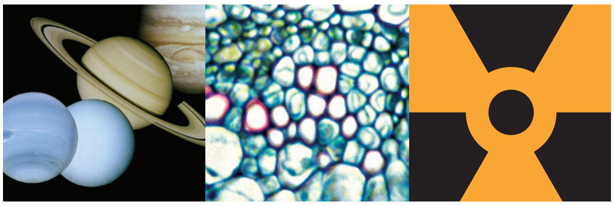
acara



National Assessment Program – Science Literacy Year 6 Report

2009

acara
AUSTRALIAN CURRICULUM,
ASSESSMENT AND
REPORTING AUTHORITY

NAP-SL 2009 Project Staff

Penny Hutton from Educational Assessment Australia (EAA) was the Project Director of NAP-SL 2009. Jennifer Cowing (EAA) was the Project Manager. The test development team was led by Sofia Kesidou (EAA). The Public Report was written by Penny Hutton, Sofia Kesidou, Goran Lazendic and Nathaniel Lewis. The School Release Materials were written by Penny Hutton, Jennifer Cowing and Philip Thompson (EAA).

The sampling and data analysis tasks were undertaken by Goran Lazendic, Nathaniel Lewis and Jennifer Lau from EAA and Margaret Wu and Mark Dulhunty from Educational Measurement Solutions (EMS). The Technical Report was written by Margaret Wu (EMS), Penny Hutton, Goran Lazendic and Nathaniel Lewis (EAA).

Cover Photograph
Top: courtesy of International Grammar School,
© Educational Assessment Australia

acara AUSTRALIAN CURRICULUM, ASSESSMENT AND REPORTING AUTHORITY

Contents

List of Tables

List of Figures

Acknowledgements

National Assessment Program – Science Literacy Review Committee Members

Listed below is the main review committee member for each jurisdiction and specialist area who participated in the Science Literacy Review Committee (SLRC) during the development and implementation of the 2009 National Assessment Program – Science Literacy, sample assessment. The assistance of others who came to the SLRC meetings is also acknowledged. The work of the committee made a valuable contribution to the success of the project.

Chair	Peter Titmanis Peter Bruce
State and Territory Representatives	
Australian Capital Territory	Dianne Grantham
New South Wales	Glen Sawle
Northern Territory	Louise Fogg
Queensland	Emily Sangster
South Australia	Jan Brooks
Tasmania	Meriloy Horsham
Victoria	Maria James
Western Australia	Marilyn McKee
Educational measurement specialist	Andrew Smith
Language background other than English specialist	Hanya Stefaniuk
Indigenous education specialist	Penny Hamilton
Australian Government representative	Mike Turner
Catholic school sector representative	Michael Doyle
Independent school sector representative	Glenda Leslie
Benchmarking and Educational Measurement Unit	Peter Bruce
Performance, Measurement and Reporting Taskforce member	Susan Dennett
Consultants	Mark Hackling Peter Fensham
Educational Assessment Australia	Penny Hutton
Educational Measurement Solutions	Margaret Wu

Foreword

This report presents the findings from the 2009 round of the National Assessment Program – Science Literacy (NAP-SL). It was carried out under the auspices of the national council of education ministers, the Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA).

The National Assessment Program – Science Literacy is one of a suite of national assessments (with Civics & Citizenship and ICT) which are conducted with a random sample of students on three-year cycles. The findings allow the measurement and reporting on how our students are progressing towards the achievement of the Educational Goals for Young Australians.

This third report on Year 6 Science Literacy provides a national comparison of student performance against the science literacy scale, and an analysis of various findings across states and territories and student sub-groups. It also allows the achievement of Year 6 students to be compared with the findings from the previous 2003 and 2006 science assessments.

A student survey was introduced in 2009 and the results provide an interesting insight into students' level of science participation, the activities undertaken and the importance students put on science and its role in our everyday lives.

Complementing this public report is a set of Science Literacy School Release Materials which will be a valuable resource for schools to use with their classes, allowing them to conduct a similar assessment and compare their students' achievements with those achieved nationally. A separate technical report detailing the assessment process has been produced and both resources are available on the MCEECDYA website.

The Australian Curriculum, Assessment and Reporting Authority (ACARA) 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.

ACARA acknowledges the work of the Science Literacy Review Committee (SLRC) in the development and implementation of this National Assessment Program. This report is commended to teachers, educators, the community and to all those interested in improving the level of student skills, knowledge and understandings in primary science.

Professor Barry McGaw AO
Chair
Australian Curriculum, Assessment and Reporting Authority (ACARA)

Executive Summary

In July 2001, the Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA, now the Ministerial Council for Education, Early Childhood Development and Youth Affairs, MCEECDYA) agreed to the development of assessment instruments and key performance measures for reporting on student skills, knowledge and understandings in primary science. It directed the newly established Performance Measurement and Reporting Taskforce (PMRT), a nationally representative body, to undertake the national assessment program.

The PMRT established a number of national committees to advise it on critical aspects of the study and ensure that the assessments and results were valid across the states and territories. The main function of these committees 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.

The National Assessment Program – Science Literacy measures scientific literacy. This is the application of broad conceptual understandings of science to make sense of the world, 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. The 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)

The first national assessment was conducted in 2003. The Primary Science Assessment Program (PSAP) – as it was then known – tested a sample of Year 6 students. PSAP results were reported in 2005.

In 2006, a consortium of Educational Assessment Australia (EAA) and Curriculum Corporation (now Educational Services Australia) conducted the second national science assessment. The National Assessment Program – Science Literacy (NAP-SL) tested a sample of Year 6 students. In 2009, this assessment was repeated by EAA with a new sample of Year 6 students in order to identify trends over time. The findings in this report describe the scientific literacy of Year 6 Australian students from the latest 2009 Science Literacy assessment, with comparisons made to 2003 and 2006.

Assessment domain

The assessment domain and instruments were developed in consultation with curriculum experts from each state and territory and representatives from the Catholic and independent school sectors.

The domain outlined the development of scientific literacy across three main areas:

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

Strand B: interpreting evidence and drawing conclusions from their 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 items drew on four concept areas:

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

These evolved from a review of the 'National Statements and Profiles' and were common across Australian curricula at the time of test development.

In August 2006 the Statements of Learning were endorsed by Ministers of Education in all states and territories. The 2009 National Assessment Program – Science Literacy tests drew on the concepts and content of the Statements of Learning in Science. It is expected that future National Assessment Program – Science Literacy tests will also draw on the Australian Curriculum – Science.

Assessment instruments

The assessment instruments were administered to a random sample consisting of 5 per cent of the total Australian Year 6 student population. The students' regular classroom teachers administered the National Assessment Program – Science Literacy on the following dates:

- 14 October 2009 – Northern Territory, Queensland, Tasmania, Victoria
- 21 October 2009 – Australian Capital Territory, New South Wales, South Australia, Western Australia.

The assessment instruments consisted of seven pen-and-paper assessments, including multiple-choice and open-ended items, and two practical tasks. Each student completed one of the pen-and-paper assessments and one of the practical tasks. Students were allowed 60 minutes for the pen-and-paper assessment 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. For the 2009 assessment, students also completed a 30-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

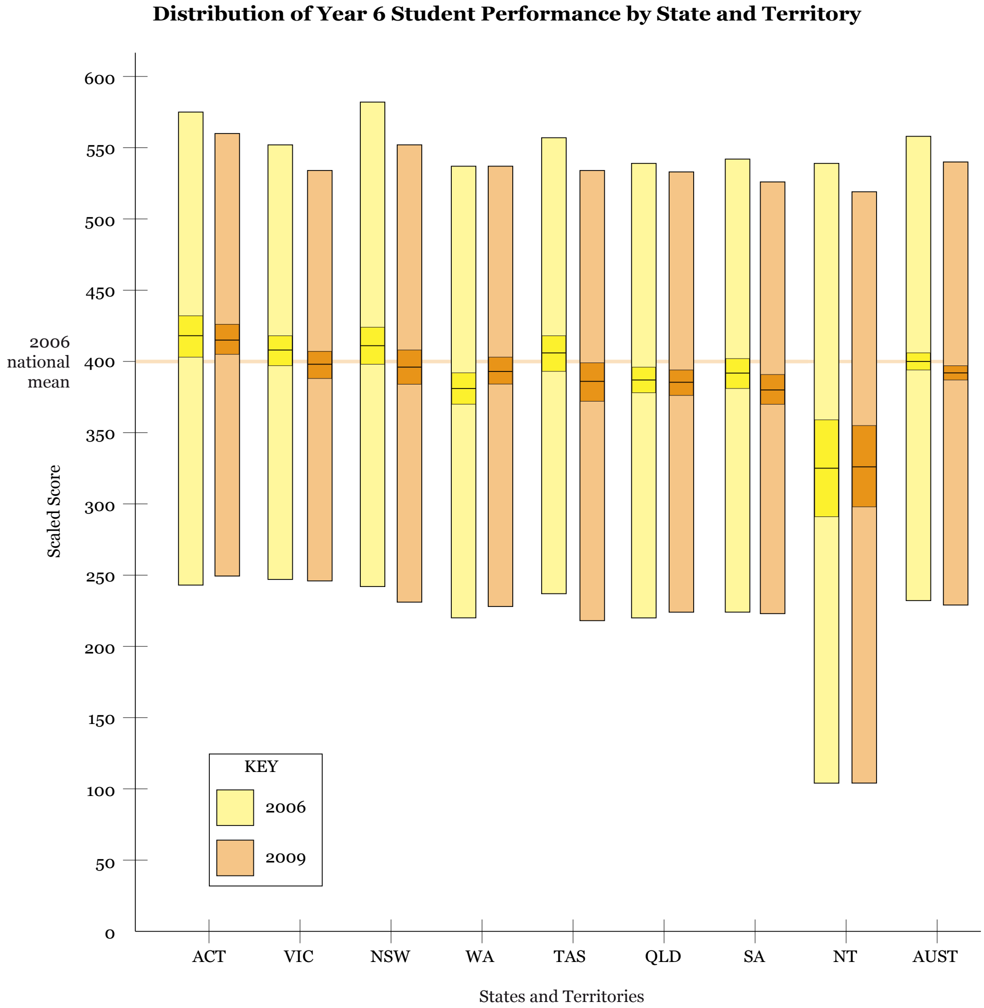
One of the main objectives of the National Assessment Program – Science Literacy is to monitor trends over time. The scientific literacy scale was initially established in 2003. However in 2006 a more robust test design was implemented, the sample frame was more inclusive of remote schools and items provided better discrimination of students. Consequently, the results of the 2006 assessment were utilised to establish a new baseline scientific literacy scale and the 2003 results were re-scaled onto it.

In 2009, student performance was scaled to the baseline established in 2006. Since the 2009 assessment has been aligned specifically to the 2006 assessment, only the differences between the 2006 and 2009 results were subjected to tests of statistical significance in all comparisons presented in this report. For this reason the majority of the trend commentary will be based on comparisons between 2006 and 2009.

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

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

Figure ES.1 Comparison of distributions of Year 6 student performance by state and territory in 2006 and 2009



Note: 2009 results scaled to 2006.

In technical terms, the darker coloured bands mark the likely range of the mean of the true population score. This is known as the confidence interval (CI).

The mean student achievement at the national level was somewhat lower in 2009 in the National Assessment Program – Science Literacy compared to that of 2006, however this difference is not statistically significant. Students from the Australian Capital Territory again had the highest mean while students from the Northern Territory had the lowest mean. The same outcomes were observed in 2006. Students from Western Australia increased their mean achievement in 2009 in comparison to 2006 results, albeit the difference is not statistically significant, whereas the mean score for Tasmania was significantly lower in 2009 compared to 2006.

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

- Indigenous students had a statistically significantly lower mean achievement than non-Indigenous students (see Table ES.1)
- students in remote and very remote areas had a statistically significantly lower mean achievement than students in all other geographic locations (see Table ES.2).

Table ES.1 Mean scores for Indigenous and non-Indigenous students in 2009

Student group	Mean score	95 per cent CI
Indigenous	297	±16.0
Non-Indigenous	397	±5.0

Table ES.2 Mean scores of students by school geographic location in 2009

MCEECDYA geographic location category	Percentage of students	Mean score
Metropolitan districts	72.3	395 (±6.2)
Provincial areas	24.7	389 (±7.9)
Remote and very remote areas	3.0	336 (±23.6)
AUST	100.0	392 (±5.1)

Note: figures in parentheses refer to 95 per cent confidence intervals and the percentages 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 and all other tables in this report 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 please refer to the 2009 Technical Report.

A Student Survey was administered for the first time as part of the 2009 assessment. Analyses showed no reliable correlations between student performance in scientific literacy and student responses to survey items. However, the survey provided interesting insights about student experiences with science learning at school and student perceptions of and attitudes towards science. Interestingly, over 50 per cent of students responded that they would like to do more science indicating a positive attitude towards this subject area exists.

Distribution of students across Proficiency Levels

A proficient standard for scientific literacy was established after the 2003 assessment to provide parents, educators and the community with a clear picture of the 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, university science educators, curriculum officers and experienced primary teachers in all states and territories, from government, Catholic and independent schools, were 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. This standard informed the development of the tests for the 2006 and 2009 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 have not achieved the proficient standard have demonstrated only partial mastery of the skills and understandings expected for Year 6; these students are on the way to becoming proficient. There are also students who have shown superior results and exceeded the proficient standard.

Initially, in 2003, three Proficiency Levels, corresponding with Levels 2, 3 and 4 of the scientific literacy Progress Map, were identified.

However, as 90 per cent of students’ scores were within Level 3 in 2003, three further Proficiency Levels within Level 3 were created, providing five levels for reporting student performance in the assessment. The proficient standard was deemed to be Level 3.2 on the Proficiency Level continuum.

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 areas such as reading, writing and numeracy where deficiencies will have significant effects on students’ future learning and functioning in society.

Table ES.3 Percentages of students in Proficiency Levels by state and territory in 2009

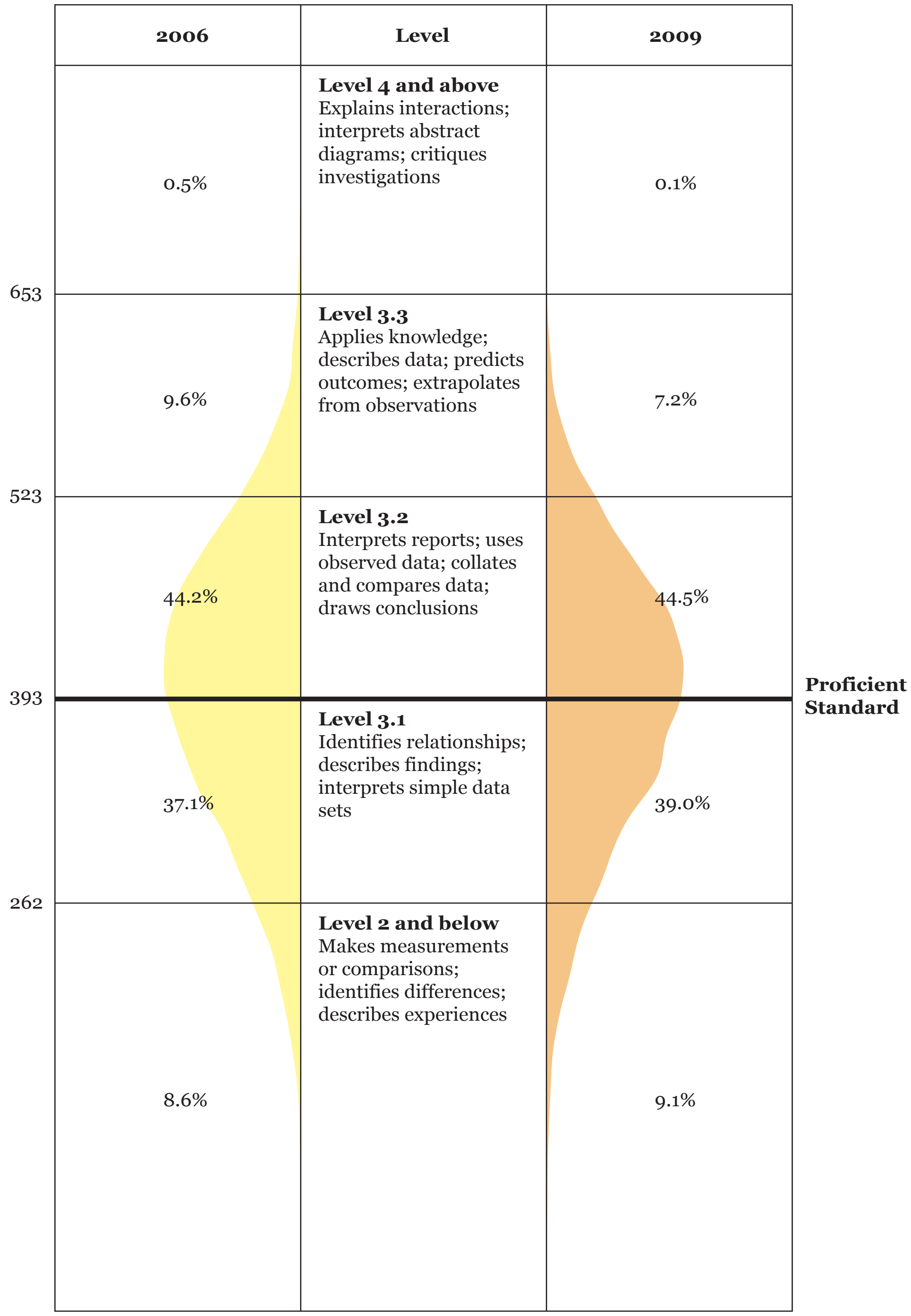
State/ Territory	Proficiency Level					At or above the proficient standard
	2 and below	3.1	3.2	3.3	4 and above	
NSW	8.7 (±2.6)	38.3 (±4.2)	43.8 (±3.9)	9.0 (±3.0)	0.2 (±0.3)	53.0 (±5.0)
VIC	6.8 (±2.3)	38.6 (±3.6)	48.0 (±4.1)	6.5 (±1.7)	0.1 (±0.1)	54.6 (±4.6)
QLD	10.5 (±2.9)	40.7 (±3.4)	42.7 (±3.3)	6.0 (±1.7)	0.1 (±0.2)	48.8 (±3.8)
SA	10.5 (±2.7)	43.0 (±4.0)	41.1 (±4.3)	5.3 (±1.6)	0.0 (±0.2)	46.5 (±5.0)
WA	9.1 (±2.4)	37.6 (±3.5)	46.2 (±3.8)	7.0 (±1.7)	0.1 (±0.2)	53.3 (±4.5)
TAS	11.2 (±3.8)	39.0 (±4.0)	43.2 (±5.5)	6.5 (±2.2)	0.1 (±0.3)	49.8 (±6.0)
NT	31.3 (±9.8)	35.1 (±5.4)	29.2 (±6.3)	4.4 (±2.6)	0.0 (±0.2)	33.6 (±7.5)
ACT	6.3 (±2.1)	32.5 (±4.0)	49.3 (±4.0)	11.8 (±3.1)	0.2 (±0.3)	61.2 (±4.8)
AUST	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)	51.9 (±2.2)

Note: 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 were observed in the Northern Territory. New South Wales, Victoria, Queensland, South Australia, Western Australia and Tasmania have very similar percentage distributions across Proficiency Levels and at the proficient standard.

