



MCEETYA



National  
Year 6  
Science  
Assessment  
Report

2003



Ministerial Council on Education,  
Employment, Training and Youth Affairs

© 2004 Curriculum Corporation as the legal entity for the Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA).

Curriculum Corporation as the legal entity for the Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA) owns the copyright in this publication. This publication or any part of it may be used freely only for non-profit education purposes provided the source is clearly acknowledged. The publication may not be sold or used for any other commercial purpose.

Other than as permitted above or by the Copyright Act 1968 (Commonwealth), no part of this publication may be reproduced, stored, published, performed, communicated or adapted, regardless of the form or means (electronic, photocopying or otherwise), without the prior written permission of the copyright owner. Address inquiries regarding copyright to:

MCEETYA Secretariat, PO Box 202,  
Carlton South, VIC 3053, Australia.



**Ministerial Council on Education,  
Employment, Training and Youth Affairs**

# Contents

---

<b>Foreword</b>	<b>vii</b>
<b>Executive Summary</b>	<b>ix</b>
Assessment domain	x
Assessment instruments	x
Student performance in scientific literacy	xi
Standard for Year 6 scientific literacy	xii
<b>Chapter 1</b>	
<b>Overview of the National Assessment</b>	<b>1</b>
Introduction	1
The National Science Assessment	3
What does the National Science Assessment measure?	3
What aspects of scientific literacy were assessed?	4
Who participated in the National Science Assessment?	4
What did the National Science Assessment participants have to do?	5
How are the National Science Assessment results reported?	6
How is this report organised?	6
<b>Chapter 2</b>	
<b>The Scientific Literacy Scale</b>	<b>9</b>
Introduction	9
Scale construction	9
More about the assessment program	12
The assessment booklets	12
The sampling procedures	13
Assessment administration procedures	13
Marking of responses to open-ended items	14
Data entry procedures	14
School reports	15
<b>Chapter 3</b>	
<b>Profile of Student Performance in Scientific Literacy</b>	<b>17</b>
Introduction	17
Interpreting the results	17
Metric for reporting summary performance	19
Reading the bar charts	19
A comparison of scientific literacy among States and Territories	21
Multiple comparisons of achievement	23
Results by gender	24
<b>Chapter 4</b>	
<b>Adding Meaning to the Scientific Literacy Results</b>	<b>27</b>
Introduction	27
Establishing Proficiency Levels	27
Describing the Proficiency Levels	29
Sample items illustrating Proficiency Levels	29
Sample item illustrating performance at Proficiency Level 3.3	29
Sample item illustrating performance at Proficiency Level 3.2	32

Sample items illustrating performance at Proficiency Levels 3.1 and 3.2	33
Sample item illustrating performance at Proficiency Level 3.1	36
Sample item illustrating performance at Proficiency Level 2 and below	39
Results on sample items	40
Distributions of students within Proficiency Levels	40
Distributions of student groups across Proficiency Levels	42
Females and males	42
Indigenous students	43
Students from language backgrounds other than English	45
School location	46
<b>Chapter 5</b>	
<b>The Standard for Year 6 Scientific Literacy</b>	<b>49</b>
Introduction	49
The standard	49
Student groups achieving the Year 6 standard	50
<b>Chapter 6</b>	
<b>Further Consideration of Results on Scientific Literacy Items</b>	<b>53</b>
Introduction	53
Coverage of scientific literacy	53
Major science concepts	54
Types of assessment items	54
The strands	55
Analysis of scientific literacy performance at the item level	56
<b>References</b>	<b>59</b>
<b>Appendix 1</b>	
<b>National Year 6 Primary Science Assessment Domain</b>	<b>61</b>
Scientific literacy	63
Scientific literacy: the assessment domain	64
Scientific concepts in the National Science Assessment	68
<b>Appendix 2</b>	
<b>Sampling</b>	<b>71</b>
Sampling results	73
Sample achieved	74
The main sample	74
School-level student exclusions	75
Population coverage	75
<b>Appendix 3</b>	
<b>Proficiency Levels, Assessment Domain Descriptors, Illustrative Items and Item Descriptors</b>	<b>77</b>

