior to sample selection.

Table A2.1 shows the number of educational institutions and students in the sampling frame for each jurisdiction. In this and the following tables percentages have been rounded and may not add up to 100.

Table A2.1 Estimated 2012 Year 6 enrolment figures as provided by Department of Education, Employment and Workplace Relations (DEEWR)

State/ Territory	Institutions	Students	Percentage of students
ACT	97	4503	1.6
NSW	2364	87709	32.1
NT	158	3198	1.2
QLD	1392	59231	21.7
SA	605	19326	7.1
TAS	221	6643	2.4
VIC	1766	62916	23.0
WA	883	29486	10.8
AUST	7486	273012	100.0

Table A2.2 shows the proportions of large, moderately small and very small schools within each jurisdiction. Schools with Year 6 enrolment sizes larger than or equal to the Target Cluster Size (25) were classified as large schools. Those with enrolment sizes smaller than the Target Cluster Size (TCS) but larger than 12 (TCS/2) were classified as moderately small schools. Schools with an enrolment of 12 (TCS/2) or less were classified as very small. It can be seen that there are many small schools in each jurisdiction. It was important that an appropriate strategy was utilised to prevent an over-selection of small schools, resulting in a sample size smaller than the desired target sample size.

Table A2.2 Proportions of schools by school size and jurisdiction

State/ Territory	School size	No. Schools	Percentage of schools	No. Students	Percentage of students
ACT	Large	77	79.4	4190	93.0
	Moderately small	14	14.4	268	6.0
	Very small	6	6.2	45	1.0
	Total	97	100.0	4503	100.0
NSW	Large	1386	58.6	77362	88.2
	Moderately small	379	16.0	7001	8.0
	Very small	599	25.3	3346	3.8
	Total	2364	100.0	87709	100.0
NT	Large	55	34.8	2355	73.6
	Moderately small	26	16.5	454	14.2
	Very small	77	48.7	389	12.2
	Total	158	100.0	3198	100.0
QLD	Large	799	57.4	53691	90.6
	Moderately small	171	12.3	3161	5.3
	Very small	422	30.3	2379	4.0
	Total	1392	100.0	59231	100.0
SA	Large	324	53.6	15923	82.4
	Moderately small	132	21.8	2464	12.7
	Very small	149	24.6	939	4.9
	Total	605	100.0	19326	100.0
TAS	Large	118	53.4	5453	82.1
	Moderately small	40	18.1	747	11.2
	Very small	63	28.5	443	6.7
	Total	221	100.0	6643	100.0
VIC	Large	1021	57.8	54229	86.2
	Moderately small	328	18.6	6131	9.7
	Very small	417	23.6	2556	4.1
	Total	1766	100.0	62916	100.0
WA	Large	518	58.7	25847	87.7
	Moderately small	119	13.5	2163	7.3
	Very small	246	27.9	1476	5.0
	Total	883	100.0	29486	100.0

Class selection

One class containing Year 6 students was sampled per school. In some schools where there were several Year 6 classes, each with a small number of Year 6 students, the classes were combined to create a pseudo-class, where possible. Classes generally had equal probabilities of selection. The overall procedure for class selection was as follows:

1. Each class in a school was assigned a random number.
2. The classes in a school were ordered by the assigned random numbers.
3. The first class on each school's ordered list was chosen for the sample.

More detail about the sampling process may be found in the 2012 Technical Report.

Sample achieved

The NAP–SL specifications set the target sample size at 12 000 students. The total achieved sample size for 2012 was 13 236.