In 2009, approximately 52 per cent of students, at the national level, attained the proficient standard or better in scientific literacy. Figure ES.2 shows a comparison, at the national level, of the percentage of students in each of the Proficiency Levels and the proportion of students performing at or above the proficient standard in 2006 and 2009.

Figure ES.2 Distribution of students in Proficiency Levels in 2006 and 2009



At the national level the difference between 2009 and 2006 in the proportion of students performing at or above the proficient standard is 2.5 per cent, which is not a statistically significant difference. Similarly, there was no significant difference in the distribution of students across Proficiency Levels in 2006 and 2009.

In conclusion, results of the 2009 National Assessment Program – Science Literacy, at the national level, remained the same as those observed in the 2006 assessment cycle, both in terms of mean student achievement and in the proportion of students performing at or above the proficient standard in scientific literacy.

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) (available online on the Ministerial Council for Education, Early Childhood Development and Youth Affairs, MCEECDYA website at www.mceecdya.edu.au).

The National Goals and now the Educational Goals provide the framework for reporting on student achievement through the annual MCEECDYA publication, the *National Report on Schooling in Australia*.

In July 2001, MCEECDYA agreed to the development of assessment instruments and key performance measures for reporting on student skills, knowledge and understandings in primary science. It directed the Performance Measurement and Reporting Taskforce (PMRT) to undertake the National Assessment Program.

The PMRT set the policy objectives and established a Steering Committee to manage the assessment and a Consultative Committee to facilitate discussion among the jurisdictions and school sectors. The Consultative Committee also provided feedback about the appropriateness of the conceptual framework and reviewed the assessment items to ensure that they were inclusive of all the state and territory curricula.

The National Science Assessment was the first assessment program designed specifically to provide information about performance against the National Goals (now the Educational Goals). MCEECDYA has also endorsed similar assessment programs to be conducted for Civics and Citizenship, and Information and Communications Technology Literacy. The intention is that each assessment program will be repeated every three years so that performance in these areas of study can be monitored over time.

Of the three subject areas, science is the only program that focuses entirely on primary school performance. This is because MCEECDYA has agreed to use the Program for International Student Assessment (PISA) as the measure of performance for secondary science literacy.

In January 2008, PMRT awarded the contract for the third cycle of science testing, due in 2009, to Educational Assessment Australia (EAA). The Benchmarking and Educational Measurement Unit (BEMU) was nominated by PMRT to liaise between the contractor and PMRT in the delivery of the project. Both PMRT and BEMU were incorporated into the Australian Curriculum, Assessment and Reporting Authority (ACARA) in 2009.

The National Assessment Program – Science Literacy

Implementation of the National Assessment Program – Science Literacy involved a large number of separate but related steps, including the development of items and instruments to assess the assessment domain; the trialling of those items and assessment instruments; the administration of the 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, describes the testing process, presents the results at the national and state and territory levels and includes comparisons with 2006. Where valid, comparisons are also made to the 2003 testing cycle.

What does the National Assessment Program – Science Literacy measure?

The National Assessment Program – Science Literacy measures scientific literacy.

Scientific literacy has been defined by the OECD – Programme for International Student Assessment (OECD–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 2009 National Assessment Program – Science Literacy in accord with the Ball et al. 2000 report recommendation.

The science items and instruments therefore assess outcomes that contribute to scientific literacy, such as conceptual understandings, rather than focusing solely on facts. They also assess student competence in carrying out investigations in realistic situations.

The National Assessment Program – Science Literacy 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 the 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 were assessed?

Three main areas of scientific literacy were assessed:

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

Strand B: interpreting evidence and drawing conclusions from their 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.

A conscious effort was made to develop assessment items that related to everyday contexts.

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

The intention was to ensure that all Year 6 students were familiar with the materials and experiences to be used in the National Assessment Program – Science Literacy 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, university science educators, curriculum officers and experienced primary teachers in all states and territories, from government, Catholic and independent schools, were 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

Levels 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 have not achieved the proficient standard have demonstrated 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 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 will be conducted to gauge whether student proficiency has improved.

Information about students' performances in relation to the Year 6 standard from the third (2009) National Assessment Program – Science Literacy is reported with comparisons to 2003 and 2006 data by proficiency levels in Chapter 5.

Who participated in the 2009 National Assessment Program – Science Literacy?

Approximately 5 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.

A grade-based population of students enrolled at schools was chosen. This is consistent with the other strands of the National Assessment Program. 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 structural differences that may assist interpretation of the results of the testing is summarised in Table 3.1.

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.

Table 1.1 Number of schools and students by state and territory in the final sample 2009

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	56	55 (98.2%)	1311	1199 (91.5%)
NSW	92	91 (98.9%)	2258	2092 (92.6%)
NT	50	38 (76%)	831	743 (89.4%)
QLD	92	92 (100%)	2228	2043 (91.7%)
SA	95	93 (97.9%)	2005	1848 (92.2%)
TAS	63	63 (100%)	1276	1167 (91.5%)
VIC	93	93 (100%)	2243	2040 (90.9%)
WA	94	93 (98.9%)	2208	2030 (91.9%)
AUST	635	618 (97.3%)	14360	13162 (91.7%)

From Table 1.1 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, 12 were excluded from the final sample for various reasons. Four schools were exempted, one school was deemed ineligible and seven schools had insufficient eligible students present on test day.

What did the National Assessment Program – Science Literacy participants have to do?

There were seven pen-and-paper (objective) assessments which included multiple-choice and open-ended items. There were also two practical assessment tasks. The assessment papers included common items. The papers were distributed randomly so that each of the students in a class completed one of the seven pen-and-paper assessments.

However, all students in the same class undertook the same practical task. The practical tasks were assigned to classes across Australia in a way that ensured approximately equal numbers of classes attempted each of the two tasks.

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 had collected as a group.

The students then answered a set of items independently, based on their observations and the data they had collected. The individual student responses were the only ones 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 into one instrument using results obtained from students doing the same objective assessment.

Students were allowed 60 minutes to complete the pen-and-paper assessments and 45 minutes for the practical tasks.

In addition to the assessment tasks, students were also asked to respond to a 30-item survey. The Student Survey sought to obtain information about students' perceptions of and attitudes to science and their experiences of science learning at their school. Results of the survey are summarised in Chapter 7.

The students' regular classroom teachers administered the National Assessment Program – Science Literacy on:

- 14 October 2009 – Northern Territory, Queensland, Tasmania, Victoria
- 21 October 2009 – Australian Capital Territory, New South Wales, South Australia, Western Australia.

How are the National Assessment Program – Science Literacy results reported?

The results of the National Assessment Program – Science Literacy 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, Interpreting the Scientific Literacy Results. Chapter 3 includes results in relation to the 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 also presented in this report in Chapter 6.

How is this report organised?

Chapter 2 provides detailed information about the assessment, including explanations of the assessment domain and the assessment procedures.

Chapter 3 provides a description of the scientific literacy scale. It also includes results in terms of means and distributions of student performance for each state and territory as well as the Australian population.

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 tasks accompanies the discussion of students' results.

Chapter 5 examines comparisons in achievement by Proficiency Levels between the tests in the 2003, 2006 and 2009 cycles.

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 reproduces items from the Student Survey which was undertaken for the first time in 2009 and includes an analysis of their responses to a number of questions and statements.

Chapter 8 provides a brief summary of the main findings of the 2009 National Assessment Program – Science Literacy.

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

In the process of constructing the National Assessment Program – Science Literacy test, the following steps were undertaken, involving a number of inter-related tasks:

1. clarifying the assessment strands for scientific literacy
2. constructing assessments that comprised items and tasks which defined the assessment strands operationally
3. administering the assessments to students
4. using the measurement model and technical standards to analyse the results.

As in 2006, the PMRT established a number of national committees to ensure that the assessments and results were valid across the states and territories and to advise it on critical aspects of the study.

The main function of these committees and groups was to ensure that the assessment strands of scientific literacy were 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.

For the 2009 cycle, a Science Literacy Review Committee (SLRC) was consulted about item development and review, as well as other issues as they arose.

A brief description of the steps involved in developing the assessment is provided here.

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

   A common understanding of the Progress Map, the descriptions of each strand, and the hierarchy of students' understandings and skills in the concept areas was developed (see Appendix 1). The names of the conceptual strands were also adjusted to match those used in the Statements of Learning.

2. **Constructing assessments that comprised items and tasks which defined the assessment strands operationally and covered the full range of proficiency expected to be represented in Year 6 classes**

   In consultation with EAA, BEMU (on behalf of PMRT) approved the more technical aspects of the assessment design including, for example, the number of assessment booklets, 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 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 advisory committees and other key staff in the states and territories. 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. Feedback received was used to refine the assessment items.

3. **Administering the assessments to students**

   Once the items and tasks had been written and reviewed, they were trialled with a sample of students in 30 schools selected from the government, Catholic and independent sectors in the Australian Capital Territory, New South Wales, Queensland and Western Australia.

   The results were analysed to determine the degree to which the items and tasks measured the scientific literacy domain. The committees then reviewed the data from the trial testing, gauged the validity of the assessments and suggested modifications where necessary. These modifications were included in the revised assessments.

The final assessments were administered to a stratified random sample of students in October 2009. The final sample contained a total of 13 162 students at 618 schools. Information about the achieved sample is shown in Appendix 2.

4. **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 the National Assessment Program – Science Literacy. These statistical models are used in all state and territory testing programs and in major international testing programs such as PISA and the Trends in International Mathematics and Science Study (TIMSS).

   Details of the application of the Rasch model can be found in the 2009 Technical Report for the National Assessment Program – Science Literacy.

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

The assessment booklets

In 2009, the National Assessment Program – Science Literacy involved the use of seven assessment booklets. A booklet rotation design similar to that used in other sample-based international assessments was implemented. The effect of such a design is to reduce the possibility that an item's location in a test booklet has an impact on its difficulty.

In addition to those items written specifically for the 2009 cycle, 29 items from the 2003 and 2006 cycles which had been held secure were incorporated into the item pool to enable trend analysis to be undertaken.

To achieve the rotation design for the National Assessment Program – Science Literacy, the items were first written in units. Each unit had a context and contained between one and five items. Clusters were then constructed by grouping four to six units together. From there, booklets were compiled by arranging three clusters in every booklet. In total there were seven different clusters across the seven booklets. Each booklet had approximately 40 items in the objective section.

The multiple-choice items in the booklets had only single correct answers. The open-ended items required students to construct their own responses. These were categorised into those that required a single word or short sentence response (short-answer items) and those that required more substantive responses (extended-response items).

Each booklet contained an objective (pen-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. The practical task required students to undertake an activity in groups of three students and collect and record data from that activity. Students then responded individually to either ten or twelve items related to the activity they had completed.

Project objectives

For the 2009 cycle, the same assessment domains were tested as in 2003 and 2006 so that student achievement levels could be compared against the same proficiency levels previously reported. The project specifications required that EAA:

- develop an equating design that would enable the results of the 2009 testing cycle to be compared to the 2006 testing cycle
- develop assessment instruments, including pen-and-paper tests and practical tasks
- pilot draft assessment instruments, where applicable, to assess the practicality of instruments
- conduct a trial of all the proposed materials in a sample of Australian schools in 2008
- review and refine draft assessment instruments and marking keys based on empirical evidence and feedback from PMRT's Science Literacy Review Committee (SLRC)
- develop and trial a Student Survey instrument in consultation with the SLRC
- review and refine the Student Survey based on trial data and SLRC advice
- administer the main assessment to Year 6 students in a stratified sample of schools across all states and territories
- collect and analyse the data from the main assessment
- report on the 2009 performance with comparisons, where applicable, to the 2006 and 2003 test cycles
- produce a Public Report, Technical Report and select a set of items suitable for School Release.

Coverage of scientific literacy

The distribution of items across the assessment domain for scientific literacy (each strand and major conceptual area) is shown in Table 2.1. There were 113 items distributed across the seven objective pen-and-paper tests and two practical tasks. Each student had to complete one pen-and-paper test and one practical task.

Table 2.1 Distribution of assessment items across the assessment strands for scientific literacy 2009

Domain	**Item type and number of items**			
	Multiple choice	**Short answer**	**Extended response**	**Total**
Distribution of items by strand				
Strand A	5	4	9	18
Strand B	20	3	18	41
Strand C	22	9	23	54
Total	47	16	50	113
Distribution of items by major science conceptual area				
Earth and Space	18	0	7	25
Energy and Force	9	5	17	31
Living Things	7	1	18	26
Matter	13	10	8	31
Total	47	16	50	113

The domain

The scientific literacy domain comprises three strands. 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 school 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.

Table 2.1 shows that 59 of the items assessed the process strands (18 for Strand A and 41 for Strand B) and 54 assessed the conceptual understanding strand (Strand C).

Major science concepts

Items were developed in relation to four major concept areas: Living Things; Earth and Space; Matter; and Energy and Force as shown in Table 2.1 on page 13. These concept areas were selected in consultation with BEMU and the SLRC as being those that are found most widely in state and territory curricula.

Types of assessment items

The strands and major concepts of scientific literacy were assessed through a range of item types (Table 2.1). Of the 113 items, 47 were classified as being multiple-choice, 16 as short-answer and 50 as extended-response.

All the items were presented in item sets or units, with between one and five items pertaining to each stimulus text and/or diagram(s).

The sampling procedures

As in 2006, the sample design for National Assessment Program – Science Literacy 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. 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 2009 (as for 2006), as many schools 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 pen-and-paper tasks 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 2009 Technical Report (available online on the MCEECDYA website at www.mceecdya.edu.au).

Assessment administration procedures

Students' regular class teachers administered the National Assessment Program – Science Literacy, 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 National Assessment Program – Science Literacy, 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 (refer to Table A2.6 in Appendix 2).

Teachers were able to review the Test Administrator's Manual before the assessment date and raise questions with the coordinators of the National Assessment Program – Science Literacy in their jurisdiction. EAA provided a toll-free telephone number and 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 Assessment in a random sample of classes in 32 of the 618 schools involved. The invigilators reported conformity with the administration procedures.

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.

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 the scores agreed to by expert scorers (the group leaders, the Test Development Manager and the Professional Leader). 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 Professional Leader. 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 10 per cent of the 2006 National Assessment Program – Science Literacy trend items were also marked by the 2009 markers to assure 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 the National Assessment Program – Science Literacy were provided with feedback about the performance of their students on the Assessment prior to the close of the 2009 school year. The reports showed the results for each student on an item-by-item basis and comparative data showing the percentage of the school and the national sample of students responding correctly to the item or, 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.

National Assessment Program – Science Literacy School Release Materials

Some assessment items have been released from the 2009 National Assessment Program – Science Literacy to enable teachers to administer the tasks under similar conditions and gauge their own students' proficiency in relation to the national standards. The School Release Materials comprise an objective test containing 37 multiple-choice, short-answer and extended-response questions as well as a practical task. The School Release Materials will be made available on the MCEECDYA website at www.mceecdya.edu.au

The remaining 2009 assessment items have been secured for the purpose of equating the next National Assessment Program – Science Literacy (which is to be undertaken in 2012) and, together with the 2003 and 2006 assessments, will allow longitudinal data on student performance to be obtained.

Chapter 3
Student Performance in Scientific Literacy for 2009

Introduction

In this chapter, summary statistics for the 2009 National Assessment Program – Science Literacy are shown in terms of students' 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 results of the National Assessment Program – Science Literacy is provided. This chapter also contains comparisons of performance of Year 6 students over the 2006 and 2009 assessment cycles. Analyses indicate that 2009 results to a large extent correspond to those observed in 2006.

Scientific literacy scale

Setting 2006 results as the baseline

A scientific literacy scale was constructed in 2003, using the Rasch model. The Rasch analysis produced information about the relative difficulty of 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 the National Assessment Program – Science Literacy 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. These reasons included:

1. The 2006 test design was more robust than the 2003 test design.

2. There were considerably more items in 2006 than in 2003, resulting in a better coverage of the assessment domain in 2006.

3. The 2006 items were generally more discriminating than the 2003 items.

4. The 2006 sampling was more comprehensive, as remote schools were also included in the sample (see 2006 Technical Report for more information).

The Rasch 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 2009 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.

Establishing Proficiency Levels

One of the main objectives of the National Assessment Program – Science Literacy 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 and possess the understandings and skills of lower Proficiency Levels. The National Assessment Program – Science Literacy 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 obtained a score equal to or above the Level 3.2 cut-point of 393 were deemed to have attained the proficient standard in scientific literacy.

An overview of 2009 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.

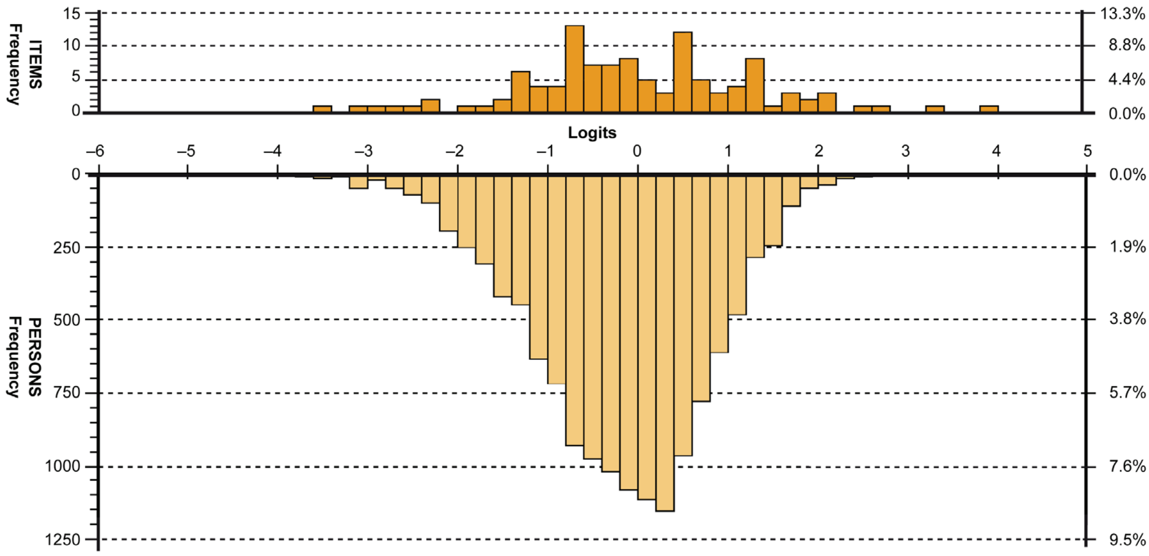
Analysing the 2009 results

The measurement model for analysing students' responses in 2009 was the same as that of the 2006 assessment. The common item equating methodology was used to place the 2009 results on the scientific literacy scale.