# List of Tables

---

<b>Table ES1</b> Percentages of students at each Proficiency Level on the Scientific Literacy Scale, by State and Territory	xiii
<b>Table ES2</b> Percentages of students at or above each Proficiency Level on the Scientific Literacy scale by State and Territory	xiii
<b>Table 1.1</b> Number of schools and students in the final sample by State and Territory	5
<b>Table 3.1</b> Years of formal schooling by State and Territory	18
<b>Table 3.2</b> Age of participants State and Territory	18
<b>Table 3.3</b> Percentile scores by State and Territory	22
<b>Table 3.4</b> Multiple comparisons of scientific literacy results by State and Territory	23
<b>Table 3.5</b> Scientific literacy of males and females by State and Territory	25
<b>Table 4.1</b> Selected results (percentage correct scores and mean scores on polytomously-scored items, expressed as percentage) on illustrative science items by Proficiency Level, State and Territory	40
<b>Table 4.2</b> Percentages of students at each Proficiency Level on the Scientific Literacy Scale by State and Territory	41
<b>Table 4.3</b> Percentages of students at each Proficiency Level on the Scientific Literacy Scale by gender, State and Territory	43
<b>Table 4.4</b> Mean scores and confidence intervals for Indigenous and non-Indigenous students on scientific literacy	44
<b>Table 4.5</b> Percentages of Indigenous and non-Indigenous students at each Proficiency Level on the Scientific Literacy Scale	44
<b>Table 4.6</b> Mean score and confidence intervals of scientific literacy scores by main language spoken by students at home	45
<b>Table 4.7</b> Percentage distribution of scientific literacy Proficiency Levels by main language spoken at home	46
<b>Table 4.8</b> Distributions of mean scores by MCEETYA geolocation categories	46
<b>Table 4.9</b> Percentage distribution of scientific literacy Proficiency Levels by MCEETYA geolocation category	47
<b>Table 5.1</b> Percentages of students achieving the proficient standard or better, by State and Territory	51
<b>Table 5.2</b> Percentages of students achieving the proficient standard or better across male/female; Indigenous/non-Indigenous; and Home language English/Home language not English groups	51

<b>Table 5.3</b> Percentages of students at or above each Proficiency Level on the Scientific Literacy scale by State and Territory	52
<b>Table 6.1</b> Distribution of assessment items across the assessment domain for scientific literacy	54
<b>Table 6.2</b> Percentages of students omitting responses by item type by State and Territory	55
<b>Table 6.3</b> Numbers of items by State and Territory for which the results were the highest and within the given range from the State or Territory with the highest result	56
<b>Table A1.1</b> Scientific literacy assessment domain	67
<b>Table A1.2</b> Major scientific concepts	69
<b>Table A2.1</b> Schools excluded from the National Science Assessment	73
<b>Table A2.2</b> School sampling frame for the National Science Assessment	74
<b>Table A2.3</b> Student sampling frame for the National Science Assessment	76
<b>Table A3.1</b> Proficiency levels, assessment domain descriptors, illustrative items and item descriptors	79

## List of Figures

---

<b>Figure 3.1</b> Example of a bar chart	20
<b>Figure 3.2</b> Student achievement by State and Territory on the Scientific Literacy Scale	22
<b>Figure 3.3</b> Proficiency of male and female students on the Scientific Literacy Scale	24
<b>Figure 4.1</b> Proficiency levels and cut-off scores	28
<b>Figure 4.2</b> Item set, 'Mosquitos', illustrating the type of item at Proficiency Level 3.3.	30
<b>Figure 4.3</b> Item set, 'Craters', illustrating the type of items at Proficiency Level 3.2	32
<b>Figure 4.4</b> Item set, 'Bush Pond', illustrating the type of item at Proficiency Levels 3.1 and 3.2.	34
<b>Figure 4.5</b> Item set, 'Paper Clips', illustrating the type of item at Proficiency Level 3.1.	37
<b>Figure 4.6</b> Item set, 'Sandpaper', illustrating the type of item in Proficiency Level 2 and below.	39
<b>Figure 4.7</b> Proficiency levels on the Scientific Literacy Scale by State and Territory	41

# Acknowledgements

---

## Steering Committee/Review Committee Members

Listed below are those individuals who participated on either the Steering Committee or the Review Committee at some stage during the development and implementation of the National Science Assessment. Their work made a valuable contribution to the success of the project.