Table A2.3 School participation rates by jurisdiction

State/ Territory	School population	Number of schools sampled	Number of excluded and non-participating schools	Number of schools that participated	School participation (per cent)
ACT	97	54	0	54	100.0
NSW	2364	92	2	90	97.8
NT	158	50	9	41	82.0
QLD	1392	92	0	92	100.0
SA	605	94	0	94	100.0
TAS	221	64	0	64	100.0
VIC	1766	93	3	90	96.8
WA	883	94	2	92	97.9
AUST	7486	633	16	617	97.5

In total, three schools were excluded prior to the test date. These schools gave various reasons for not participating. Another five schools refused to participate. A further eight schools with a low participation rate were removed from the final sample. From Table A2.3 it can be seen that the participation rate for Northern Territory schools was lower than that for other jurisdictions. From the original target sample of 50 schools, nine were excluded from the final sample for various reasons. Two schools were exempted, two schools were deemed ineligible and five schools had insufficient eligible students present on test day.

More detail about the achieved sample may be found in the 2012 Technical Report.

Table A2.4 provides a breakdown of the sample according to jurisdiction. The target sample is the number of Year 6 students enrolled in the sampled classes at the time of testing. The achieved sample is the number of Year 6 students who participated.

Table A2.4 NAP–SL target and achieved sample sizes by jurisdiction

State/ Territory	Target sample		Achieved sample	
	Students	Percentage of students	Students	Percentage of students
ACT	1305	8.9	1242	9.4
NSW	2246	15.3	2060	15.6
NT	959	6.5	710	5.4
QLD	2207	15.0	2052	15.5
SA	2082	14.2	1926	14.6
TAS	1420	9.7	1259	9.5
VIC	2112	14.4	1854	14.0
WA	2344	16.0	2133	16.1
AUST	14675	100.0	13236	100.0

Table A2.5 provides a breakdown of the achieved sample in comparison with the number of Year 6 students in each jurisdiction.

Table A2.5 Achieved sample by student participation

State/ Territory	Student population	Number of students in sampled classes	Number of students who participated	Within-school exclusions	Within-school exclusions (per cent)	Within-school student participation (per cent)
ACT	4503	1305	1242	9	0.7	95.2
NSW	87709	2246	2060	5	0.2	91.7
NT	3198	959	710	22	3.1	74.0
QLD	59231	2207	2052	31	1.5	93.0
SA	19326	2082	1926	15	0.8	92.5
TAS	6643	1420	1259	18	1.4	88.7
VIC	62916	2112	1854	43	2.3	87.8
WA	29486	2344	2133	21	1.0	91.0
AUST	273012	14675	13236	164	1.2	90.2

Sample characteristics

Table A2.6 provides a breakdown of the achieved sample across states and territories according to gender, Indigenous status, students' language background and geographic location.

Table A2.6 Percentage distribution of Year 6 sample characteristics by jurisdiction

	State/Territory (per cent)								AUST (per cent)
	ACT	NSW	NT	QLD	SA	TAS	VIC	WA	
Student gender									
Male	49.8	51.1	46.5	50.0	50.4	50.0	52.3	51.8	50.6
Female	50.2	48.9	53.2	50.0	49.6	50.0	47.7	48.2	49.4
Missing	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0
Indigenous status									
Non-Indigenous	97.5	95.6	73.8	93.1	86.4	88.4	88.9	89.8	90.3
Indigenous	2.3	3.8	24.6	4.7	2.8	6.8	1.0	5.7	5.0
Missing	0.2	0.6	1.5	2.2	10.8	4.8	10.1	4.5	4.7
Language background									
English speaking background	82.1	67.4	60.8	86.0	83.9	92.9	77.5	74.6	78.7
Language background other than English	17.5	28.5	21.3	7.2	10.3	4.8	16.8	15.6	15.2
Missing	0.4	4.0	17.9	6.8	5.8	2.3	5.8	9.8	6.1
Geographic location									
Metropolitan areas	99.4	75.4	0.0	73.7	75.6	42.6	71.3	71.9	69.1
Provincial areas	0.6	24.4	66.9	25.0	20.4	56.3	28.7	21.3	27.1
Remote and very remote areas	0.0	0.1	33.1	1.3	4.0	1.1	0.0	6.8	3.8
Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of students	1242	2060	710	2052	1926	1259	1854	2133	13236

Data has not been presented for parental education and occupation as there was substantial missing data across each state and territory.