In 2009, care was taken to increase the number of trend items that could be used to equate the 2009 results to the 2006 scale. A total of 29 trend items were included in the 2009 assessment. This included all 9 trend items from the 2003 assessment and a further 20 items from the 2006 assessment. The trend items covered a range of scientific literacy strands, concept areas and range of item difficulties. From a total of 29 potential trend items, 20 items were used to equate the 2009 item and student parameters to the scientific literacy scale established in 2006 (for detailed information see 2009 Technical Report).

Before presenting data for the application of the Rasch model it is important to ensure that the test has appropriately targeted the student population. As can be seen from Figure 3.1, the 2009 assessment achieved a good spread of item difficulties and was appropriately matched to the Year 6 cohort. This demonstrates the items were of excellent quality and were able to discriminate between achievements at the highest level while still catering for less able students.

Figure 3.1 2009 National Assessment Program – Science Literacy Item: Person map



Achievements by state and territory in 2009

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 1 mths
NSW	12 yrs 0 mths
NT	11 yrs 10 mths
QLD	11 yrs 5 mths
SA	11 yrs 11 mths
TAS	12 yrs 3 mths
VIC	12 yrs 2 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.

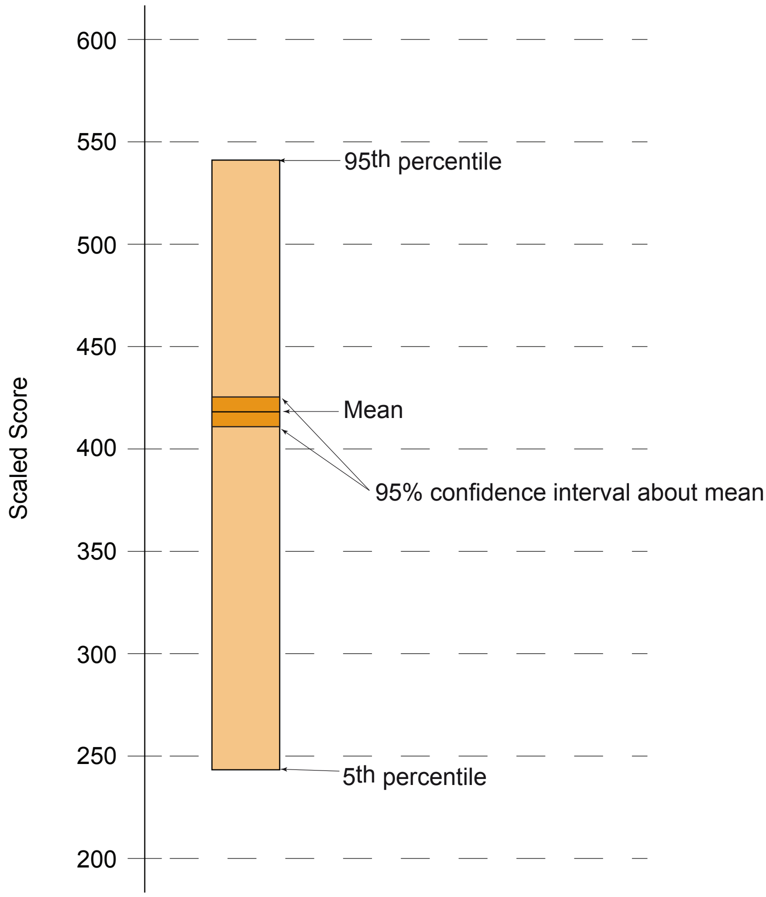
Reading bar charts

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 5 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 5 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 the National Assessment Program – Science Literacy 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.

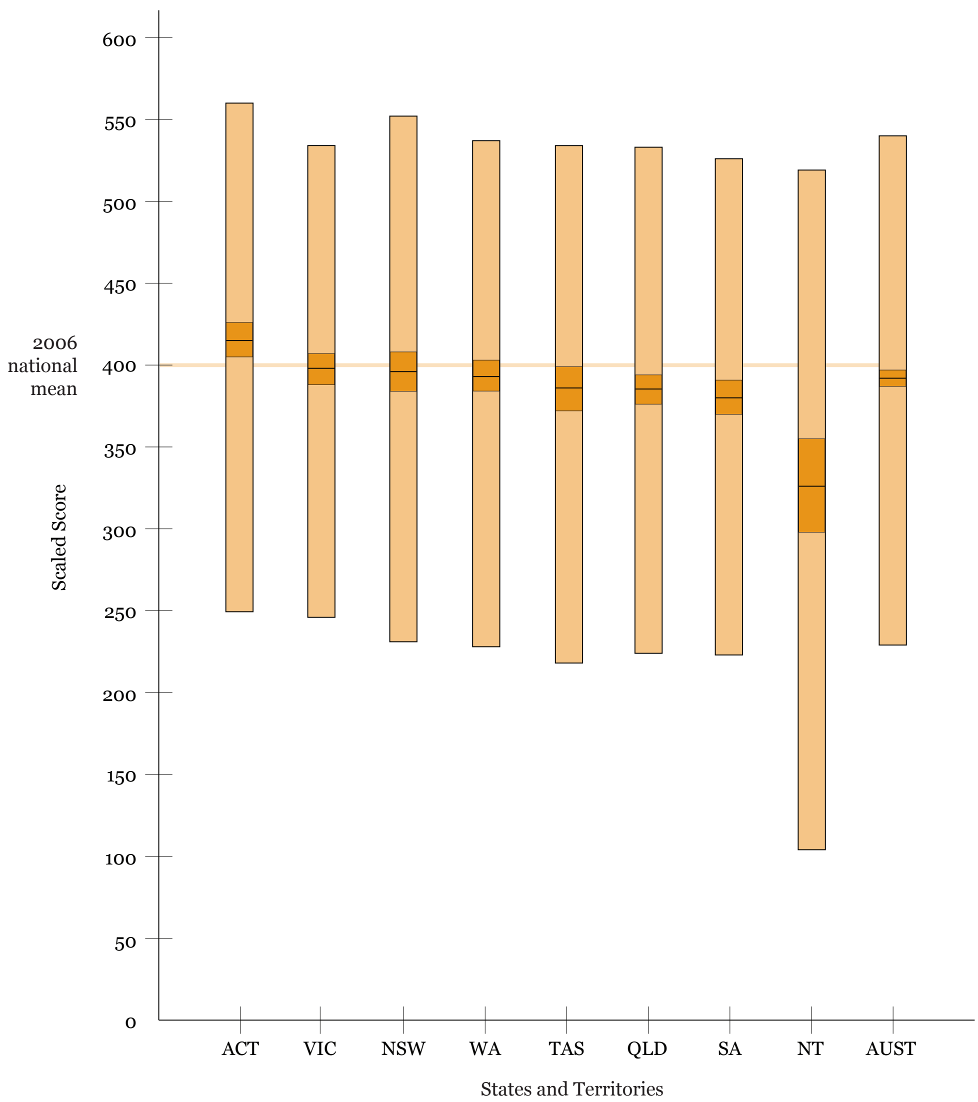
Confidence intervals in this report were constructed with a standard statistical precision. This precision is such that the probability that a confidence interval does not contain the true population mean equals the probability of obtaining five incorrect measurements out of 100 random draws. Such a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

Mean and range of students' scores in 2009

Figure 3.3 shows student performance in scientific literacy for each state and territory in 2009 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 2009 distribution of student performance by state and territory

Distribution of Year 6 Student Performance by State and Territory



	ACT	VIC	NSW	WA	TAS	QLD	SA	NT	AUST
Mean	415	398	396	393	386	385	380	326	392
95% CI	±10.6	±9.2	±12.1	±9.6	±13.5	±8.9	±10.4	±28.6	±5.1

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 the score range and confidence interval.

Additional information about the range of students' scores in 2009 is provided by listing the scaled scores corresponding to the standard range of percentile values by each state and territory in Table 3.2.

Table 3.2 Distribution of percentile scores by state and territory

State/ Territory	Mean score	95 per cent confidence interval	Percentile						
			5th	10th	25th	50th	75th	90th	95th
ACT	415	±10.6	249	291	354	420	481	533	560
VIC	398	±9.2	246	283	342	401	458	507	534
NSW	396	±12.1	231	270	332	400	463	519	552
WA	393	±9.6	228	268	335	400	460	509	537
TAS	386	±13.5	218	255	322	392	453	504	534
QLD	385	±8.9	224	259	325	389	449	499	533
SA	380	±10.4	223	259	319	384	446	496	526
NT	326	±28.6	104	149	233	338	423	486	519
AUST	392	±5.1	229	268	332	397	457	509	540

Table 3.2 shows that the Northern Territory has lower percentile scores than all other jurisdictions, whereas the Australian Capital Territory has the highest percentile scores.

Comparisons of means by state and territory in 2009

Table 3.3 contains results of a series of pair-wise comparisons between means for states and territories to determine if the jurisdictional differences were 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 Table 3.3 it is possible to draw a comparison between any two jurisdictions. Results below the diagonal (the lower left-hand half) do not include the Bonferroni adjustment, while the results above the diagonal (the upper right-hand half) incorporate the Bonferroni adjustment. Comparisons that are statistically significant are shown by an upward or downward symbol.

Table 3.3 Multiple comparisons of scientific literacy results by state and territory for 2009 with and without Bonferroni adjustment

			ACT	VIC	NSW	WA	TAS	QLD	SA	NT
		Mean score	415	398	396	393	386	385	380	326
	Mean score	95% CI	±10.6	±9.2	±12.1	±9.6	±13.5	±8.9	±10.4	±28.6
ACT	415	±10.6		•	•	▲	▲	▲	▲	▲
VIC	398	±9.2	▼		•	•	•	•	•	▲
NSW	396	±12.1	▼	•		•	•	•	•	▲
WA	393	±9.6	▼	•	•		•	•	•	▲
TAS	386	±13.5	▼	•	•	•		•	•	▲
QLD	385	±8.9	▼	▼	•	•	•		•	▲
SA	380	±10.4	▼	▼	•	•	•	•		▲
NT	326	±28.6	▼	▼	▼	▼	▼	▼	▼	

▲	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 3.3 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 in Victoria and students from either South Australia or Queensland were also statistically significant when the test was conducted without the Bonferroni adjustment.

However, when the Bonferroni adjustment was implemented, the mean score for students from the Australian Capital Territory was no longer significantly different from those for Victoria and New South Wales. Similarly Victoria was no longer shown to be significantly different to South Australia or Queensland.

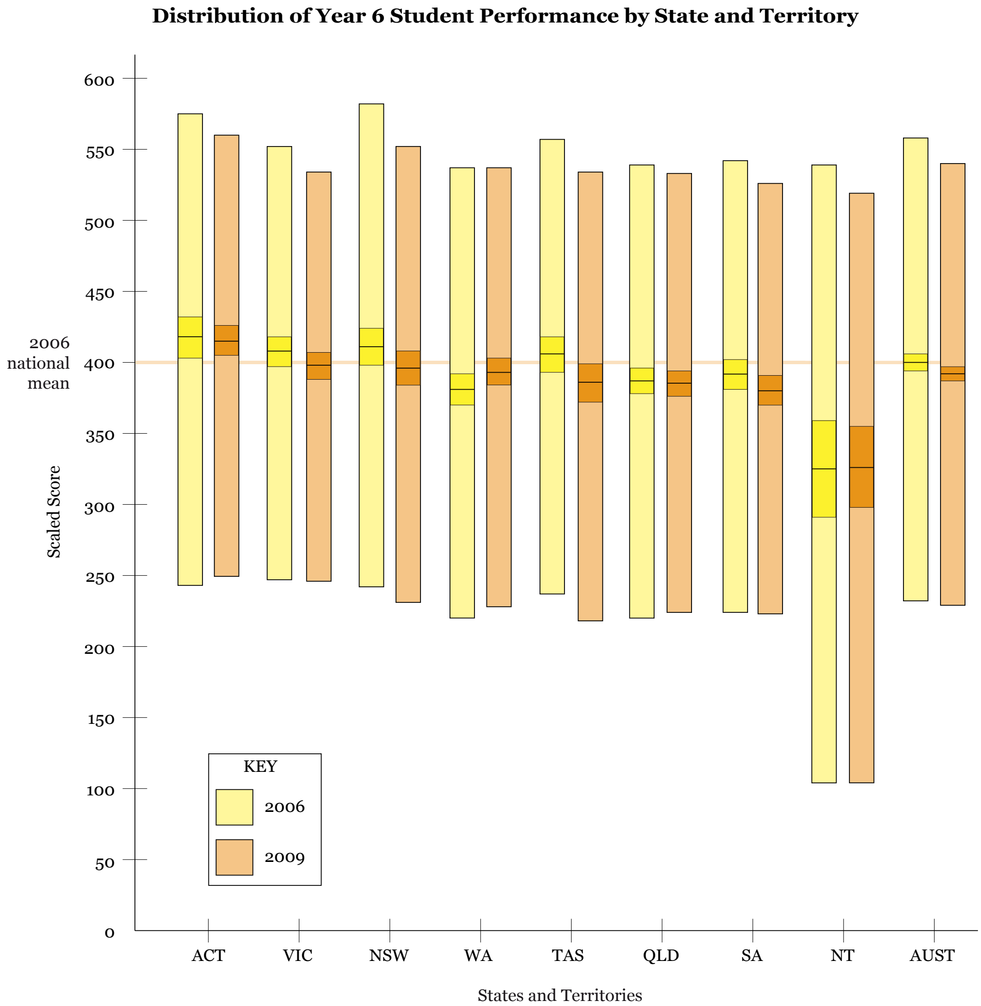
Students from the Northern Territory achieved a substantially lower mean score than students in all other states and territories with all comparisons producing a statistically significant difference, with and without Bonferroni adjustment.

Comparisons of student results in 2009 and 2006

The 2009 National Assessment Program – Science Literacy was the third time the science domain had been assessed in the national assessment programs, with the first being carried out in 2003 (Primary Science Assessment Program). Given that the 2003 assessment differed from the 2006 assessment in terms of item booklet design, 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 2009 cycles. To test whether the 2009 results differ from that of previous assessment cycles, a link error is 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 2009 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 means 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 and 2009



As depicted in Figure 3.4, overall the 2009 means at the national level and at the state and territory level were somewhat lower than those of 2006. The exception was Western Australia which showed an increase of 12 points in 2009 compared to the mean student achievement in 2006.

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

Table 3.4 Comparison of 2006 and 2009 jurisdiction mean scores ranked by 2009 achievement

State/ Territory	Mean score		Change from 2006 to 2009	Statistically significant
	2006	2009		
ACT	418 (±14.3)	415 (±10.6)	−3	NO
VIC	408 (±10.2)	398 (±9.2)	−10	NO
NSW	411 (±12.5)	396 (±12.1)	−15	NO
WA	381 (±10.0)	393 (±9.6)	12	NO
TAS	406 (±12.1)	386 (±13.5)	−20	YES
QLD	387 (±8.6)	385 (±8.9)	−2	NO
SA	392 (±10.0)	380 (±10.4)	−12	NO
NT	325 (±33.7)	326 (±28.6)	1	NO
AUST	400 (±5.4)	392 (±5.1)	−8	NO

Note: 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 2006 and 2009 at the national level nor at the state and territory level. The only exception is Tasmania for which the difference of 20 points represents a statistically significant decrease in students' mean achievement in 2009. The 2009 increase in mean achievement of students in Western Australia, while greater than that for any other jurisdiction, was not 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 and 2009.

Table 3.5 State and territory mean scores in 2006 and 2009 rankings

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

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

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. In 2009, Victoria obtained second rank and New South Wales moved to third rank. Tasmania moved from fourth position in 2006 to fifth position in 2009. South Australia moved from fifth position in 2006 to seventh position in 2009. The Australian Capital Territory, Queensland and the Northern Territory did not change positions. However, given that the differences in mean achievement were not statistically significant between New South Wales, South Australia, Western Australia, Victoria, Tasmania and Queensland, the change in ranking order for these states should be regarded as indicative only.

Trends in mean achievement in scientific literacy

Tests of statistical significance between 2003 and 2009 results were deemed not to be sound owing to the reasons detailed at the beginning of this chapter. Nevertheless, an overview of the trends in scientific literacy at the national level for 2003, 2006 and 2009 is provided below in Table 3.6.

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

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

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

Summary

In summary, the 2009 results are similar to those of 2006 both in terms of student mean achievement and the distribution of student scores. That is, the analysis shows that, there are no statistically significant differences between the 2006 and 2009 results at either jurisdiction or national levels. The only exception is Tasmania, where the decrease in 2009 was statistically significant compared to that of the 2006 mean.

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'.

The next section discusses the establishment of the Proficiency Levels and the cut-off scores for each of the levels.

This chapter also 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 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 and possess the understandings and skills of lower Proficiency Levels.

Initially, in 2003, 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 2003, three further Proficiency Levels within Level 3 were created, providing five levels for reporting student performance in the assessment.

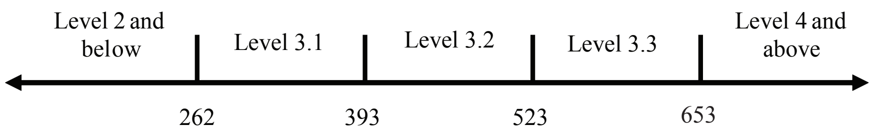
The cut-off points, which denote the boundaries between Proficiency Levels, were established in 2003 using a combination of experts' knowledge of the skills required to answer each scientific literacy item and information from the analysis of students' responses. The same cut-off points were used to determine the Proficiency Levels for the 2009 test.

The difficulty range spanned by each proficiency level was such that students whose scores were at the top of a level had 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 were at the bottom of a level had 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.