### Chair

Peter Titmanis

### New South Wales

Michelle Bruniges

Phil Daniels

Jim Scott

### Victoria

Neil Champion

Helen Conley

### Queensland

Margo Bampton

Terri Burnet

Ian Cosier

Shelley Peers

### South Australia

Peter Russo

### Western Australia

Fred Deschon

### Tasmania

Meriloy Horsham

Andrew Smith

### Northern Territory

Lisa Scarfe

### Australian Capital Territory

Pauline Brown

Dianne Granthan

### Australian Government

Irene Kaspar

Wendy Whitham

### Catholic school sector

Pauline Sharma

### Independent school sector

Estelle Lewis

### Indigenous education specialist

Paul Newman

### Language background other than

### English specialist

Cheryl Geoghegan

### Australian Academy of Science

Marian Heard

### Australian Science Teachers'

### Association

Deborah Crossing

### PMRT Secretariat

Sandra van Schagen

Christopher Dean

### BEMU

Perette Price

### Australian Council for Educational

### Research

Professor Jim Tognolini

Chris Freeman

Professor Mark Hackling (Consultant)



# Foreword

---

In 2003 a nationally comparable science assessment – the first designed, developed and carried out under the auspices of the national council of education ministers, the Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA) – was carried out.

This assessment represents a new direction in national approaches to reporting on and celebrating the achievements of Australian students and schools.

Development of nationally comparable assessments stands as one of the most notable recent achievements of MCEETYA. They are a key result of a collaborative venture between the States and Territories and the Australian Government to measure and report on how our students are progressing towards the achievement of the National Goals of Schooling in the Twenty-First Century.

Over the next two years, similar reporting at primary and middle-secondary level will occur in civics and citizenship and information and communications technology (ICT) literacy. Subsequent studies of science literacy, civics and citizenship and ICT literacy are planned on a rolling three-yearly basis.

This report on the 2003 National Year 6 Science Assessment provides the key results from the national sample assessment. It gives a snapshot of student results across the national science literacy scale, and an analysis of various trends across states and territories and student sub-groups. The report sets out a range of information that will be useful for continual reflection and improvement in the science curriculum across the nation.

In addition to this report, a separate technical report will be available on the MCEETYA website to provide researchers and others with information on the processes underlying the results.

While part of the assessment instrumentation will be kept confidential to enable its re-use in the next cycle, other test items will be released. These will serve as a valuable resource for teacher professional development and the enhancement of student learning.

The design, development and implementation of the 2003 National Year 6 Science Assessment have relied extensively on the expertise and support of members of the National Education Performance Monitoring Taskforce and its successor, the Performance Measurement and Reporting Taskforce (PMRT) along with the national advisory committees and state bodies responsible for administering these assessments on behalf of MCEETYA. The leadership and work of Peter Titmanis of the PMRT's Benchmark and Measurement Unit (BEMU) have also been invaluable throughout the project's entirety.

Thank you also to the principals, teachers and students at government, Catholic and independent schools across Australia who undertook the trial assessment in 2002 and the first assessment proper in 2003.

This report is testament to the efforts of all involved and serves as an example of what can be achieved by willing collaboration between jurisdictions backed by the vision, dedication and hard work of people in every Australian State and Territory.

Reports such as this enable us to better understand and improve on our children's skills and knowledge. I strongly urge members of the Australian schooling community to make use of this report.

**Ken Smith**

Chair

Performance Measurement and Reporting Taskforce

November 2004

# Executive Summary

---

In July 2001, the Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA) 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 the students, irrespective of where they attended school.

The Australian Council for Educational Research was contracted by PMRT to undertake the National Science Assessment at Year 6.

The report from the National Science Assessment describes the development of the assessment domain and the instruments used to assess that domain; the administration of those assessment instruments; and the marking, analysis and reporting of the results.

The National Science Assessment 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)

## Assessment domain

The assessment domain and instruments were developed in consultation with curriculum experts from each State and Territory and representatives of 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, critiquing the trustworthiness of evidence and claims made by others, and communicating findings.
- Strand C: using science understandings for describing and explaining natural phenomena, interpreting reports and making decisions.

The assessment items drew on four concept areas: Life and Living; Earth and Beyond; Natural and Processed Materials; and Energy and Change. These evolved from a review of the 'National Statements and Profiles' and were common across Australian curricula.