Table A2.7 provides a breakdown of the number of students in the achieved sample by Indigenous status across the three geographic location categories.

Table A2.7 Achieved sample size by Indigenous status and geographic location

Geographic location	Number of students by Indigenous status			Total
	Indigenous	Non-Indigenous	Missing	
Metropolitan areas	226	8428	494	9148
Provincial areas	284	3188	117	3589
Remote and very remote areas	147	339	13	499
Total	657	11955	624	13236

Comparisons of mean scores for Indigenous and non-Indigenous students by geographic location have not been provided in this report. The relatively small sample size for Indigenous students and the amount of missing data will result in artificially inflated estimates of the measurement error thus rendering comparisons of mean scores unsound.

School-level student exclusions

Within-school exclusions may have occurred for the following reasons:

Table A2.8 Within-school exclusion categories

Functional Disability	Student has a moderate to severe permanent physical disability such that he/she cannot perform in the NAP–SL testing situation. Functionally disabled students who could respond to the assessment were included.
Intellectual Disability	Student has a mental or emotional disability and is cognitively delayed such that he/she cannot perform in the NAP–SL testing situation. This includes students who are emotionally or mentally unable to follow even the general instructions of the assessment. Students were NOT excluded solely because of poor academic performance or disciplinary problems.
Limited Language Proficiency	The student is unable to read or speak the language of the assessment (i.e. English) and would be unable to overcome the language barrier in the testing situation. Typically a student who has received less than one year of instruction in the language of the assessment may be excluded.
Refusal	Parent/caregiver requested that student not participate OR student refusal.

The numbers of non-participating students are provided in Table A2.9 broken down by jurisdiction and reason for non-participation.

Table A2.9 Student non-participation by jurisdiction

State/ Territory	Non-inclusion code					Total
	Absent	Functional disability	Intellectual disability	Limited language proficiency	Student or parent refusal	
ACT	54	1	6	2	0	63
NSW	181	0	3	0	2	186
NT	227	1	4	13	4	249
QLD	124	5	16	7	3	155
SA	141	3	6	3	3	156
TAS	143	2	8	5	3	161
VIC	215	1	9	7	26	258
WA	190	3	9	8	1	211
AUST	1275	16	61	45	42	1439

Appendix 3
Proficiency Levels, Assessment Strand Descriptors, Illustrative Items and Item Descriptors

Table A3.1 Proficiency Levels, assessment strand descriptors, illustrative items and item descriptors

Proficiency Level (scaled location)	Assessment strand descriptors	Descriptor: a student at this level may display skills like	Illustrative items and item descriptors
Level 4 and above (scaled score > 653)	Strand A: Formulates scientific questions, identifies the variable to be changed, the variable to be measured and in addition identifies at least one variable to be controlled. Uses repeated trials or replicates. Collects and records data involving two or more variables.	When provided with an experimental design involving multiple variables can identify the questions being investigated.	Understands that variables need to be controlled in fair tests in order to compare results meaningfully and to draw valid conclusions [in the context of an experiment on the rate of evaporation of two liquids]. Q4 Evaporating liquids
	Strand B: Calculates averages from repeat trials or replicates, plots line graphs where appropriate. Interprets data from line graph or bar graph. Conclusions summarise and explain the patterns in the science data. Able to make general suggestions for improving an investigation (e.g. make more measurements).	Conclusions summarise and explain the patterns in the data in the form of a rule and are consistent with the data.	Identifies a suggestion that would result in an improvement to the experimental method and justifies choice [in the context of investigating how long it takes to react and catch a falling ruler]. Q6 Reaction time (practical task)
	Strand C: Explains interactions, processes or effects that have been experienced or reported, in terms of a non-observable property or abstract science concept.	Explains interactions that have been observed in terms of an abstract science concept.	Explains the rising of a solar balloon in terms of the reduced density of the air inside the balloon [in the context of an experiment on the effect of sunlight on objects]. Q2 Solar energy