Figure 4.1 Cut-off scores

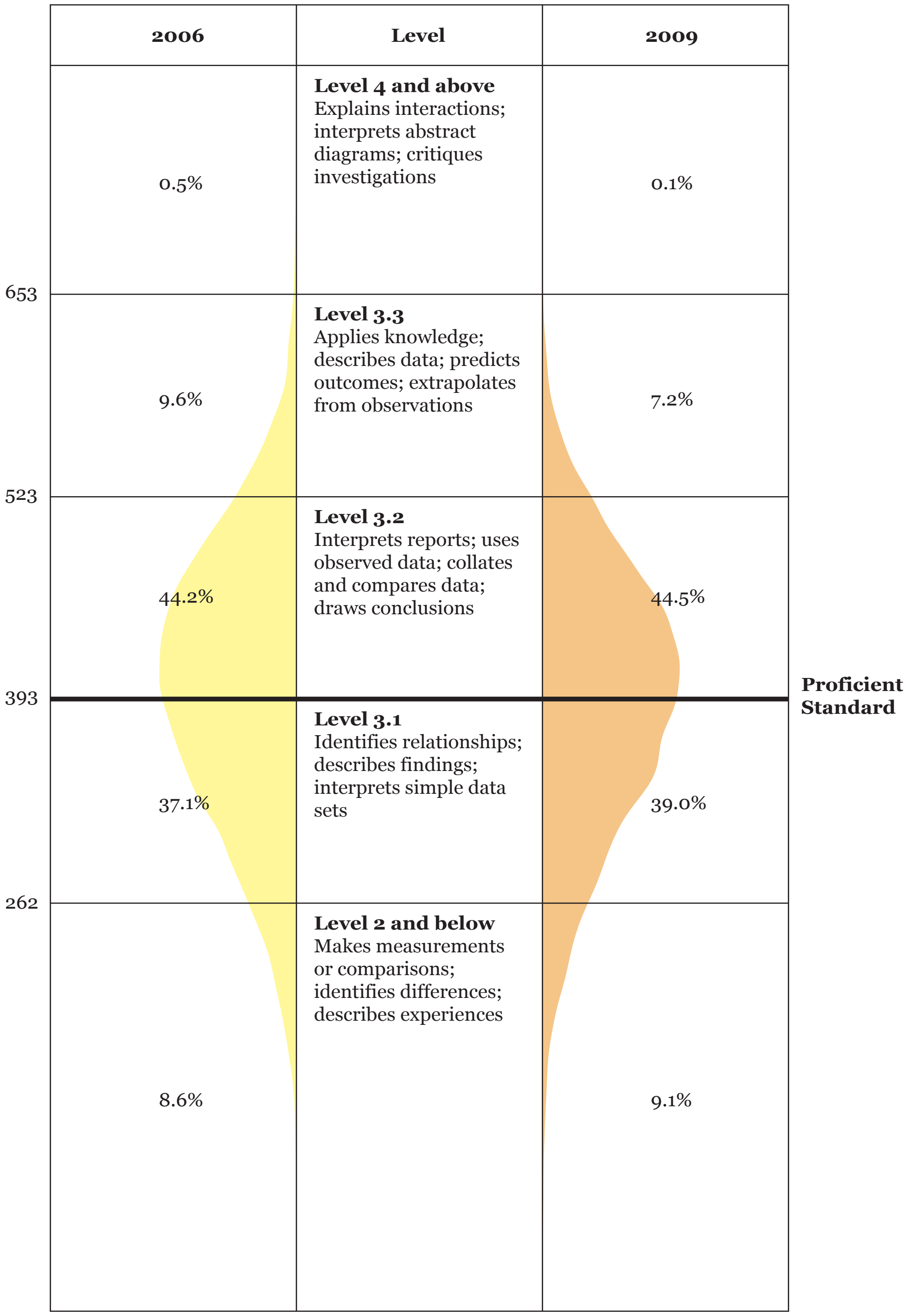


A score of 653 or more locates students in Proficiency Level 4 and above.

Similarly, scores in the range of 262 to 653 relate to Proficiency Level 3 on the assessment framework.

Figure 4.2 shows comparisons, at the national level, of the percentage of students in each of the Proficiency Levels in 2006 and 2009.

Figure 4.2 Distribution of students in Proficiency Levels for 2006 and 2009



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

Between 2006 and 2009, at the national level, the difference between the proportion of students performing at or above the proficient standard is 2.4 per cent, which is not a statistically significant difference. Similarly, there is no significant difference in the distribution of the percentages of students across corresponding Proficiency Levels in 2006 and 2009.

Describing Proficiency Levels

Appendix 3 provides the descriptions of the knowledge 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 example items from the 2009 testing at each Proficiency Level.

Sample items from the 2009 test 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 successfully.

At each Proficiency Level, a wide range of items that varied in context, format and difficulty was 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 and a further set have been included in the School Release Materials; 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 3 in the item set ‘Burning foods’ (see Figure 4.3 on page 36) illustrates performance at Level 4 and above. This extended-response item assesses Strand C, and the concepts are from Energy and Force. It assesses students’ ability to give reasons for the low efficiency of energy transfer from burning food to water in terms of heat loss to the environment, in the context of an experiment to measure which types of food give off most heat when burnt.

Students were provided with a table which displayed the change in water temperature when a food substance was burned. The question required students to recognise that some of the energy heated the surrounding air and/or was transferred to light energy.

Students who can complete items of this level of scientific literacy could be expected to draw on their own knowledge or experiences to explain interactions or effects that have been reported in terms of a non-observable property or abstract science concept.

This item is located at 754 on the scientific literacy scale.

Figure 4.3 Item illustrating performance at Proficiency Level 4 and above

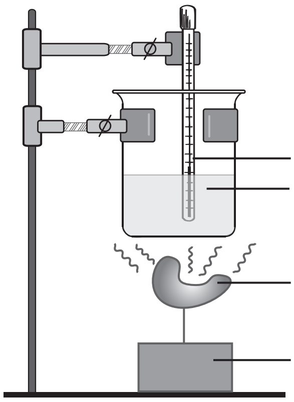
Burning foods

Mr. George compared different foods to see which one gives off the most heat energy when burnt.

He identified samples of four different foods with the same mass. He filled four containers with 200 mL of water and measured the temperature of the water in each container.

Then he

- lit each food sample with a match;
- placed each food sample under a container with water as soon as the food sample was burning;
- measured the temperature of the water in the container after the food sample had burnt completely.



His results are shown in the table below.

Table: Temperature of water before and after burning each food sample

Food sample	Starting temperature of water (°C)	Final temperature of water (°C)
marshmallow	20	35
cashew	20	45
potato chip	20	35
jelly bean	20	40

Q3 Did Mr. George's experiment measure all the energy given off by each food sample?

◯ Yes ◯ No

Give reasons for your answer.

Figure 4.4 contains a second example illustrative of Level 4 and above which assesses Strand B for the concept area Living Things. This extended-response item required students to explain a pattern in data which was presented in tabular form. In order to complete this item successfully, students needed to interpret the information provided and draw a conclusion which summarised the perceived pattern.

This item is located at 686 on the scientific literacy scale.

Figure 4.4 Item illustrating performance at Proficiency Level 4 and above

Tomato plants

A farmer wants to grow taller tomato plants with more tomatoes. He decides to test two methods.

Method 1: buy seeds and add more fertiliser to the soil.

Method 2: save seeds from his tallest tomatoes and plant them the following year.

Results for Method 1

	Average height of plants (cm)	Average weight of plants with tomatoes attached (kg)	Average weight of tomatoes per plant (kg)
Group A (usual amount of fertiliser)	116	8.0	3.5
Group B (double the amount of fertiliser)	116	8.0	2.0

Q1 The average weight of a plant with the tomatoes attached is mostly made up of tomatoes and leaves.

Circle the group of plants which you would expect to have more leaves.

Group A Group B

Use the results for Method 1 to explain your answer.

____

____

____

A third example which is illustrative of Level 4 and above also assesses Strand B. Question 4 in 'Tomato plants' is an extended-response item. This item required students to provide a suggestion for additional data that needed to be collected to determine which method would produce more tomatoes.

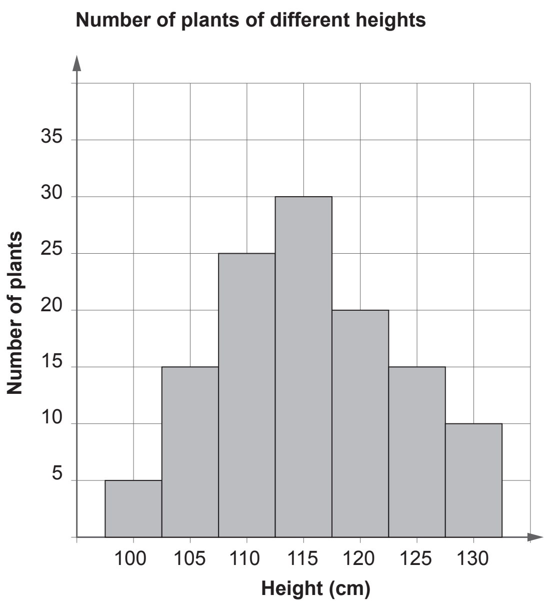
This item is located at 699 on the scientific literacy scale. For this item, students were presented with second-hand data in the form of a table (see Figure 4.4) and a graph (Figure 4.5) that displayed results for two methods used by a farmer to grow more productive tomato plants. Students were required to use the given data in order to suggest additional data that the farmer needed to collect. At a more general level, students who can complete items requiring the same level of scientific literacy as this item and the previous one would be able to draw a conclusion that is consistent with data provided (in tabular, graphical or descriptive form) and that summarises the patterns in the data in the form of a rule.

Figure 4.5 Item illustrating performance at Proficiency Level 4 and above

Results for Method 2

The farmer collected seeds from his tallest tomato plants **(130 cm tall)**.

He planted the seeds and measured the heights of the plants that grew from the seeds.



Look at the column (bar) graph.

Q4 Name one other piece of information that the farmer needs to collect to decide which of the two methods of growing plants is better.

____

____

Sample items illustrating performance at Proficiency Level 3.3

Questions 2 and 3 of 'Cola fountain' (see Figure 4.6) assess Strand A and are illustrative of Level 3.3. This item set drew on the conceptual area of Matter. Question 2 is located at 612 and Question 3 is located at 607 on the scientific literacy scale. For these items, students were presented with information about an experiment designed to measure the height of a fountain produced by adding different numbers of lollies to a bottle of diet cola.

Question 2 required students to identify two things that needed to be kept the same when carrying out the experiment in order to ensure it was a 'fair test'.

For Question 3 students needed to provide an extended response which identified the variable to be changed and the variable to be measured in order to test a hypothesis. Students who responded correctly stated that lollies should be added to bottles of both regular and diet cola and the fountains produced needed to be measured and compared.

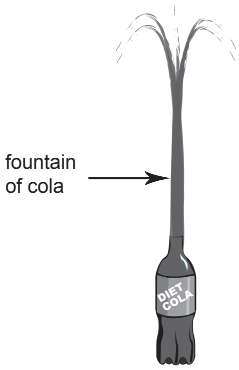
Students who can complete items requiring the same level of scientific literacy would be able to demonstrate an awareness of the principles of conducting an experiment and controlling variables.

Item 1 (shown on page 48) in this set is illustrative of Level 3.1 and is discussed further on page 48.

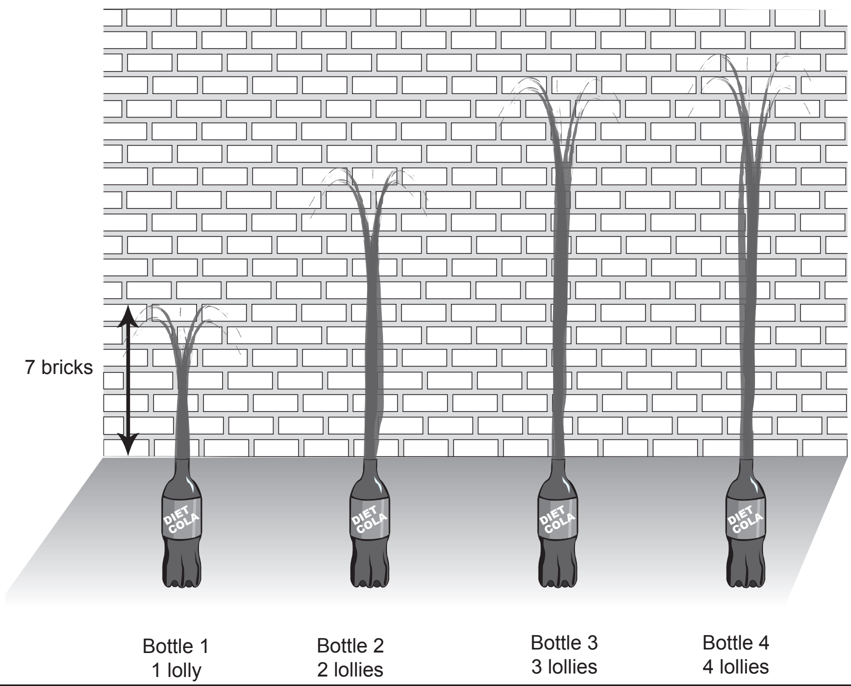
Figure 4.6 Stimulus and items illustrating performance at Proficiency Level 3.3

Cola fountain

Dropping a mint lolly into a bottle of diet cola produces a fountain.



Sam and Michael investigated whether they could make a higher fountain by adding different numbers of lollies to bottles of diet cola. They set up four bottles against a brick wall as shown below. They measured the height of the fountains using the brick wall.



Q2 What should Sam and Michael keep the same in this experiment to make it a fair test?

Tick all possible answers.

- ☐ the number of lollies dropped in each bottle
- ☐ the type of diet cola used
- ☐ the amount of diet cola in each bottle
- ☐ the number of bottles used

Q3 Regular cola contains sugar. Diet cola contains an artificial sweetener.

Sam suggested that this sweetener caused the fountain when the lolly was dropped in the diet cola. Explain in detail how Sam could test his idea.

____

____

____

____

Question 5 from the practical task 'Which beak works best?' is drawn from Strand B and relates to the Living Things concept area. It is also illustrative of Level 3.3.

For this practical task, students worked in groups of two or three to conduct an investigation to determine which type of beak works best for gathering different types of food. Students were provided with two craft sticks to simulate a sieve-type beak, a toothpick to simulate a spear-type beak and a plastic spoon to simulate a net-type beak. 'Foods' were represented by beads and small pieces of paper. Students used each of the 'beaks' in turn to collect each type of 'food' in a 10-second timeframe. Each student recorded the results of each group member for each 'food' in a table. Students then individually responded to a set of items based on their records and observations.

Figure 4.7 Extract from stimulus and items illustrating performance at Proficiency Level 3.3 and Level 4 and above

Part A

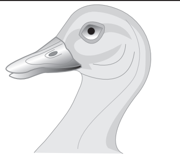
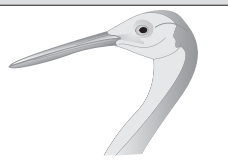
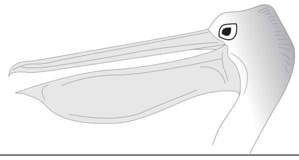
Group work (Work as a group of three students.)

Experiment: **Which beak works best?**

You will investigate which beak type works best to gather different types of food.

The pictures below show three birds, each with a different type of beak: sieve (filter), spear and net. In the experiment, you will use two craft sticks as a sieve-type beak, a toothpick as a spear-type beak and a spoon as a net-type beak.

You will use beads and pieces of paper as food.

Beak type	Example of bird with this beak type	Photo of bird
Sieve (two craft sticks)	**Duck** 	
Spear (toothpick)	**Heron** 	
Net (plastic spoon)	**Pelican** 	

Setting up the experiment

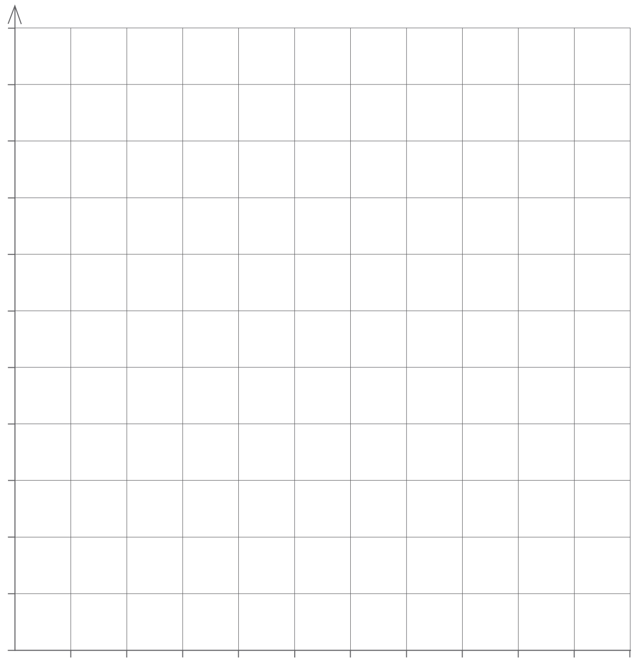
1. Place the plastic plate on a flat surface.
2. Place the two craft sticks, the toothpick and the plastic spoon next to the plate.
3. Place the plastic cup next to the plate.

Q5–Q7

Use the **TOTAL number of beads gathered** to draw a column (bar) graph that displays your group's data from **Table 1**.

Draw your graph in the space below. Remember to label the axes of your graph.

Total number of beads gathered by each beak type



Q8 How is a bird's beak matched to what the bird eats?

Give an example to support your answer.

____

____

____

____

Questions 5–7 required students to construct a column or bar graph based on the data gathered by the group for the 'bead' food. Question 5 specifically required students to correctly label the axes of the graph, using appropriate conventions of scientific literacy. It is located at 622 on the scientific literacy scale. In order to be marked as correct, students needed to label the axes using the variables for which data were collected, i.e. the 'amount of food' (or 'total number of beads') and 'beak type' (craft sticks, toothpicks, plastic spoon).

Questions 6 and 7 are also from Strand B and are illustrative of Level 3.2. They are further discussed on page 44.

Question 8 (see previous page) is also from Strand B and is an example of a polytomous item. That is, the item had a score range of 0, 1 or 2. At Score 2 students were required to provide a response which demonstrated their understanding of the relationship between the shape of the beak and the type of food eaten. They also needed to include an example which supported their generalisation (e.g. 'A bird's beak determines what it eats. For example, a pelican's beak is net shaped so that it can trap fish inside'). A response of this calibre was at Level 4 and the scaled score was 733. A response which achieved a Score 1 was one that provided only a generalisation or only an example (e.g. 'Birds with large beaks can eat large foods.' [generalisation only] 'Hawks and owls have sharp beaks for tearing meat.' [example only]). Students who provided responses such as these are demonstrating achievement at Level 3.3.

Question 2 from the item set 'Collecting Ants' (see Figure 4.8) is representative of Strand C. It is located at 606 on the scientific literacy scale. This item required students to identify the differences in air pressure in the insect pooter which caused an ant to be sucked up.