## Assessment instruments

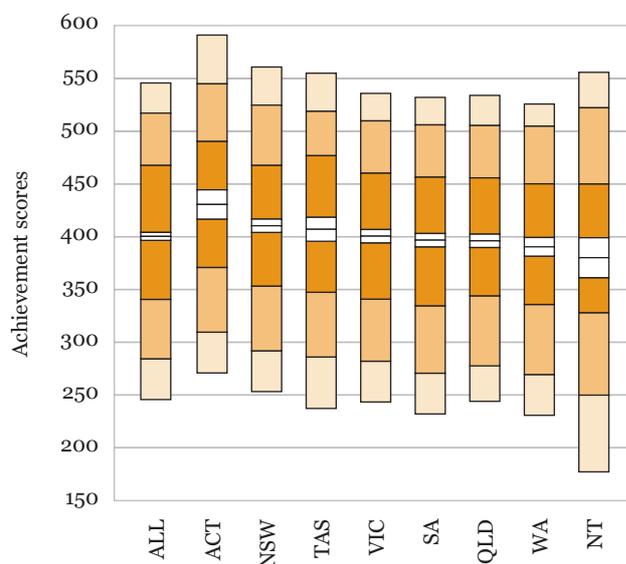
The assessment instruments were administered to a random sample consisting of 6 per cent of the total Australian Year 6 student population. The students' regular classroom teachers administered the National Science Assessment between 20 and 31 October 2003.

Nationally, the assessment instruments consisted of two pencil-and-paper assessments including multiple-choice and short answer type items and two practical assessment tasks. The practical task required the students to conduct an experiment in groups of three and then respond individually to a set of questions about the experiment. Each student completed one of the pencil-and-paper assessments and one of the practical tasks. Students were allowed 60 minutes for the pencil-and-paper assessment and 45 minutes for the practical task.

## Student performance in scientific literacy

One of the main objectives of the National Science Assessment is to monitor trends in scientific literacy over time. One way of doing so is to compare mean achievement scores and the distribution of student scores on the scientific literacy scale.

The mean scores and distributions of scores are shown here:



The National Science Assessment showed that the Australian Capital Territory was the only State or Territory with performance significantly above the national mean.

The highest-achieving students were from the Australian Capital Territory, New South Wales, Tasmania and the Northern Territory. The Northern Territory had students with some of the highest as well as the lowest levels of scientific literacy.

At the national level, the results across the Scientific Proficiency Levels showed the following trends:

- for males and females, there were no significant differences in proficiency;
- the proficiency of non-Indigenous students was significantly higher than that of Indigenous students;
- students whose home language was English showed significantly higher levels of proficiency than those whose home language was not English; and
- students from the MCEETYA geolocation 'remote zone' performed significantly worse than students from any other location.

## Standard for Year 6 scientific literacy

A standard for scientific literacy has been established to provide parents, educators and the community with a clear picture of the proficiency 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.

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 showed superior results and exceeded the proficient standard.

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

The proficient standard will be the main reference point (key performance measure) for monitoring scientific literacy in Australian primary schools over time. Every three years a new national Year 6 science assessment will be conducted to gauge whether student proficiency has improved.

Information about students' performances in relation to the Year 6 standard from the first national Year 6 science assessment is summarised below. The results in Table ES1 show the percentage of students in each of the levels established for scientific literacy while Table ES2 shows the percentage of students who achieved or bettered each of the levels.

**Table ES1** Percentages of students at each Proficiency Level on the Scientific Literacy Scale, by State and Territory

State/Territory	Proficiency Level				
	2 and Below	3.1	Proficient 3.2	3.3	4 and Above
NSW	3.4 (±0.8)	33.7 (±2.1)	52.6 (±2.4)	10.1 (±1.6)	0.1 (±0.2)
VIC	4.4 (±1.0)	36.9 (±2.7)	52.3 (±2.7)	6.3 (±1.2)	0.0 (±0.1)
QLD	5.1 (±0.9)	40.0 (±2.2)	49.0 (±2.0)	5.8 (±1.1)	0.0 (±0.0)
SA	4.4 (±1.2)	38.6 (±2.5)	50.1 (±2.3)	6.8 (±1.3)	0.0 (±0.1)
WA	5.1 (±1.0)	40.3 (±2.2)	48.7 (±2.3)	5.9 (±1.2)	0.0 (±0.0)
TAS	5.0 (±1.4)	35.7 (±2.9)	49.9 (±2.9)	9.3 (±1.8)	0.1 (±0.3)
NT	10.7 (±3.6)	39.9 (±5.6)	42.5 (±4.8)	6.9 (±2.8)	0.0 (±0.0)
ACT	2.7 (±1.1)	27.5 (±3.9)	56.1 (±4.8)	13.3 (±2.7)	0.2 (±0.5)
ALL	4.6 (±0.4)	37.2 (±0.9)	50.5 (±0.9)	7.6 (±0.5)	0.1 (±0.1)