Proficiency Level (scaled location)	Assessment strand descriptors	Descriptor: a student at this level may display skills like	Illustrative items and item descriptors
Level 3.3 (scaled score 523–653)	Strand A: Formulates simple scientific questions for testing and makes predictions. Demonstrates awareness of the need for fair testing and appreciates scientific meaning of 'fair testing'. Identifies variable to be changed and/or measured but does not indicate variables to be controlled. Makes simple standard measurements. Records data as tables, diagrams or descriptions.	Demonstrates an awareness of the principles of conducting an experiment taking into account variables to be changed and/or measured.	Formulates the question investigated in an experiment, identifying what needs to be changed and what needs to be measured [in the context of an experiment on time taken for jelly to set]. Q1 Making jelly
	Strand B: Displays data as tables or constructs bar graphs when given the variables for each axis. Identifies and summarises patterns in science data in the form of a rule. Recognises the need for improvement to the method. Applies the rule by extrapolating and predicting.	Extrapolates from an observed pattern to describe an expected outcome or event.	Draws a conclusion that summarises the pattern in the data presented in a supplied table [in the context of an experiment on processes that break down rocks]. Q1 Changing rocks
	Strand C: Describes relationships between individual events (including cause and effect relationships) that have been experienced or reported. Can generalise and apply the rule by predicting future events.	Applies knowledge of relationship to explain reported phenomenon.	Explains the changing length of shadows during the day. Q4 Light and shadows

Proficiency Level (scaled location)	Assessment strand descriptors	Descriptor: a student at this level may display skills like	Illustrative items and item descriptors
Level 3.2 (scaled score 393–523)	Strand A: Formulates simple scientific questions for testing and makes predictions. Demonstrates awareness of the need for fair testing and appreciates scientific meaning of 'fair testing'. Identifies variable to be changed and/or measured but does not indicate variables to be controlled. Makes simple standard measurements. Records data as tables, diagrams or descriptions.	Collates and compares data set of collected information. Selects experimental design that represents a fair test.	Understands the need for fair testing by identifying the experiment that is a fair test [in the context of designing an experiment to determine whether marks on apples affect the apples' taste]. Q2 Food and energy
	Strand B: Displays data as tables or constructs bar graphs when given the variables for each axis. Identifies and summarises patterns in science data in the form of a rule. Recognises the need for improvement to the method. Applies the rule by extrapolating and predicting.	Interprets data and identifies patterns in – and/or the relationships between – elements of the data.	Provides a justification for disagreeing with a statement that incorrectly summarises data in a supplied column graph [in the context of investigating reaction time]. Q9 Reaction time (practical task)
	Strand C: Describes the relationships between individual events (including cause and effect relationships) that have been experienced or reported. Can generalise and apply the rule by predicting future events.	Interprets information in a contextualised report by application of relevant science knowledge.	Explains why decomposers are important in composting [in the context of recycling]. Q3 Recycling

Proficiency Level (scaled location)	Assessment strand descriptors	Descriptor: a student at this level may display skills like	Illustrative items and item descriptors
Level 3.1 (scaled score 262–393)	Strand A: Formulates simple scientific questions for testing and makes predictions. Demonstrates awareness of the need for fair testing and appreciates scientific meaning of 'fair testing'. Identifies variable to be changed and/or measured but does not indicate variables to be controlled. Makes simple standard measurements. Records data as tables, diagrams or descriptions.	Makes simple standard measurements and records data as descriptions.	All items addressing this strand at this level have been held secure.
	Strand B: Displays data as tables or constructs bar graphs when given the variables for each axis. Identifies and summarises patterns in science data in the form of a rule. Recognises the need for improvement to the method. Applies the rule by extrapolating and predicting.	Interprets simple data set requiring an element of comparison.	Interprets a column graph to identify the number of categories that match a specified criterion [in the context of investigating reaction time]. Q8 Reaction time (practical task)
	Strand C: Describes the relationships between individual events (including cause and effect relationships) that have been experienced or reported. Can generalise and apply the rule by predicting future events.	Selects appropriate reason to explain reported observation related to personal experience.	Identifies the relationship between blocking the path of light and the formation of a shadow. Q1 Light and shadows