At a more general level, students who complete items requiring the scientific literacy skills and understandings at Level 3.3 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

Questions 6 and 7 from the practical task 'Which beak works best?' (see Figure 4.7) are illustrative of Strand B and the concept area of Living Things. Question 6 is located at 406 on the scientific literacy scale. Question 6 required students to use an appropriate scale for the graph, beginning at 0 on the baseline. Question 7 is located at 462 on the scientific literacy scale. This question required students to accurately plot data, that they had previously gathered and recorded during their investigation, onto the graph.

Students who completed these items correctly could be expected to both plot data on a graph and interpret such data. They can also identify relationships between elements of the data e.g. the variation in the number of beads collected when using different 'beaks'.

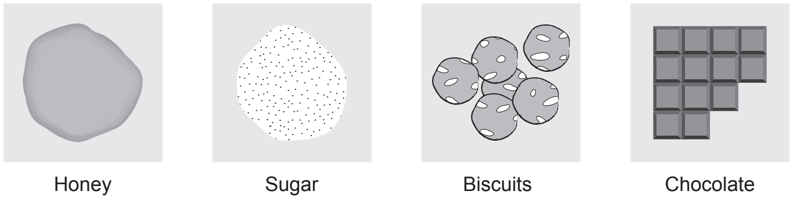
Question 1 in the 'Collecting ants' item set (Figure 4.8) is taken from Strand A and relates to the Living Things concept area. It is an extended-response item and is located at 501 on the scientific literacy scale. This question required students to name two additional things that needed to be 'kept the same' to ensure that the experiment was 'fair'. Students who responded correctly to this question would be expected to understand the need for fair testing when designing an experiment, although they may not necessarily know the term 'controlling variables'.

Figure 4.8 Items illustrating performance at Proficiency Levels 3.2 and 3.3

Collecting ants

Kayla wanted to find out which food would attract the most ants. She set up some traps to collect ants. The traps were pieces of cardboard with different types of food on them.

Kayla made sure all four cardboard pieces were the same size.



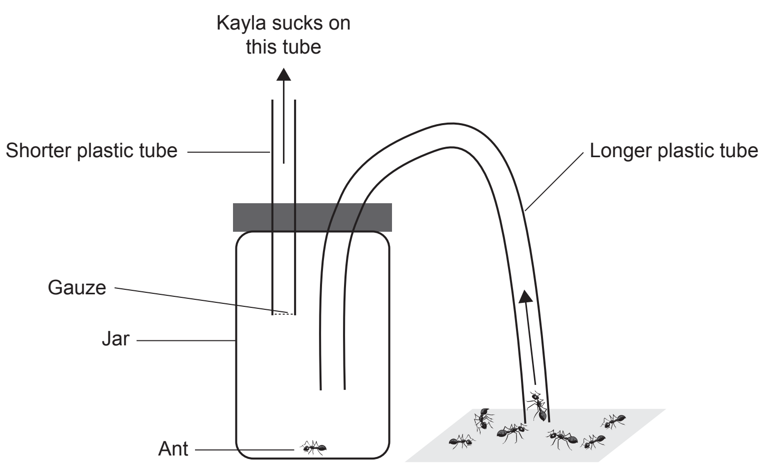
Q1 Name **two** other things that Kayla needs to keep the same to make the experiment fair.

1 ____

2 ____

Q2 Kayla made a device called an insect pooter to collect the ants. She made the pooter by making two holes in the lid of an empty jar and inserting a plastic tube tightly into each hole.

Insect pooter



Kayla placed the longer tube over an ant and sucked on the shorter plastic tube. The ant ended up in the bottom of the jar.

Explain how this happened.

____

____

____

____

Question 1 in the 'Phases of the Moon' item set assesses Strand C and relates to the concept area Earth and Space. This item is located at 434 on the scientific literacy scale. Students were provided with a series of illustrations depicting the phases of the Moon over a three-week period. Students were then asked to identify the pattern of change and draw in a missing shape.

Question 2 in this item set is illustrative of Level 4 and above. It assesses Strand C and relates to the Earth and Space concept area.

Figure 4.9 Items illustrating performance at Proficiency Level 3.2 and Level 4 and above

Phases of the Moon

Sally recorded the shape of the Moon on different days in April. On some days, however, it was cloudy and she could not see the Moon. The table shows her results.

Table: Shape of the Moon on different days

Crescent		Quarter Moon	Full Moon	Quarter Moon	
					
April 4	**April 6**	**April 8**	**April 15**	**April 22**	**April 25**

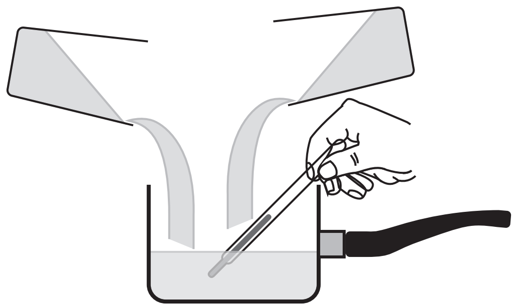
Q1 In the table above, draw what Sally would have seen on April 6 and April 25 if it had not been cloudy.

Q2 The Moon takes about 28 days to orbit Earth. What did the Moon look like on May 8 of the same year? Give reasons for your answer.

Another example of a Strand C item at this level is Question 3 from ‘Heating and cooling’ (Figure 4.10). The item set is drawn from the concept area Energy and Force. It is a multiple-choice question and is located at 455 on the scientific literacy scale. Students were given the temperature of two glasses of water and asked to predict the temperature when the water was mixed.

Figure 4.10 Item illustrating performance at Proficiency Level 3.2

Q3 John has a glass of water at a temperature of 60 °C and a glass with the same amount of water at 20 °C. He pours the water from both glasses into a pot and measures the temperature of the mixed water immediately.



What is the most likely temperature of the mixed water?

- ◯ 20 °C
- ◯ 40 °C
- ◯ 60 °C
- ◯ 80 °C

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 1 (Figure 4.11) from the item set 'Cola fountain' (see Figure 4.6) is illustrative of Level 3.1. It is drawn from Strand A and relates to the concept area Matter. This item is located at 322 on the scientific literacy scale. Students were required to record simple non-standard measurements of the various fountain heights in a table.

Figure 4.11 Item illustrating performance at Proficiency Level 3.1

Q1 Sam measured the height of each fountain using the bricks on the wall.
Complete the table below to record the results of this experiment.

Table: Height of fountain compared to number of lollies

	Number of lollies	Height of fountain (bricks)
Bottle 1	1	7
Bottle 2	2	
Bottle 3	3	
Bottle 4	4	

Question 2 in the item set 'Greenhouse gas emissions' (Figure 4.12) is indicative of Strand B and the concept area Earth and Space. It is located at 388 on the scientific literacy scale. It required students to use information available in a table to determine a plausible reason for Household 2 producing less carbon dioxide than Household 1. In general students who can answer questions requiring the same level of scientific literacy as this item could be expected to formulate simple scientific questions for testing and make predictions. They could also be expected to make simple measurements and to record data as a table, diagram or description.

Question 1 in this set is a further example of an item located at Proficiency Level 3.2.

Figure 4.12 Items illustrating performance at Proficiency Levels 3.1 and 3.2

Greenhouse gas emissions

The table below shows the amount of carbon dioxide produced per year for two different households in New South Wales. Both households have four people. The carbon dioxide is created when electricity is produced to power electrical devices in these households.

Table: Amount of carbon dioxide produced per year for two households

	Amount of carbon dioxide (tonnes/year)	
Activities	**Household 1**	**Household 2**
Heating and cooling the home	2.8	0.0
Cooking	0.4	0.6
Water heating	2.4	2.2
Lighting	0.6	0.8
Using kitchen appliances	1.6	2.0
Leaving kitchen appliances in standby mode	0.5	0.3
Total	8.3	5.9

Q1 Household 1 decided to switch off the kitchen appliances instead of leaving them in standby mode. By how much would they reduce their carbon dioxide emissions?

- ◯ 0.3 tonnes/year
- ◯ 0.5 tonnes/year
- ◯ 1.6 tonnes/year
- ◯ 2.0 tonnes/year

Q2 Household 2 produces less carbon dioxide than Household 1.

Which of the following reasons **best** explains this?

- ◯ Household 2 uses more energy efficient light bulbs compared to Household 1.
- ◯ Household 2 uses fewer kitchen appliances than Household 1.
- ◯ Household 2 takes shorter showers requiring less water to be heated.
- ◯ Household 2 does not use air conditioning or heating devices.

Question 1 from the item set 'Native and introduced animals' (Figure 4.13) is representative of Strand C and relates to the Living Things concept area. It is located at 276 on the scientific literacy scale.

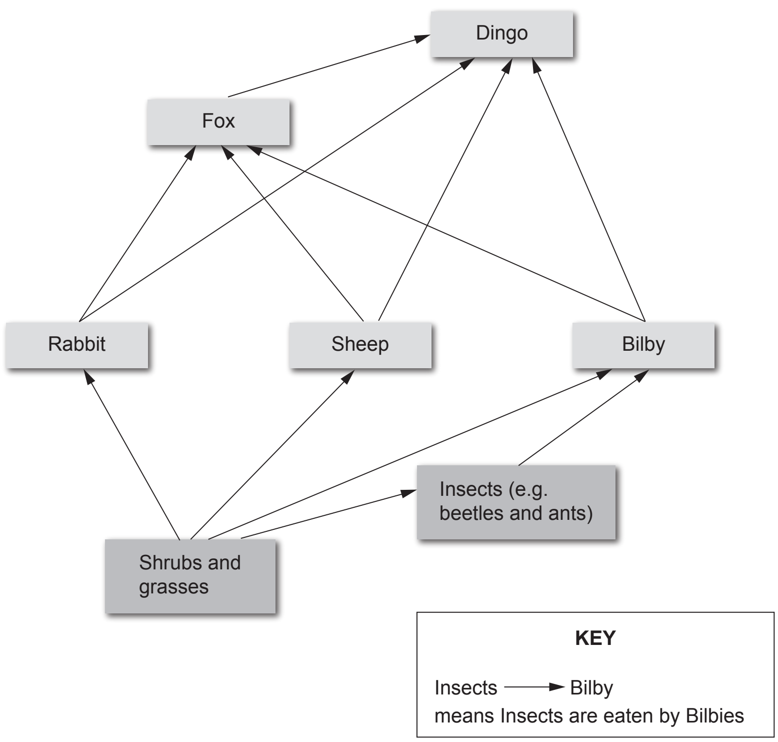
This item required students to interpret information provided in the form of a simple food web in order to draw a conclusion. Students at this level could be expected to describe cause and effect relationships between individual events in a familiar context.

Figure 4.13 Item illustrating performance at Proficiency Level 3.1

Native and introduced animals

Native animals occur naturally in Australia. Introduced animals have been brought into Australia from other countries.

The food web below contains four introduced animals: rabbits, sheep, foxes and dingoes.



Q1 For this food web, which of the following statements is correct?

- ◯ Bilbies eat foxes.
- ◯ Dingoes eat rabbits.
- ◯ Dingoes eat insects.
- ◯ Foxes eat shrubs and grasses.

Two other items which also addressed Strand C at this level are Questions 2 and 3 from the item set ‘Separating mixtures’ (Figure 4.14) which drew on the Matter concept area. Question 2 is located at 356 and Question 3 is located at 389 on the scientific literacy scale.

These items required students to select the appropriate method for separating mixtures based on their knowledge of the basic properties of materials, i.e. sand and iron filings and mud in water. At a more general level, students who responded correctly to these items, or those requiring the same level of scientific literacy, could be expected to apply known rules to explain specific instances related to personal experience.

Figure 4.14 Items illustrating performance at Proficiency Level 3.1

Three methods for separating mixtures into their parts are:

Dissolving | Using a magnet | Filtering

Which method should Jan use to separate the mixtures listed below into their parts?

Q2 Separate sand and iron filings (tiny pieces of iron):

____

Q3 Separate mud in water from water:

____

Sample items illustrating performance at Proficiency Level 2 and below

Question 1 (Figure 4.15) from the practical task ‘Which beak works best?’ (see Figure 4.7) is a Strand B item and relates to the concept area Living Things. It is located at 173 on the scientific literacy scale. Students needed to locate a piece of data in a table they had previously constructed during their group investigation.

Figure 4.15 Item illustrating performance at Proficiency Level 2 and below

Q1 Look at your group’s results in **Table 1:** *‘Food’ (number of beads) gathered*.

When Person 1 used the **craft sticks**, how many beads did they gather?

____

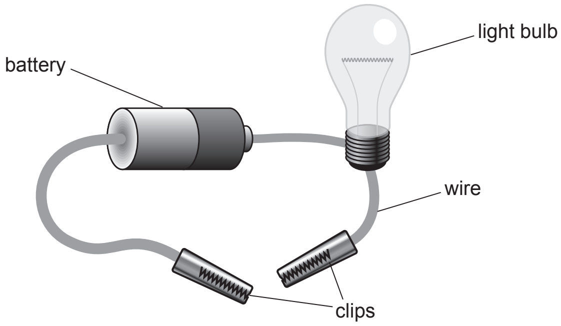
Question 1 from the item set 'Energy efficient light bulbs' (Figure 4.16) is another example of an item at Proficiency Level 2 or below of scientific literacy and is illustrative of Strand C. It is drawn from the concept area Energy and Force and is located at 146 on the scientific literacy scale. This multiple-choice item required students to apply their knowledge, based on first-hand concrete experiences, to identify which object would complete an electrical circuit.

Figure 4.16 Item illustrating performance at Proficiency Level 2 and below

Energy efficient light bulbs

The diagram below shows an incomplete electric circuit containing a battery, wires and a light bulb.

A light bulb in an incomplete electric circuit



Q1 Which item could be placed between the clips to complete the circuit and cause the light bulb to glow?

- ○ paper clip
- ○ piece of paper
- ○ straw
- ○ toothpick

Another item which is illustrative of this level of scientific literacy is shown in Figure 4.17. It is an extended-response item from Strand C and relates to the concept area Living Things. It is located at 262 on the scientific literacy scale. This item from the set 'Collecting ants' (see Figure 4.8) required students to use their real-world knowledge or experience to describe the reason for a defensive behaviour (spraying formic acid) of a black ant.

Figure 4.17 Item illustrating performance at Proficiency Level 2 and below

Q3 Black ants can spray formic acid at people or other animals. Formic acid stings the eyes and skin. Why do ants spray formic acid?

____

____

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 and complete simple graphs and tables.

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.

Table 4.1 Summary of results for sample items 2009

Page	Figure	Unit	Question	% correct	Level	Strand	Scaled score
46	4.9	Phases of the Moon	2	2.7	≥ 4	C	877
36	4.3	Burning foods	3	8.2	≥ 4	C	754
42	4.7	Which beak works best?	8	15.1	≥ 4	B	733
38	4.5	Tomato plants	4	11.3	≥ 4	B	699
37	4.4	Tomato plants	1	12.8	≥ 4	B	686
42	4.7	Which beak works best?	5	22.2	3.3	B	622
40	4.6	Cola fountain	2	20.5	3.3	A	612
40	4.6	Cola fountain	3	22.0	3.3	A	607
45	4.8	Collecting ants	2	23.6	3.3	C	606
45	4.8	Collecting ants	1	41.3	3.2	A	501
49	4.12	Greenhouse gas emissions	1	40.0	3.2	B	494
42	4.7	Which beak works best?	7	51.2	3.2	B	462
47	4.10	Heating and cooling	3	48.6	3.2	C	455
46	4.9	Phases of the Moon	1	53.5	3.2	C	434
42	4.7	Which beak works best?	6	63.2	3.2	B	406
51	4.14	Separating mixtures	3	60.2	3.1	C	389
49	4.12	Greenhouse gas emissions	2	58.3	3.1	B	388
51	4.14	Separating mixtures	2	66.8	3.1	C	356
48	4.11	Cola fountain	1	73.6	3.1	A	322
50	4.13	Native and introduced animals	1	77.9	3.1	C	276
52	4.17	Collecting ants	3	83.8	≤ 2	C	262
51	4.15	Which beak works best?	1	92.8	≤ 2	B	173
52	4.16	Energy efficient light bulbs	1	91.4	≤ 2	C	146

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

Page	Figure	Unit	Question	Proficiency Level	Strand	Scaled score	Percentage correct								
							AUST	NSW	VIC	QLD	SA	WA	TAS	NT	ACT
46	4.9	Phases of the Moon	2	≥ 4	C	877	2.7	3.0	3.2	2.9	1.8	3.0	2.0	2.4	2.9
36	4.3	Burning foods	3	≥ 4	C	754	8.2	7.6	6.8	8.6	5.6	10.5	10.0	5.8	10.4
42–43	4.7	Which beak works best?	8	≥ 4	B	733	15.1	15.3	16.3	14.9	13.7	15.8	13.6	12.3	16.9
38	4.5	Tomato plants	4	≥ 4	B	699	11.3	9.7	12.1	10.7	9.8	12.1	12.2	7.6	16.4
37	4.4	Tomato plants	1	≥ 4	B	686	12.8	12.5	14.5	13.0	9.6	12.8	13.2	8.9	17.6
42–43	4.7	Which beak works best?	5	3.3	B	622	22.2	19.2	19.2	31.5	20.4	25.3	18.5	14.4	21.9
40–41	4.6	Cola fountain	2	3.3	A	612	20.5	18.8	18.9	19.3	19.3	25.8	21.5	13.5	24.2
40–41	4.6	Cola fountain	3	3.3	A	607	22.0	25.5	24.1	18.7	19.4	19.3	24.1	15.7	28.0
45	4.8	Collecting ants	2	3.3	C	606	23.6	26.5	22.9	22.9	21.6	22.3	24.2	17.2	30.0
45	4.8	Collecting ants	1	3.2	A	501	41.3	38.9	35.6	43.2	38.0	51.4	43.1	31.5	44.3
49	4.12	Greenhouse gas emissions	1	3.2	B	494	40.0	41.6	41.7	37.9	37.7	38.8	40.4	35.9	45.8
42–43	4.7	Which beak works best?	7	3.2	B	462	51.2	51.9	54.2	53.2	48.6	53.8	47.0	39.3	52.1
47	4.10	Heating and cooling	3	3.2	C	455	48.6	52.3	50.4	42.2	47.2	50.4	45.4	43.6	54.9
46	4.9	Phases of the Moon	1	3.2	C	434	53.5	51.9	55.2	52.4	52.3	54.9	52.6	45.3	60.8
42–43	4.7	Which beak works best?	6	3.2	B	406	63.2	63.3	63.3	69.0	63.1	65.2	57.5	50.0	61.8
51	4.14	Separating mixtures	3	3.1	C	389	60.2	61.8	63.5	58.4	58.4	63.0	57.4	47.9	63.2
49	4.12	Greenhouse gas emissions	2	3.1	B	388	58.3	60.5	61.9	52.6	56.8	57.9	59.4	48.9	66.1
51	4.14	Separating mixtures	2	3.1	C	356	66.8	70.2	69.6	65.3	66.5	67.8	62.2	52.1	70.0
48	4.11	Cola fountain	1	3.1	A	322	73.6	74.5	74.2	71.9	74.3	75.4	71.6	69.9	73.7
50	4.13	Native and introduced animals	1	3.1	C	276	77.9	76.9	77.3	79.6	78.0	76.7	79.6	71.3	82.6
52	4.17	Collecting ants	3	≤ 2	C	262	83.8	84.8	86.1	83.5	84.0	84.5	82.8	71.9	85.4
51	4.15	Which beak works best?	1	≤ 2	B	173	92.8	91.9	93.9	92.1	93.7	92.8	92.8	90.1	94.3
52	4.16	Energy efficient light bulbs	1	≤ 2	C	146	91.4	91.9	90.8	90.6	92.0	91.8	91.3	86.1	94.3

Chapter 5
Distribution of students within Proficiency Levels for 2009 with comparisons to previous cycles

Introduction

In 2003, the National Assessment Program – Science Literacy determined that student achievement would be 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 2009 distributions of students within Proficiency Levels are shown in Table 5.1.