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

**Table ES2** Percentages of students at or above each Proficiency Level on the Scientific Literacy scale by State and Territory

State/Territory	Proficiency Level			
	3.1 or Above	Proficient 3.2 or Above	3.3 or Above	4 or Above
NSW	96.6 (±0.8)	62.8 (±2.1)	10.2 (±1.7)	0.1 (±0.2)
VIC	95.6 (±1.0)	58.7 (±2.5)	6.4 (±1.2)	0.0 (±0.1)
QLD	94.9 (±0.9)	54.9 (±2.1)	5.9 (±1.1)	0.0 (±0.0)
SA	95.6 (±1.2)	57.0 (±2.4)	6.9 (±1.3)	0.0 (±0.1)
WA	94.9 (±1.0)	54.6 (±2.2)	6.0 (±1.2)	0.0 (±0.0)
TAS	95.0 (±1.4)	59.3 (±2.9)	9.4 (±1.8)	0.1 (±0.3)
NT	89.3 (±3.6)	49.4 (±5.5)	6.9 (±2.8)	0.0 (±0.0)
ACT	97.3 (±1.1)	69.8 (±3.9)	13.6 (±2.8)	0.2 (±0.5)
ALL	95.4 (±0.4)	58.2 (±0.9)	7.7 (±0.5)	0.1 (±0.1)

Note: figures in parentheses refer to 95 percent confidence intervals

The results show the same trends as the distributions of students' scores and mean scores, with ACT having the highest proportion of students at the proficient standard or above, i.e. Level 3.2 and above. It should be noted that, as with mean scores, when confidence intervals are taken into account, it is unlikely that there will be a significant difference between the ACT, NSW or Tasmania in terms of the proportion of students achieving the proficient standard.

Nationally, 58.2 percent of students achieved or bettered the proficient standard while approximately 95.4 percent achieved Level 3.1 or above. The highest proficiency levels (Levels 3.3 & 4 and above) were achieved by approximately 7.7 percent of students.

# Chapter 1

## Overview of the National Assessment

---

### Introduction

In 1999, the State, Territory and Commonwealth Ministers for Education agreed to the Adelaide Declaration on National Goals for Schooling in the Twenty-first Century (Ministerial Council on Education, Employment, Training and Youth Affairs [MCEETYA] 1999) (<http://www.mceetya.edu.au/nationalgoals/natgoals.htm>).

The National Goals provide the framework for reporting on student achievement through the annual MCEETYA publication, the National Report on Schooling in Australia (ANR).

The Education Ministers also established the National Education Performance Monitoring Taskforce (NEPMT) in 1999 to develop key performance measures to monitor and report on progress toward the achievement of the Goals on a nationally-comparable basis. They identified eight priority areas for the initial development of performance measures: literacy, numeracy, science, civics and citizenship, information technology, vocational education and training in schools, enterprise education and participation and attainment.

As a first step, NEPMT commissioned a project in early 2000 to develop options for the assessment and reporting of the achievements of primary-school students in science. The outcome of this process was a report to the NEPMT entitled *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 (Ball et al., 2000).*

The Ball report recommended that students' achievement of science literacy (that is, science concepts and science process skills) rather than their acquisition of factual information be assessed and reported at the primary level. In particular, the report advocated adoption of the definition of science literacy used in the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA) for the purposes of primary science monitoring.

In July 2001, MCEETYA agreed to the development of assessment instruments and key performance measures for reporting on student skills, knowledge and understandings in primary science. It directed the newly-established Performance Measurement and Reporting Taskforce (PMRT), which by then had replaced the NEPMT, to undertake the national assessment program. The PMRT commissioned the assessment in July 2001 and the Australian Council for Educational Research (ACER) was the successful tenderer.

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 States and Territories' curricula.