Proficiency Level (scaled location)	Assessment strand descriptors	Descriptor: a student at this level may display skills like	Illustrative items and item descriptors
Level 2 and below (scaled score ≤ 262)	Strand A: Given a question in a familiar context, identifies that one variable/factor is to be changed (but does not necessarily use the term 'variable' to describe the changed variable). Demonstrates intuitive level of awareness of fair testing. Observes and describes or makes non-standard measurements and limited records of data.	Makes measurements or comparisons involving information or stimulus in a familiar context.	All items addressing this strand at this level have been held secure.
	Strand B: Makes comparisons between objects or events observed. Compares aspects of data in a simple supplied table of results. Can complete simple tables and bar graphs given table column headings or prepared graph axes.	Identifies simple patterns in the data and/or interprets a data set containing some interrelated elements.	Locates a piece of data in a supplied graph displaying day and night temperatures in the Simpson Desert. Q1 Life in the desert
	Strand C: Describes changes to, differences between or properties of objects or events that have been experienced or reported.	Makes a choice for a situation based on first-hand concrete experience, requiring the application of limited knowledge.	Identifies waste material that can be added to a compost heap [in the context of recycling]. Q2 Recycling

Appendix 4
NAP–SL 2012: Student Survey

NAP-SL 2012: STUDENT SURVEY

HOW TO FILL OUT THIS SHEET:

- Use a 2B or B pencil only. USE A 2B OR B PENCIL ONLY
- Do not use a pen.
- If you change your mind or make a mistake, rub it out completely before shading your preferred response.
- Mark only one answer for each question.
- Shade bubbles like this: ○ ● ○ ○ not like this: ✔ ✗ ⊖ ○

NAME: ______________________________

SCHOOL: ______________________________

SCHOOL CODE: ________________

In this section of the assessment you will be asked for your opinions and ideas about science and learning science.

Please read each sentence carefully and answer as accurately as you can. You may ask for help if you do not understand anything or if you are not sure how to show your answer.

Remember: There are no right or wrong answers. Your answers should be the ones that you think are best for you.

How much do you agree with the statements below?	**Strongly agree**	**Agree**	**Disagree**	**Strongly disagree**
1. I would like to learn more science at school.	○	○	○	○
2. I think it would be interesting to be a scientist.	○	○	○	○
3. I enjoy doing science.	○	○	○	○
4. I enjoy learning new things in science.	○	○	○	○
5. I learn science topics quickly.	○	○	○	○
6. I can usually give good answers to science questions.	○	○	○	○
7. I can understand new ideas about science easily.	○	○	○	○
8. Science is an everyday part of my life.	○	○	○	○
9. Science is important for lots of jobs.	○	○	○	○
10. Science is important: it changes how we live.	○	○	○	○
11. Scientific information helps people make good decisions.	○	○	○	○
12. Science is about remembering facts.	○	○	○	○
13. Science is about doing experiments.	○	○	○	○
14. Science is finding out about how things work.	○	○	○	○
15. Science is about solving problems.	○	○	○	○
16. Science is quite easy for most people to understand.	○	○	○	○

How often do you do these things outside of school?	Frequently (more than 2 times a week)	Often (1 or 2 times a week)	Sometimes (less than once a week)	Never
17. I view TV programs, DVDs or websites about science topics at home.	◯	◯	◯	◯
18. I read books, newspapers or magazine articles about science topics.	◯	◯	◯	◯
19. I talk about science ideas with my friends and family.	◯	◯	◯	◯

How often do you do these things at school?	Frequently (more than 2 times a week)	Often (1 or 2 times a week)	Sometimes (less than once a week)	Never
20. I view TV programs, DVDs or websites about science topics at school.	◯	◯	◯	◯
21. We read books, newspapers or magazine articles about science topics at school.	◯	◯	◯	◯