The National Assessment Program – Science Literacy 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, only approximately 9 per cent of students did not reach Proficiency Level 3.

However, in the Northern Territory 31.3 per cent of students did not demonstrate scientific literacy corresponding to Proficiency Level 3.

Table 5.1 2009 percentage of students in Proficiency Levels by state and territory

State/ Territory	Level 2 and below	Level 3.1	Level 3.2	Level 3.3	Level 4 and above
NSW	8.7 (±2.6)	38.3 (±4.2)	43.8 (±3.9)	9.0 (±3.0)	0.2 (±0.3)
VIC	6.8 (±2.3)	38.6 (±3.6)	48.0 (±4.1)	6.5 (±1.7)	0.1 (±0.1)
QLD	10.5 (±2.9)	40.7 (±3.4)	42.7 (±3.3)	6.0 (±1.7)	0.1 (±0.2)
SA	10.5 (±2.7)	43.0 (±4.0)	41.1 (±4.3)	5.3 (±1.6)	0.0 (±0.2)
WA	9.1 (±2.4)	37.6 (±3.5)	46.2 (±3.8)	7.0 (±1.7)	0.1 (±0.2)
TAS	11.2 (±3.8)	39.0 (±4.0)	43.2 (±5.5)	6.5 (±2.2)	0.1 (±0.3)
NT	31.3 (±9.8)	35.1 (±5.4)	29.2 (±6.3)	4.4 (±2.6)	0.0 (±0.2)
ACT	6.3 (±2.1)	32.5 (±4.0)	49.3 (±4.0)	11.8 (±3.1)	0.2 (±0.3)
AUST	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)

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

Table 5.1 shows that most students demonstrated skills that placed them within Proficiency Level 3. Results further indicate that both at the national and jurisdiction levels, the proportion of students working at Proficiency Level 4 or above accounts for only 0.1 – 0.2 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 assessment.

In 2009, approximately 52 per cent of students were found to be operating at or above the proficient standard at the national level.

At the national level the difference between 2009 and 2006 in the number of students achieving at or above the proficient standard is 2.4 per cent, which is not a statistically significant difference. Table 5.2 shows that Western Australia improved its ranking from seventh in 2006 to third in 2009. However, the difference in percentages of Western Australian students achieving at or above the proficient standard in 2006 and 2009 is not statistically significant. The difference between the 2006 and 2009 results for New South Wales, Tasmania and South Australia, where the 2009 ranking is lower compared to that of 2006, are also not statistically significant; nor are those for the Australian Capital Territory, Victoria, Queensland and Northern Territory, whose ranking did not change in 2009.

Table 5.2 Jurisdictions by percentages of students at or above the proficient standard in rank order for 2006 and 2009

Rank by jurisdiction	2006		2009	
	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)
2	VIC	58.3 (±5.0)	VIC	54.6 (±4.6)
3	NSW	57.4 (±4.3)	WA	53.3 (±4.5)
4	TAS	57.4 (±5.5)	NSW	53.0 (±5.0)
5	SA	51.6 (±4.7)	TAS	49.8 (±6.0)
6	QLD	49.2 (±3.8)	QLD	48.8 (±3.8)
7	WA	46.6 (±4.7)	SA	46.5 (±5.0)
8	NT	38.4 (±6.5)	NT	33.6 (±7.5)
	AUST	54.3 (±2.1)	AUST	51.9 (±2.2)

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 2003, 2006 and 2009 is presented in Table 5.3.

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

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)

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

Table 5.3 shows that the distribution of students across Proficiency Levels at the national level remained relatively stable across the three assessments.

This table demonstrates that in 2009 approximately 52 per cent of students were proficient at Level 3.2 and above. For 2006, it was approximately 54 per cent but this difference is not statistically significant. Comparisons between 2003 figures with 2006 and 2009 figures should be interpreted with caution. As noted on pages 20, 27 and 30 in this report, there were important differences between test designs in the 2003 test cycle and the test cycles in 2006 and 2009. The assessments in 2006 and 2009 included wider coverage of the assessment domain and samples were more inclusive of students in remote geographic locations.

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 available, from the 2003 and 2006 assessments in order to allow trends in results for the National Assessment Program – Science Literacy to be investigated. However, as in 2006, 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 2009, 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 accurately 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 the sample composed of male students.

Table 6.1 Mean scores for male and female students by state and territory in 2009

State/Territory	Percentage of Males in sample	Mean score	
		Males	Females
ACT	46.4	412 (±14.1)	418 (±12.5)
NSW	51.1	398 (±14.6)	396 (±11.5)
NT	49.7	329 (±30.8)	323 (±30.9)
QLD	50.8	387 (±9.8)	382 (±11.4)
SA	51.7	381 (±11.8)	380 (±11.2)
TAS	52.3	384 (±15.7)	388 (±16.4)
VIC	49.4	395 (±11.3)	400 (±9.4)
WA	53.0	397 (±10.1)	389 (±12.1)
AUST	50.5	393 (±6.0)	391 (±5.2)

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 males is two points higher than that for females. However, this difference is not statistically significant nor are any other gender comparisons outlined in Table 6.1. The table indicates that females 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 confirms that there were no significant differences in performance. The table also indicates 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 2009

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	8.4 (±3.7)	30.8 (±5.8)	48.4 (±6.3)	12.2 (±4.0)	0.1 (±0.3)	60.7 (±6.9)
	Female	4.2 (±1.9)	34.1 (±5.7)	50.1 (±5.2)	11.3 (±4.1)	0.3 (±0.5)	61.7 (±6.1)
NSW	Male	9.6 (±3.2)	35.9 (±4.7)	44.3 (±4.8)	10.0 (±3.9)	0.3 (±0.5)	54.6 (±5.8)
	Female	7.3 (±2.8)	40.8 (±5.8)	43.5 (±4.9)	8.2 (±2.9)	0.2 (±0.3)	51.9 (±5.9)
NT	Male	31.3 (±10.9)	32.0 (±7.3)	31.6 (±7.4)	5.1 (±3.3)	0.1 (±0.4)	36.8 (±8.0)
	Female	31.3 (±11.0)	38.2 (±7.7)	26.8 (±7.6)	3.7 (±3.2)	0.0 (±0.0)	30.5 (±9.1)
QLD	Male	10.8 (±3.2)	40.0 (±4.4)	42.0 (±4.0)	7.1 (±2.3)	0.1 (±0.3)	49.2 (±4.7)
	Female	10.1 (±3.6)	41.5 (±4.6)	43.4 (±4.9)	4.9 (±2.0)	0.0 (±0.1)	48.4 (±5.2)
SA	Male	11.2 (±3.4)	41.8 (±4.8)	40.9 (±5.1)	6.0 (±2.2)	0.1 (±0.3)	47.0 (±5.6)
	Female	9.8 (±3.2)	44.3 (±5.2)	41.4 (±5.6)	4.6 (±1.9)	0.0 (±0.1)	45.9 (±5.9)
TAS	Male	12.6 (±4.5)	38.9 (±5.1)	40.7 (±7.0)	7.7 (±3.2)	0.2 (±0.5)	48.5 (±7.1)
	Female	9.7 (±4.6)	39.1 (±5.6)	45.9 (±6.2)	5.2 (±2.7)	0.1 (±0.4)	51.2 (±7.2)
VIC	Male	7.7 (±3.1)	39.0 (±4.5)	46.6 (±5.5)	6.6 (±2.1)	0.1 (±0.2)	53.3 (±5.7)
	Female	5.9 (±2.1)	38.1 (±4.6)	49.4 (±4.6)	6.4 (±2.5)	0.0 (±0.1)	55.9 (±5.1)
WA	Male	8.6 (±2.3)	37.1 (±4.1)	46.3 (±4.0)	7.9 (±2.3)	0.1 (±0.3)	54.3 (±4.9)
	Female	9.8 (±3.4)	37.7 (±5.0)	46.2 (±5.4)	6.2 (±2.1)	0.0 (±0.1)	52.5 (±5.7)
AUST	Male	9.7 (±1.5)	38.0 (±2.1)	44.2 (±2.3)	8.0 (±1.5)	0.1 (±0.2)	52.3 (±2.6)
	Female	8.3 (±1.3)	40.0 (±2.4)	45.1 (±2.2)	6.5 (±1.1)	0.1 (±0.1)	51.7 (±2.6)

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

Trend analysis by gender

Table 6.3 shows mean scores achieved by male and female students, at the national level, as observed in the 2003, 2006 and 2009 National Assessment Program – Science Literacy assessments.

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

AUST	Percentage of males in sample	Mean score	
		Males	Females
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)

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

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

As can be seen from Table 6.3, in 2009 males achieved a slightly higher mean than females, as was also the case in the 2003 and 2006 cycles. However, this difference is not statistically significant. Similarly, differences between the 2006 and 2009 results were not statistically significant for both genders.

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 and 2009

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	15.9 (±1.3)	4.1 (±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)	0.01 (±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)

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

Table 6.4 shows that there is no statistically significant difference in the percentage of male and female students achieving various Proficiency Levels between 2006 and 2009.

Table 6.5 shows the distribution of the percentage of males and females at or above the proficient standard in 2003, 2006 and 2009.

Table 6.5 Percentage of male and female students at or above the proficient standard in 2003, 2006 and 2009

AUST	At or above the proficient standard	
	Males	Female
2003	n.a.	n.a.
2006	54.9 (±2.5)	53.7 (±2.3)
2009	52.3 (±2.6)	51.7 (±2.6)

Note: 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

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 geolocations. Indigenous students' results relative to non-Indigenous students' results are shown in Table 6.6.

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

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)

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

In 2009, the mean score for Indigenous students was 297, indicating that they did not perform as well as non-Indigenous students, with a mean score of 397. These statistics are significant, as they also were in 2006 when the mean score for Indigenous students was 311 and the mean score for non-Indigenous students was 402.

The difference between the 2006 and 2009 mean scores is 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 and 2009.

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

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)

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

Table 6.7 shows that in 2009, 38.6 per cent of Indigenous students were working at Level 2 or below, whereas only 7.4 per cent of non-Indigenous students were working at the same level. This percentage is greater than the figure of 31.4 per cent for Indigenous students at this level in 2006, but is not statistically significant owing to the relatively wide confidence intervals being observed in both the 2006 and 2009 data.

Table 6.8 shows the percentage of Indigenous and non-Indigenous students at or above the proficient standard in 2003, 2006 and 2009.

Table 6.8 Percentage of Indigenous and non-Indigenous students achieving at or above the proficient standard in 2003, 2006 and 2009

AUST	At or above the proficient standard	
	Indigenous	Non-Indigenous
2003	n.a.	n.a.
2006	25.5 (±10.0)	54.7 (±2.2)
2009	19.6 (±6.0)	53.9 (±2.3)

Note: 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 between 2003 and 2006.

Table 6.8 also shows that in 2009, 19.6 per cent of Indigenous students performed at or above the proficient standard which represents a decrease of approximately six percentage points compared to 2006. This difference represents a decrease of approximately 20 per cent between 2006 and 2009 but it is not statistically significant.

Geographic location of schools

Table 6.9 shows the distributions of mean scaled scores for students attending schools in differing geographic locations. It shows that differences between the performance of students living in 'Metropolitan areas' and 'Provincial areas' were not statistically significant. However, the three per cent of students living in 'Remote and very remote areas' had significantly lower performances in scientific literacy than students from any other location.

Students attending schools in 'Metropolitan areas' achieved the highest mean scaled scores. Similar results were found in 2006, however the results cannot be compared directly as the categories for the geographic locations changed between 2003, 2006 and 2009.

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

MCEECDYA geographic location category	Percentage of students	Mean score
Metropolitan areas	72.3	395 (±6.2)
Provincial areas	24.7	389 (±7.9)
Remote and very remote areas	3.0	336 (±23.6)
AUST	100.0	392 (±5.1)

Note: figures in parentheses refer to 95 per cent confidence intervals and the percentages 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 and all other tables in this report 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 please refer to the 2009 Technical Report.

Table 6.10 shows the distribution of student results across Proficiency Levels by the geographic location of the sampled schools.

Table 6.10 Percentage distribution across Proficiency Levels by school geographic location in 2006 and 2009

MCEECDYA geographic location category	Level 2 and below		Level 3.1		Level 3.2		Level 3.3		Level 4 and above	
	2006	2009	2006	2009	2006	2009	2006	2009	2006	2009
Metropolitan areas	7.9	8.4 (±1.5)	36.7	38.1 (±2.0)	44.3	45.5 (±2.1)	10.4	7.8 (±1.4)	0.7	0.1 (±0.1)
Provincial areas	8.4	8.6 (±1.7)	38.2	41.9 (±3.4)	45.3	43.5 (±3.3)	7.9	6.0 (±1.5)	0.2	0.1 (±0.1)
Remote and very remote areas	26.5	28.2 (±8.8)	38.0	37.9 (±8.4)	31.1	29.6 (±7.2)	4.3	4.1 (±3.7)	0.1	0.2 (±0.5)
AUST	8.6	9.1 (±1.2)	37.0	39.0 (±1.7)	44.2	44.5 (±1.8)	9.6	7.2 (±1.1)	0.5	0.1 (±0.1)

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 2006 and 2009 by geographic location. While the percentages are available for 2006, these figures should be treated with caution because the standard errors are not available.

Table 6.11 Percentage of students achieving at or above the proficient standard in 2006 and 2009 by geographic location

AUST	At or above the proficient standard		
	Metropolitan areas	Provincial areas	Remote and very remote areas
2006	55.4	53.4	35.5
2009	53.4 (±2.6)	49.5 (±4.1)	33.9 (±8.2)

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

Using the figures shown in Table 6.11, it can be determined that a significant number of students, approximately 66 per cent, who attend schools located in 'Remote and very remote areas' did not succeed in reaching the proficient standard in scientific literacy. On the other hand, there seems to be only a small difference in percentage of students performing at or above the proficient standard between students attending schools in 'Metropolitan areas' and those attending schools in 'Provincial areas'.

Language background

In 2006, an online system for collecting demographic information about students participating in the National Assessment Program – Science Literacy was implemented. However, the system did not deliver accurate and complete information so only 2003 and 2009 data are presented here.

In 2009, data were collected to understand the language backgrounds of students in Year 6. 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 and 2009. 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 score by student language background in 2003 and 2009

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

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

In 2009, students from language backgrounds other than English, with a mean score of 384, did not perform as well as students from English speaking backgrounds, with a mean score of 396. However, this difference is not statistically significant.

The distribution of students across the Proficiency Levels who have a language background other than English (LBOTE) and students with an English speaking background (ESB) is given in Table 6.13.

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

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)

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

Table 6.13 shows that 12.4 per cent of LBOTE students were working at Level 2 or below, whereas only 7.7 per cent of ESB students were working at the same level. Nine per cent of LBOTE students achieved Proficiency Level 3.3 or above, compared to 7 per cent of ESB students.

Table 6.14 Percentage of students achieving at or above the proficient standard in 2009 by student language background

AUST	At or above the proficient standard	
	LBOTE	ESB
2009	48.9 (±4.9)	53.4 (±2.3)

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.

Table 6.14 shows the percentage of students achieving at or above the proficient standard in 2009 by language background. It shows that the differences between the percentages of LBOTE and ESB students performing at or above the proficient standard were minimal.

Chapter 7
Student Survey

Introduction

An innovation in the 2009 assessment was the implementation of a Student Survey. The survey was conducted following completion of the practical task. For the National Assessment Program – Science Literacy trial, a pool of 46 items were developed in consultation with the Science Literacy Review Committee (SLRC). Following analysis of the responses from the trial and feedback from the SLRC, 30 items were selected for inclusion in the final survey form. This survey required students to provide responses which varied from simple Yes/No responses to others which provided three or four options in a rating scale.

The final survey was divided into three categories:

1. Students' perceptions of and attitudes towards science.

2. Students' interests in science beyond the classroom.

3. Students' experiences of science at school.

This chapter also includes results of the statistical analysis that was conducted for the purpose of examining the relationship between students' responses to specific Student Survey items or questions and their achievements in the 2009 National Assessment Program – Science Literacy.

Distribution of students' responses to the Student Survey

The Student Survey questions can be divided into three categories:

1. Students' perceptions of and attitudes towards science.

This section included ten statements which sought to elicit what students considered to be 'science' and their attitudes towards science, using a four-point Likert scale.

2. Students' interests in science beyond the classroom.

There were two statements in this section which were intended to gather information about the frequency with which students watched television programs or DVDs about science topics at home or read books and newspaper or magazine articles about science topics.