The National Science Assessment is the first assessment program designed specifically to provide information about performance against the National Goals. MCEETYA has also endorsed similar assessment programs to be conducted for civics and citizenship, and information and communications technology (ICT). 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. The first cycle of the program provides the baseline against which future performance will be compared.

Apart from being the first subject area, science is the only program that focuses entirely on primary school performance. This is because MCEETYA has agreed to use PISA as the measure of performance for secondary science.

The Ball report recommended strongly that the assessment of science be conducted at the end of primary schooling because:

*... delay until the end of primary schooling has the advantages of being able to assess a more mature learner who has had greater opportunity to develop scientific skills and processes and develop a better understanding of basic scientific principles.*

(Ball et al., 2000, p. 44)

# The National Science Assessment

Implementation of the National Science Assessment involved a large number of separate but related steps, including the development of an assessment domain and items and instruments to assess that domain; the trialling of those items and assessment instruments; the development of key performance measures; 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 and presents the results at the national and State and Territory levels.

## What does the National Science Assessment measure?

The National Science Assessment measures scientific literacy. This is a construct that:

*...encompasses the use of broad conceptual understandings of science for making sense of the world, understanding natural phenomena, and interpreting media reports about scientific issues. It also encompasses competencies related to asking investigable questions, conducting investigations, collecting and interpreting data and making decisions.*

(Review of curriculum frameworks: assessment frameworks: reporting policies and assessment policies. Hackling, 2002, p. 1).

This construct has evolved from the definition of scientific literacy used by 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 science items and instruments therefore assess outcomes that contribute to scientific literacy, such as conceptual understandings, rather than facts, and investigate competencies in realistic situations. The National Science Assessment relates to the ability to think scientifically in a world in which science and technology are increasingly shaping children's lives.

An assessment domain was developed in consultation with curriculum experts from each State and Territory and representatives of the Catholic and independent school sectors. This domain includes the definition of scientific literacy and outlines the development of scientific literacy across three main areas (see Appendix 1).

## What aspects of scientific literacy were assessed?

Three strands 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, critiquing the trustworthiness of evidence and claims made by others, and communicating findings.
- STRAND C:** using science understandings for describing and explaining natural phenomena, interpreting reports and making decisions.

A conscious effort was made to develop assessment items that related to everyday contexts rather than to laboratory situations.

The items drew on four concept areas: Life and Living; Earth and Beyond; Natural and Processed Materials; and Energy and Change. These evolved from a review of the 'National Statements and Profiles' and were common across Australian curricula. It is interesting to note that the same concept areas are also common internationally.

The strands of scientific literacy and the concepts to be assessed were informed by a thorough analysis and mapping of the syllabuses of all States and Territories. The intention was to ensure that all Year 6 students were familiar with the materials and experiences to be used in the National Science Assessment and so avoid any systematic bias in the instruments being developed.

## Who participated in the National Science Assessment?

Approximately six per cent of the total Australian Year 6 student population was sampled randomly and assessed. The sample was drawn from all States and Territories and government, Catholic and independent schools participated. Table 1.1 shows the number of schools and students in the final sample for whom results were reported.

A grade-based population of students enrolled at schools was chosen. This is consistent with the reporting of literacy and numeracy performance in the ANR. There are differences among 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 Tables 3.1 and 3.2.

Appendix 2 provides a comprehensive summary of the sample frame, including exclusions and response rates for participating schools and students by State and Territory for the assessment.

**Table 1.1** Number of schools and students in the final sample by State and Territory

State/Territory	Number of Schools in Target Sample	Number and % <sup>1</sup> of Schools in Final Sample	Number of Students in Final Sample
NSW	122	103 (84%)	2 466
VIC	122	100 (82%)	2 130
QLD	122	110 (90%)	2 607
SA	130	115 (88%)	2 032
WA	126	103 (81%)	2 347
TAS	64	60 (94%)	1 240
NT	32	23 (72%)	496
ACT	44	36 (82%)	854
ALL	762	650 (85%)	14 172

<sup>1</sup> 'Percentage of schools' is calculated by dividing the number of schools in the final sample by the number of schools in the target sample for each State and Territory and multiplying by 100.

In a number of cases, schools nominated students to participate in the National Science Assessment who were outside the target population. These students were not included in the results. There were also instances of schools that wished to participate in the study as volunteer schools. Their students were not included in the results, although their assessments were marked and feedback was provided to the schools on the performance of their students. Some schools with multi-level classes asked that their Years 5 and 7 students also complete the assessment. Once again these students' tests were identified and removed from the sample.