How often do you do these things at school?	Always	Mostly	Sometimes	Never
22. During science lessons I get to plan and carry out my own investigations.	◯	◯	◯	◯
23. When our class investigates things in science, we work in groups to carry out the investigation.	◯	◯	◯	◯
24. I use a computer for research or to present my science ideas and findings.	◯	◯	◯	◯
25. Our class has in-depth discussions about science ideas.	◯	◯	◯	◯

Which of these science topics have you studied at school?	Yes	No	I don't know
26. Earth and Space - for example, weather, soil, rocks, gravity, using Earth's resources, the planets, Sun and Moon.	◯	◯	◯
27. Energy and Force - for example, how toys and other machines work, electricity, heat, light, sound, magnets.	◯	◯	◯
28. Living Things - for example, living and non living things, how animals and plants survive in their environment, life cycles, interdependence.	◯	◯	◯
29. Matter - for example, the different properties of materials such as plastics and metals, the different uses of materials, changes to materials (solids, liquids and gases).	◯	◯	◯

	More than once a week	Once a week	Less than once a week	Hardly ever
30. How often do you have science lessons at school?	◯	◯	◯	◯

Do you agree with the statements below?	Yes	No
31. My classroom teacher teaches science to our class.	◯	◯
32. I think my teacher enjoys teaching science.	◯	◯
33. Our teacher invites visitors to school to talk to us about science topics.	◯	◯
34. Our class goes on excursions related to the science topics we are learning about.	◯	◯

Appendix 5
Student Survey: Comparison of 2009 and 2012 Results

Table A5.1 Comparison of student responses to common questions in the 2009 and 2012 Student Surveys

Common Survey questions	2009			2012			Change from 2009 to 2012	Statistically significant
	Strongly Agree/ Agree	Disagree/ Strongly Disagree	SE	Strongly Agree/ Agree	Disagree/ Strongly Disagree	SE		
I would like to learn more science at school.	74.4	25.6	0.7	81.2	18.8	0.7	6.8	Yes
I think it would be interesting to be a scientist.	66.0	34.0	0.7	58.7	41.3	0.8	-7.3	Yes
Science is an everyday part of my life.	51.0	49.0	0.8	37.2	62.8	0.9	-13.8	Yes
Science is important for lots of jobs.	80.1	19.9	0.6	79.6	20.4	0.6	-0.5	No
Science is about remembering facts.	73.6	26.4	0.7	57.0	43.0	0.7	-16.6	Yes
Science is about doing experiments.	87.9	12.1	0.4	82.8	17.2	0.5	-5.1	Yes
Science is finding out about how things work.	96.4	3.6	0.2	94.0	6.0	0.4	-2.4	Yes
Science is about solving problems.	71.6	28.4	0.8	73.5	26.5	0.6	1.9	Yes
	Always/Mostly	**Sometimes/ Never**	**SE**	**Always/Mostly**	**Sometimes/ Never**	**SE**		
During science lessons I get to plan and carry out my own investigations.	22.2	77.8	0.8	27.0	73.0	0.8	4.8	Yes
When our class investigates things in science, we work in groups to carry out the investigation.	61.3	38.7	1.1	60.4	39.6	1.0	-0.9	No
	Yes	**No**	**SE**	**Yes**	**No**	**SE**		
Science concept area studied at school: Earth and Space	92.6	7.4	0.5	89.1	10.9	0.6	-3.5	Yes
Science concept area studied at school: Energy and Force	83.8	16.2	0.9	86.1	13.9	0.7	2.3	No
Science concept area studied at school: Living Things	86.9	13.1	0.7	85.4	14.6	0.7	-1.5	Yes
Science concept area studied at school: Matter	75.6	24.4	1.1	80.4	19.6	0.8	4.8	Yes
My classroom teacher teaches science to our class.	76.4	23.6	1.6	70.7	29.3	1.6	-5.7	No

Note: Percentages in this table are weighted and might be slightly different from those reported in Chapter 7.