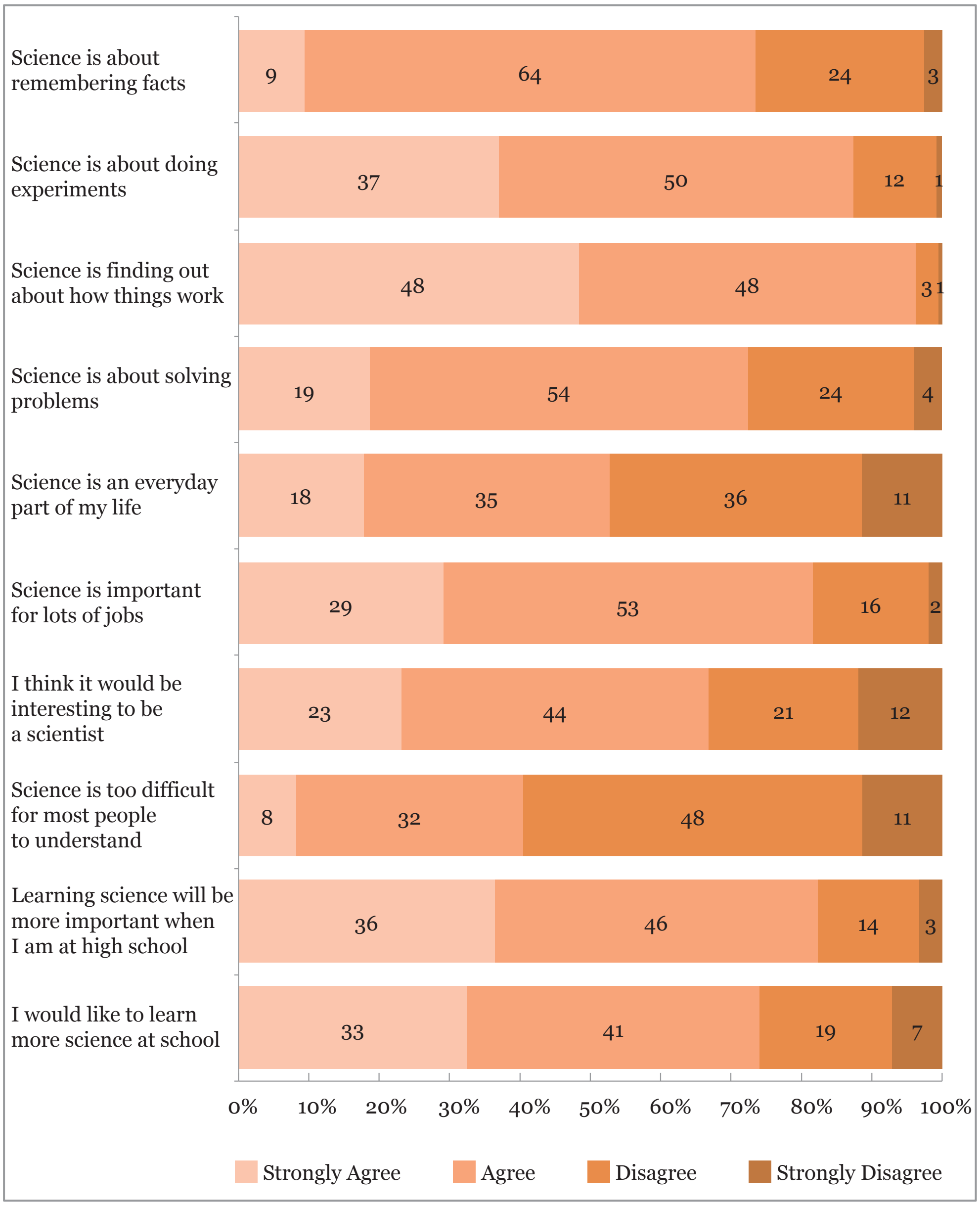
3. Students' experiences of science at school.

This section asked several questions, including the frequency of science teaching at the student's school and the time of the day when science lessons occurred. Students were also asked about the science concept areas or topics they had studied. In addition, students were asked whether science was taught by their classroom teacher and whether extra-curricula science-related activities were available at the school. The final part of this section included ten statements which sought to draw out information about the pedagogy of science teaching in each student's classroom.

Results of the survey are summarised in the following Figures. 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 beneath each Figure.

Students' perceptions of and attitudes towards science

Figure 7.1 Students' perceptions of and attitudes towards science

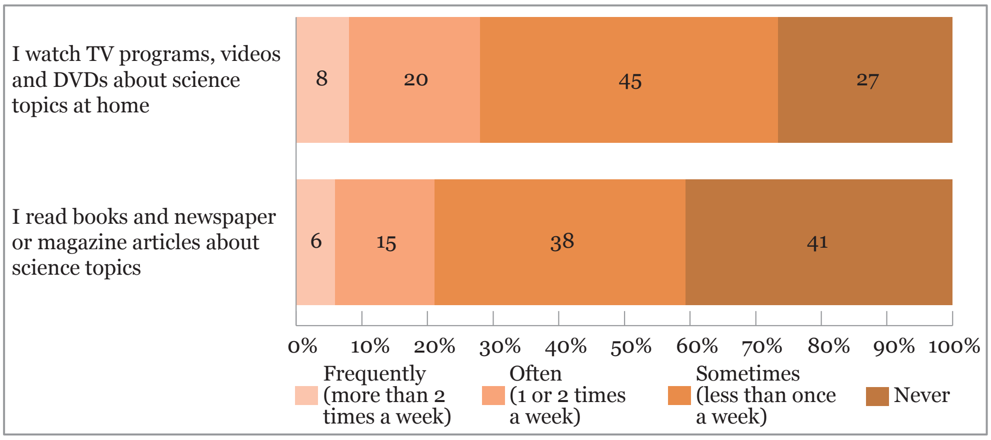


In this and following Figures percentages have been rounded and may not add to 100.

Figure 7.1 demonstrates that while the great majority of students either agreed or strongly agreed with the first four statements about science, only 53 per cent saw science as being part of their lives. In response to the statement that 'Learning science will be more important when I am at high school', 82 per cent indicated this to be the case and 74 per cent responded that they would like to 'learn more science at school'.

Students' interests in science beyond the classroom

Figure 7.2 Students' interests in science beyond the classroom



Using the prompts 'I watch TV programs, videos and DVDs about science topics at home' and 'I read books and newspaper or magazine articles about science topics', students were asked to indicate some of the ways that they learned about science beyond the school, using different mediums.

Figure 7.2 shows that when accessing audiovisual mediums, 73 per cent of the students responded that they, at least, 'sometimes' watched science programs at home, with 28 per cent of the students doing so 'often' or 'frequently'.

When using a printed medium at least 59 per cent indicated that they read about science topics, with at least 21 per cent doing so 'often' or 'frequently' whereas 41 per cent 'never' read science-related materials at home.

Students' experiences of science at school

Figure 7.3 Frequency of science lessons

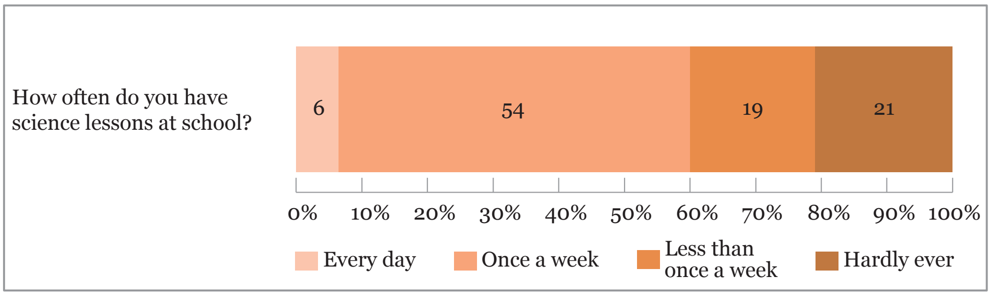


Figure 7.3 indicates that the majority of students reported to have at least one science lesson each week with 6 per cent indicating that they have a science lesson every day, while 21 per cent reported to 'hardly ever' have a science lesson.

Figure 7.4 Time of day when science is taught

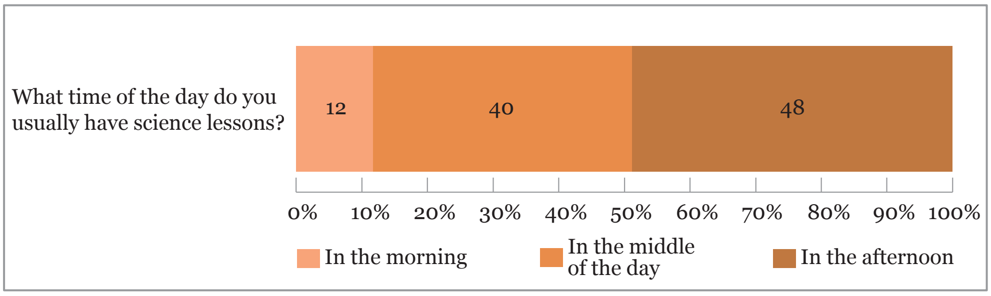


Figure 7.4 demonstrates that in this survey only 12 per cent of students reported that their science lessons occurred in the mornings, when typically the students are most alert, while 48 per cent of them reported having afternoon science lessons.

Figure 7.5 Role of science in class and school planning

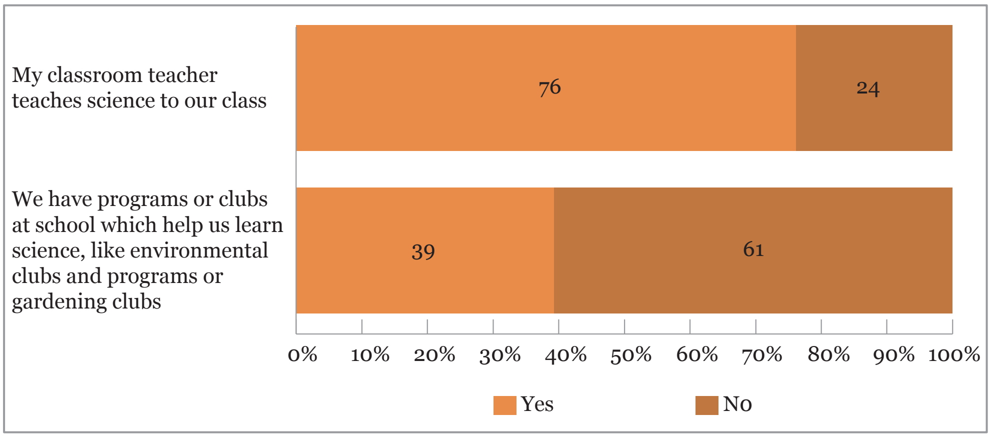


Figure 7.5 shows that the regular classroom teacher is the person who teaches science for 76 per cent of the students and only 39 per cent of the students have access to other programs or clubs where they can engage in science-related activities within the school.

Figure 7.6 Students' experiences of science at school

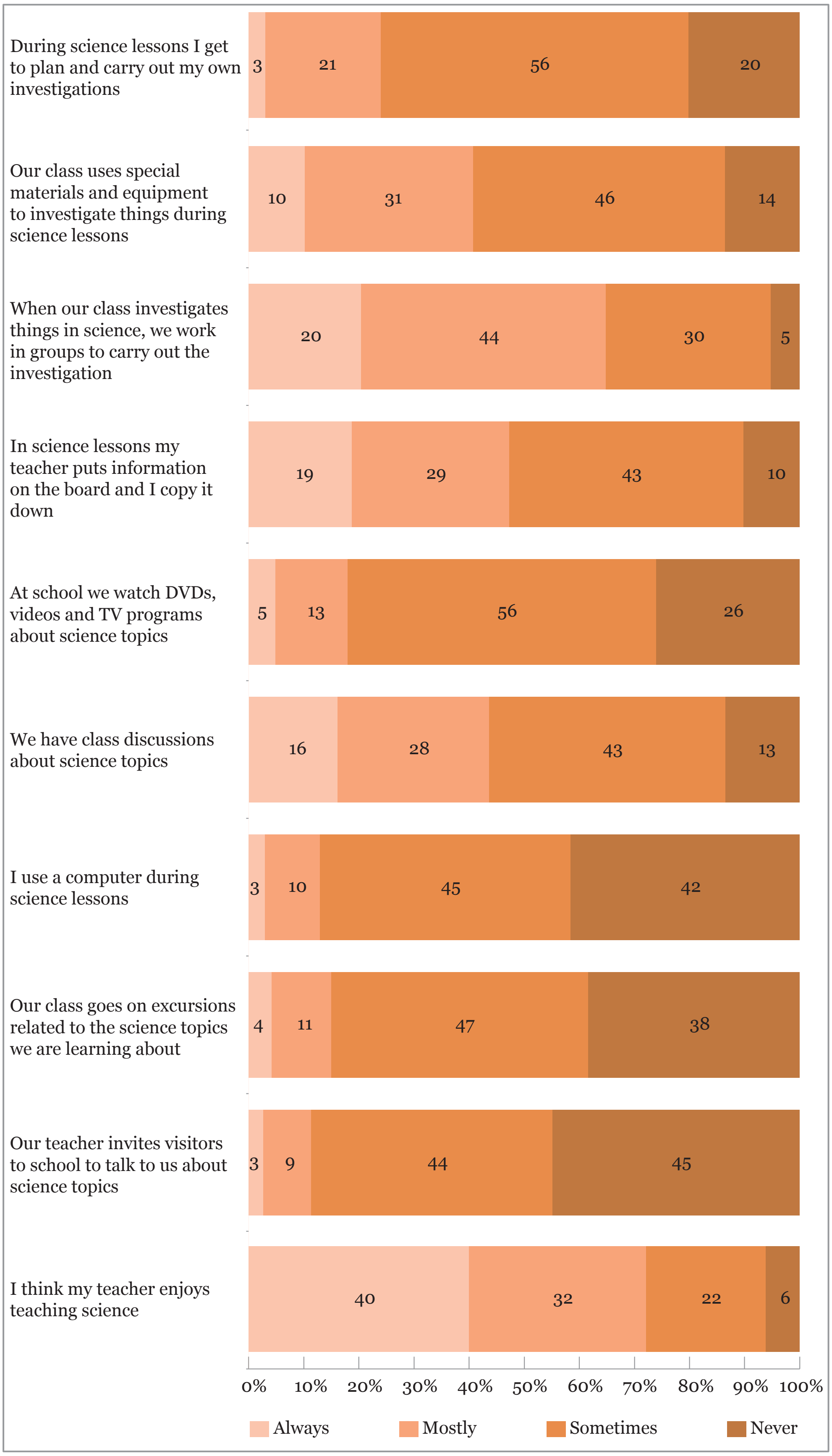
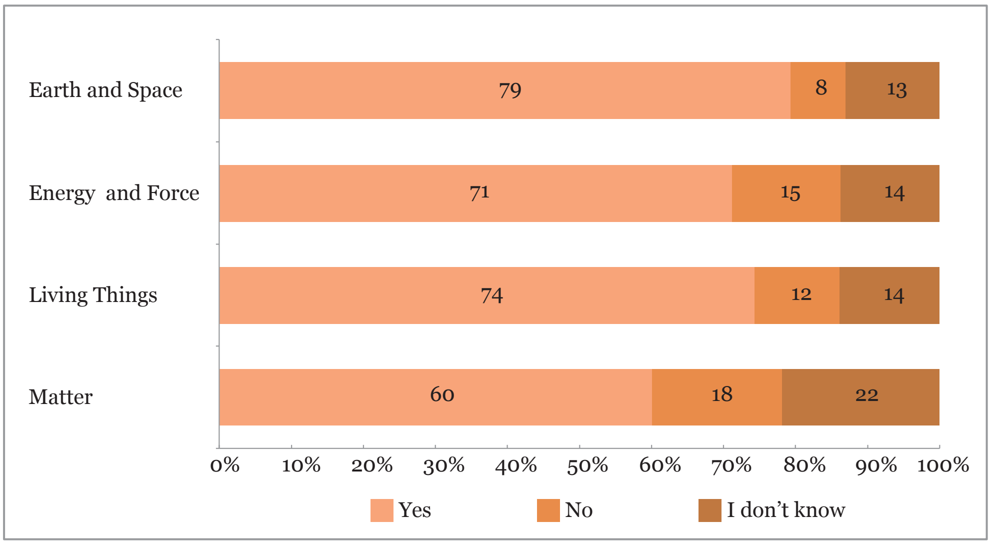


Figure 7.6 shows that while 24 per cent of students reported that they ‘always’ or ‘mostly’ carried out their own self-directed investigations in science, 64 per cent reported that they ‘always’ or ‘mostly’ worked in groups to carry out investigations. However, only 13 per cent reported that they used computers during science lessons.

Figure 7.7 Science concept areas studied at school



Students were provided with prompts to assist them in recognising the broad science conceptual areas that may have been covered.

For the concept area Earth and Space, students were prompted by the provision of 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)’.

These examples were based on advice from the SLRC regarding the most commonly studied topics in primary schools.

As Figure 7.7 shows, the most commonly studied science concept area is ‘Earth and Space’ and ‘Matter’ is the least commonly studied concept area.

Relationship between Student Survey responses and scientific literacy

To examine whether there are systematic commonalities between students' responses to the Student Survey, the students' responses were subjected to factor analysis. Factor analysis is a statistical technique for identifying latent commonalities, or latent factors, in a data file that contains answers to numerous questions or observable factors.

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 mutually correlate. Factor analysis showed that there are only weak commonalities in the Student Survey, with three extracted factors explaining only approximately 25 per cent of the variability in students' responses. Items that form the three 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	I would like to learn more science at school. I think it would be interesting to be a scientist. I watch TV programs, videos and DVDs about science topics at home. I read books and newspaper or magazine articles about science topics. Science is an everyday part of my life. Science is important for lots of jobs.	13.4	13.4
Second	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. At school we watch DVDs, videos and TV programs about science topics. I use a computer during science lessons.	7.0	20.4
Third	Our class uses special materials and equipment to investigate things during science lessons. When our class investigates things in science, we work in groups to carry out the investigation. I think my teacher enjoys teaching science. During science lessons I get to plan and carry out my own investigations.	5.0	25.4

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 the order of their contribution to the amount of explained variance in the Student Survey. Further details can be found in the 2009 Technical Report.

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 survey and students' achievement in the National Assessment Program – Science

Literacy, a regression analysis was conducted. The regression analysis used items from the survey as independent variables and students' 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 15 per cent of the variability in students' scores in scientific literacy could be predicted based on their response to items in the survey. Furthermore, only one item demonstrated a correlation of meaningful magnitude, with a regression coefficient of 0.2 (SE = 0.004). This was Item 8: 'Science is too difficult for most people to understand'. The regression analysis showed that an increase in scientific literacy achievement is followed by an increase in students' disagreement with this statement.

Such a result has rendered any further use of regression analysis in explaining students' achievement in the National Assessment Program – Science Literacy as uninformative. However, in order to provide an illustrative overview of the relationship between students' responses to survey items and students' 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.

Item 8 (below) is included in this set because it is the only item that showed a correlation with students' achievement in scientific literacy. Items 10, 28 and 22 (over leaf) are included as illustrative of the three latent factors extracted in the factor analysis. Finally, Item 13 is presented here because it shows the relationship between students' achievements in scientific literacy and the frequency of science teaching that students in the sample reported they typically receive.

Table 7.2 shows the distribution of students' responses across Proficiency Levels and response categories for Item 8. As can be seen, the proportion of students preforming at or above the proficient standard (i.e. achieving Level 3.2 and above) increases significantly when students disagreed with the statement: 'Science is too difficult for most people to understand'.

Table 7.2 Responses by Proficiency Levels to Item 8

Item 8 Science is too difficult for most people to understand.	**Level 2**	**Level 3.1**	**Level 3.2**	**Level 3.3**	**Level 4**
Strongly agree	19.1 (±4.5)	48.9 (±4.9)	29.4 (±4.4)	2.6 (±1.3)	0.0 (±0.0)
Agree	12.1 (±1.8)	46.0 (±2.3)	38.0 (±2.5)	3.8 (±1.1)	0.0 (±0.1)
Disagree	5.2 (±1.0)	34.4 (±2.1)	50.8 (±2.1)	9.4 (±1.6)	0.2 (±0.2)
Strongly disagree	6.1 (±2.0)	30.4 (±3.7)	51.4 (±4.2)	12.0 (±2.6)	0.2 (±0.4)
AUST	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)

Table 7.3 shows the distribution of students' responses across Proficiency Levels and response categories for survey Item 10. This item had the strongest relationship with the first latent factor extracted in the factor analysis.

Table 7.3 Responses by Proficiency Levels to Item 10

Item 10 I would like to learn more science at school.	**Level 2**	**Level 3.1**	**Level 3.2**	**Level 3.3**	**Level 4**
Strongly agree	8.9 (±1.8)	37.6 (±2.3)	45.8 (±2.5)	7.6 (±1.7)	0.1 (±0.2)
Agree	8.6 (±1.4)	38.9 (±2.4)	44.7 (±2.3)	7.6 (±1.4)	0.1 (±0.2)
Disagree	7.1 (±1.5)	38.9 (±3.4)	46.4 (±3.3)	7.5 (±1.9)	0.1 (±0.3)
Strongly disagree	13.4 (±3.8)	46.3 (±5.8)	36.9 (±5.5)	3.4 (±1.7)	0.0 (±0.1)
AUST	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)

As Table 7.3 shows, the proportion of students who expressed strong disagreement with the statement: 'I would like to learn more science at school' and whose achievements were below the proficient standard (i.e. achieving Level 3.1 and below) is higher than the proportion of students in the other three response categories.

Table 7.4 shows the distribution of students' responses across proficiency levels and response categories for survey Item 28. This item had the strongest relationship with the second latent factor extracted in the factor analysis.

Table 7.4 Responses by Proficiency Levels to Item 28

Item 28 Our class goes on excursions related to the science topics we are learning about.	**Level 2**	**Level 3.1**	**Level 3.2**	**Level 3.3**	**Level 4**
Always	16.1 (±5.5)	42.5 (±6.6)	36.8 (±6.0)	4.5 (±2.6)	0.0 (±0.3)
Mostly	11.0 (±2.6)	38.5 (±4.0)	43.9 (±4.5)	6.6 (±2.2)	0.0 (±0.2)
Sometimes	7.2 (±1.2)	37.5 (±2.3)	47.0 (±2.2)	8.2 (±1.4)	0.1 (±0.2)
Never	8.8 (±1.6)	40.5 (±2.8)	43.6 (±2.8)	6.9 (±1.7)	0.1 (±0.2)
AUST	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)

Table 7.4 indicates that opportunities to engage in excursions related to the science topics being studied at school has little impact on students' achievements. This may be more related to the number of excursions that a class experiences in a year than their subject relevance.

Table 7.5 shows the distribution of students' responses across Proficiency Levels and response categories for survey Item 22. This item had the strongest relationship with the third latent factor.

Table 7.5 Responses by Proficiency Levels to Item 22

Item 22 Our class uses special materials and equipment to investigate things during science lessons.	Level 2	Level 3.1	Level 3.2	Level 3.3	Level 4
Always	10.3 (±2.8)	38.7 (±4.1)	45.5 (±4.1)	5.5 (±2.1)	0.0 (±0.1)
Mostly	8.4 (±1.7)	38.3 (±2.7)	45.1 (±2.6)	8.1 (±1.7)	0.1 (±0.2)
Sometimes	8.1 (±1.3)	38.5 (±2.4)	45.2 (±2.3)	8.0 (±1.4)	0.2 (±0.2)
Never	10.0 (±2.5)	42.4 (±3.7)	42.9 (±4.1)	4.7 (±1.9)	0.0 (±0.2)
AUST	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)

The data in Table 7.5 indicate that using special materials and equipment during science lessons has no impact on students' achievements in scientific literacy as measured by the National Assessment Program – Science Literacy.

Table 7.6 shows the distribution of students' responses across Proficiency Levels and response categories for survey Item 13.

Table 7.6 Responses by Proficiency Levels to Item 13

Item 13 How often do you have science lessons at school?	Level 2	Level 3.1	Level 3.2	Level 3.3	Level 4
Every day	11.2 (±3.9)	39.9 (±7.0)	41.4 (±6.3)	7.5 (±3.2)	0.0 (±0.2)
Once a week	7.9 (±1.3)	37.3 (±2.2)	46.0 (±2.1)	8.7 (±1.6)	0.2 (±0.2)
Less than once a week	9.0 (±2.1)	39.0 (±3.4)	45.5 (±3.8)	6.5 (±1.6)	0.1 (±0.2)
Hardly ever	10.7 (±2.7)	43.0 (±3.3)	42.1 (±3.7)	4.2 (±1.4)	0.0 (±0.0)
AUST	9.1 (±1.2)	39.0 (±1.7)	44.5 (±1.8)	7.2 (±1.1)	0.1 (±0.1)

Item 13 asked students to report on the frequency of science lessons in their class room. Table 7.6 indicates that students who received at least one science lesson each week, had slightly higher achievements than those who reported having 'less than one' or 'hardly ever' having a science lesson. However, the differences are not significant and may also indicate that students are not always aware that what they are learning is 'science', particularly when much primary teaching is done in integrated cross-curriculum units.

Conclusion

It is difficult to draw any reliable conclusions from the survey data collected. This could perhaps be explained by several factors. It may be that students, regardless of their level of achievement, provided answers that they thought would please their teachers. It may also be that they did not fully understand what was being asked by each item and therefore chose a response at random. Such a result does indicate the need for more detailed investigation into students' engagement with questionnaires accompanying the main assessments in National Assessment Programs. Nevertheless, the Student Survey in its first administration in the National Assessment Program – Science Literacy provides some interesting insights into students' perspectives of science. It could, therefore, be used as a basis for designing a Student Survey to be administered in the 2012 cycle.

Chapter 8
Conclusion

The 2009 cycle of the National Assessment Program – Science Literacy provided the first opportunity to report on the progress of Year 6 students in scientific literacy over a six year period (2003 – 2009). One of the main objectives of the National Assessment Program – Science Literacy 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 the sample frame being 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 and the 2003 results were re-scaled onto it.

The Assessment Domain for scientific literacy and science concept areas have remained stable over that period. Secure items from the 2003 and 2006 cycles were retained and new items were developed for the 2009 cycle.

While the number of items assessed and test booklets increased between 2003 and 2006, it remained similar for 2009. In 2003 students were assessed using two test booklet forms. In 2006 and 2009, seven test booklet forms were implemented allowing for a rotational 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 location within a test booklet.

Student achievement in scientific literacy 2003 – 2009

In 2009, approximately 52 per cent of students, at the national level, attained the proficient standard or better in scientific literacy. In 2006, the percentage was approximately 54 per cent and in 2003 it was approximately 59 per cent. However, it should be noted that the 2006 and 2009 samples were more inclusive of students in remote geographic locations. Chapter 3 of this report contains detailed information about the performance of students nationally and at a state or territory level in the National Assessment Program – Science Literacy. 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. Chapter 5 provides detailed breakdowns of student performance across the Proficiency Levels at a state and territory level including trend data.

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 boys slightly outperformed girls, although the difference was not statistically significant. This was the case in the two previous cycles.

Nationally non-Indigenous students achieved significantly higher levels of scientific literacy than Indigenous students, as was the case in 2006 and 2003. 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.

While students living in ‘Metropolitan areas’ achieved the highest mean score in scientific literacy, their results were not significantly different from those living in ‘Provincial areas’. However, students living in ‘Remote and very remote areas’ recorded results which were significantly lower. Similar findings are evident in the National Assessment Programs in literacy, numeracy, civics and citizenship and information and communication technologies.

Students from English-speaking backgrounds also achieved slightly higher means nationally than students from language backgrounds other than English. However, the difference is not statistically significant.

Student Survey

As discussed in Chapter 7, a Student Survey was administered for the first time as part of the 2009 assessment. While the survey provided interesting insights into students' perceptions of and attitudes to science and their experiences with science learning at school, there were no reliable correlations between student performance in scientific literacy and particular survey responses. The responses to the Student Survey will guide further survey development for the 2012 cycle and 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.

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.

Copyright Acknowledgements

We would like to thank the following copyright holders who have granted permission to use the texts and graphics cited below.

Item	Page	Acknowledgement
Chestnut Teal	42	© Wendy Opie/Viridans Images

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 the 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).

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 in accord 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 the National Assessment Program – Science Literacy, as the sample testing methodology to be used 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 have been 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 communications technologies.

Strand B: Interpreting evidence and drawing conclusions from their 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 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.

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 – July 2004 version from DEST Science Education Assessment Resource (SEAR) project

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 their 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

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 the National Assessment Program – Science Literacy

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 National Assessment Program – Science Literacy test instruments are constructed within the constraints of test length, it will not be feasible to include all the listed concepts in instruments constructed for a specific testing cycle.

Table A1.2 Major scientific concepts in the National Assessment Program – Science Literacy 2009

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, reproductions 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 National Assessment Program – Science Literacy consisted of all students enrolled in Year 6 in Australian schools in 2009.

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

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 National Assessment Program – Science Literacy 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 location.

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.

Table A2.1 Estimated 2009 Year 6 enrolment figures as provided by Department of Education, Employment and Workplace Relations (DEEWR)

State/ Territory	Institutions	Students	Percentage of students
ACT	103	4501	1.7
NSW	2338	87 112	32.1
NT	150	3005	1.1
QLD	1379	56 879	21.0
SA	615	19 245	7.1
TAS	227	6756	2.5
VIC	1813	65 573	24.2
WA	883	28 017	10.3
AUST	7508	271 088	100.0

In this and the following tables percentages have been rounded and may not add to 100.

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 enrolment of less than 12 (TCS/2) 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	74	71.8	4082	90.7
	Moderately small	18	17.5	352	7.8
	Very small	11	10.7	67	1.5
	Total	103	100.0	4501	100.0
NSW	Large	1390	59.5	76 900	88.3
	Moderately small	374	16.0	6999	8.0
	Very small	574	24.6	3213	3.7
	Total	2338	100.0	87 112	100.0
NT	Large	54	36.0	2212	73.6
	Moderately small	25	16.7	450	15.0
	Very small	71	47.3	343	11.4
	Total	150	100.0	3005	100.0
QLD	Large	759	55.0	50 827	89.4
	Moderately small	207	15.0	3796	6.7
	Very small	413	29.9	2256	4.0
	Total	1379	100.0	56 879	100.0
SA	Large	316	51.4	15 639	81.3
	Moderately small	132	21.5	2479	12.9
	Very small	167	27.2	1127	5.9
	Total	615	100.0	19 245	100.0
TAS	Large	121	53.3	5418	80.2
	Moderately small	53	23.3	996	14.7
	Very small	53	23.3	342	5.1
	Total	227	100.0	6756	100.0
VIC	Large	1070	59.0	56 773	86.6
	Moderately small	336	18.5	6255	9.5
	Very small	407	22.4	2545	3.9
	Total	1813	100.0	65 573	100.0
WA	Large	486	55.0	23 890	85.3
	Moderately small	142	16.1	2682	9.6
	Very small	255	28.9	1445	5.2
	Total	883	100.0	28 017	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 2009 Technical Report.

Sample achieved

The National Assessment Program – Science Literacy specifications set the target sample size at 12 000 students. The total achieved sample size for 2009 was 13 162.

Table A2.3 School participation rates by jurisdiction

State/ Territory	School population	Number of schools sampled	Number of excluded schools	Number of schools that participated	School participation (per cent)
ACT	103	56	1	55	98.2
NSW	2338	92	1	91	98.9
NT	150	50	6	38	76.0
QLD	1379	92	0	92	100.0
SA	615	95	2	93	97.9
TAS	227	63	0	63	100.0
VIC	1813	93	0	93	100.0
WA	883	94	1	93	98.9
AUST	7508	635	11	618	97.3

In total, eleven schools were excluded prior to test date. Of these, eight gave various reasons for not participating. Another three schools refused to participate. A further six 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 NT schools was lower than that for other jurisdictions. From the original target sample of 50 schools, 12 were excluded from the final sample for various reasons. Four schools were exempted, one school was deemed ineligible and seven schools had insufficient eligible students present on test day.

More detail about the achieved sample may be found in the 2009 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 National Assessment Program – Science Literacy target and achieved sample sizes by jurisdiction

State/ Territory	Target sample		Achieved sample	
	Students	Percentage of students	Students	Percentage of students
ACT	1311	9.1	1199	9.1
NSW	2258	15.7	2092	15.9
NT	831	5.8	743	5.6
QLD	2228	15.5	2043	15.5
SA	2005	14.0	1848	14.0
TAS	1276	8.9	1167	8.9
VIC	2243	15.6	2040	15.5
WA	2208	15.4	2030	15.4
AUST	14360	100.0	13162	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	4501	1311	1199	11	0.8	91.5
NSW	87112	2258	2092	8	0.4	92.6
NT	3005	831	743	4	0.5	89.4
QLD	56879	2228	2043	33	1.5	91.7
SA	19245	2005	1848	16	0.8	92.2
TAS	6756	1276	1167	19	1.5	91.5
VIC	65573	2243	2040	37	1.6	90.9
WA	28017	2208	2030	9	0.4	91.9
AUST	271088	14360	13162	137	1.0	91.7

Sample characteristics

Table A2.6 provides a breakdown of the achieved sample across states and territories according to gender, Indigenous status, student's main language 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	46.3	50.5	49.7	50.8	51.6	52.3	49.4	52.1	50.5
Female	53.5	48.4	50.3	49.2	48.3	47.6	50.6	46.2	49.0
Missing	0.3	1.1	0.0	0.0	0.2	0.1	0.0	1.7	0.5
Indigenous status									
Non Aboriginal or Torres Strait Islander	95.5	93.4	67.6	87.7	78.4	82.0	84.0	85.9	85.5
Aboriginal or Torres Strait Islander	3.3	3.8	26.8	5.4	2.2	6.9	0.9	4.1	5.0
Missing	1.2	2.9	5.7	6.9	19.4	11.1	15.0	10.0	9.5
Main language									
English	81.7	60.9	55.3	83.9	63.5	85.9	71.4	62.9	70.6
Other than English	15.1	29.3	24.0	10.3	18.5	12.6	21.0	17.6	18.7
Missing	3.3	9.7	20.7	5.8	18.0	1.5	7.6	19.5	10.7
Geographic location									
Metropolitan areas	100.0	78.0	0.0	69.2	71.5	42.1	74.5	74.8	69.1
Provincial areas	0.0	20.9	66.9	27.1	25.1	57.3	25.5	20.7	27.1
Remote and very remote areas	0.0	1.1	33.1	3.7	3.5	0.6	0.0	4.4	3.8
Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of students	1199	2092	743	2043	1848	1167	2040	2030	13 162

Data has not been presented for parental education and occupation as there is between 30 per cent – 60 per cent 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 locations.

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

Geographic location	ATSI			Total
	Indigenous	Non-Indigenous	Missing	
Metropolitan areas	223	7980	892	9095
Provincial areas	268	2943	351	3562
Remote and very remote areas	163	332	10	505
Total	654	11 255	1253	13 162

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 National Assessment Program – Science Literacy 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 National Assessment Program – Science Literacy testing situation. This includes students who are emotionally or mentally unable to follow even the general instructions of the Assessment.
Limited language proficiency	The student is unable to read or speak any of the languages of the Assessment in the country 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 languages of the Assessment may be excluded.
Refusal	Parent requested that student not participate OR student refused.

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	100	0	7	3	1	111
NSW	158	0	2	3	3	166
NT	86	0	4	0	0	90
QLD	152	3	23	4	3	185
SA	140	0	6	2	8	156
TAS	89	4	11	4	0	108
VIC	166	1	17	5	14	203
WA	168	1	4	3	1	177
AUST	1059	9	74	24	30	1196

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.	All items addressing this strand at this level have been held secure.
	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.	Makes a suggestion for collecting additional data to decide which method produces more tomatoes [in the context of determining additional data required]. Q4 Tomato plants
	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 low efficiency of energy transfer from burning food to water in terms of heat loss to the environment [in the context of measuring the transfer of heat from burning food to water]. Q3 Burning foods

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 and controlling variables.	Identifies requirements for a 'fair test' by selecting what needs to be measured and what needs to be changed. Q2 & 3 Cola fountain
	Strand B: Demonstrates awareness of the need for fair testing by keeping a variable controlled when changing a second variable. Records data as tables or constructs bar graphs from collected or given data.	Extrapolates from an observed pattern to describe an expected outcome or event.	Records data as a column/bar graph, using the conventions of science literacy to label the axes Q5 Which beak works best? practical task
	Strand C: Describes relationships between individual events (including cause and effect relationships) that have been experienced or reported.	Applies knowledge of relationship to explain reported phenomenon.	Identifies the effect of differences in air pressure in an insect pooter [in the context of an ant being sucked up a tube]. Q2 Collecting ants

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. Gives reason for controlling a single variable.	Understands the need for fair testing when conducting an experiment by recognising what elements should 'stay the same' [in the context of designing an experiment to determine which food would attract most ants]. Q1 Collecting ants
	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.	Constructs a scale for the horizontal and vertical axes of a graph and plots data accurately. Q 6 & 7 Which beak works best? 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.	Given initial temperatures of water in two containers, predicts the temperature of combined water. Q3 Heating and cooling

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.	Makes non-standard measurements and limited records of data [in the context of measuring the height of a fountain produced during an experiment]. Q1 Cola fountain
	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.	Identifies a plausible reason for the difference in two families' carbon footprints. Q2 Greenhouse gas emissions
	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.	Interprets information from a food web [in the context of recognising some of the impacts of introduced animals on native flora and fauna]. Q1 Native and introduced animals

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.	Identifies the number of beads (collected) in a simple table of results [in the context of retrieving data from a participant-constructed table]. Q1 Which beak works best? practical task
	Strand C: Describes changes to, differences between or properties of objects or events that have been experienced or reported.	Describes a choice for a situation based on first-hand concrete experience, requiring the application of limited knowledge.	Identifies an item which conducts electricity to complete an electrical circuit. Q1 Energy efficient light bulbs