## What did the National Science Assessment participants have to do?

There were two pencil-and-paper assessments which included multiple-choice and short answer type items; there were also two practical assessment tasks. The assessment papers were distributed randomly so that half the students in each class completed one of the pencil-and-paper assessments and the other half the second pencil-and-paper assessment.

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

The practical task required the students to work in groups of three. The teachers, using a procedure outlined in the Assessment Administrators Manual, allocated students randomly to groups. Students conducted the experiment in

these groups and responded to a set of questions designed to stimulate group discussion about the experiment.

The students then answered a further set of items independently. The individual student responses were the only ones used in the analysis and generation of proficiency data.

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

Students were allowed 60 minutes for the pencil-and-paper assessments and 45 minutes for the practical tasks.

The students' regular classroom teachers administered the National Science Assessment between 20 and 31 October 2003.

## How are the National Science Assessment results reported?

The results of the National Science Assessment are reported as mean scores and distributions of scores. 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 scales, including results in relation to the scales by State and Territory, are contained in Chapter 4, 'Adding Meaning to the Scientific Literacy Results'.

Results for groups such as Indigenous students, students from language backgrounds other than English and students from different geographic locations are presented in this Report.

## How is this report organised?

Chapter 2 provides a more detailed discussion of the assessment domain and the procedures used.

Results in terms of means and distributions of students' performances are presented in Chapter 3.

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. Results achieved by Indigenous students, by students with language backgrounds other than English and by students from diverse geographic locations are also included in this chapter.

Chapter 5 identifies the standard for Year 6 scientific literacy and describes students' proficiency in relation to the standard.



# Chapter 2

## The Scientific Literacy Scale

---

### Introduction

This chapter provides a brief description of the steps that were used to define and construct the scientific literacy scale.

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.

### Scale construction

A measurement and reporting scale for scientific literacy was developed that would enable standards and changes in student proficiency over time to be monitored.

The construction of the scale involved a number of interrelated steps:

1. defining the assessment domain for scientific literacy, i.e. clarifying what was to be measured;
2. describing increasingly complex student understandings and skills within the domain;

3. constructing assessments that comprised items and tasks defining the assessment domain operationally and covering the full range of proficiency expected to be represented in Year 6 classes;
4. administering the assessments to students; and
5. using the measurement model and technical standards to develop the Scientific Literacy Scale.

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.

In addition, the contractor set up a number of committees and ad hoc advisory groups.

The main function of all of these committees and groups 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 the students irrespective of where they attended school.

A brief description of the steps involved in developing the scientific literacy measurement and reporting scale is provided here.

### **1. Defining the assessment domain for scientific literacy**

The PISA definition formed the basis of the work to assess the scientific literacy of Year 6 students in Australia (see Appendix 1).

Associate Professor Mark Hackling of Edith Cowan University prepared a draft of the assessment domain that included descriptions of the strands, the initial hierarchy of students' understandings and skills and the concept areas. This was then made available for consultation.

The development of the National Science Assessment was characterised by a high level of communication with stakeholders and regular feedback and consultation with representatives of each of the jurisdictions and nominated members of the PMRT. Through various committees and groups, experts from the States and Territories contributed to and commented on all aspects of the study.

### **2. Describing increasingly complex student understandings and skills within the domain**

The assessment of students' scientific literacy required the development of a measurement scale.

The scale was conceptualised by describing the main understandings and skills that students were expected to develop during the years of schooling. Descriptions were developed to form a hierarchy of increasingly complex understandings and skills. Difficult tasks or items that would challenge the most able students were located at the upper end of the hierarchy

and define the upper end of the scale. Conversely, those items that were relatively easy and could be answered by students with little scientific literacy understandings and skills were located toward the base of the hierarchy and defined the lower end of the scale.

### **3. Constructing assessments that comprised items and tasks defining the assessment domain operationally and covered the full range of proficiency expected to be represented in Year 6 classes**

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 domain. They also had to ensure that the tasks intended to assess higher-order understandings and skills at the top of the scale were, in fact, harder than those at the bottom of the scale.

The items were reviewed first by ACER and its internal panels, then, after trialling with samples of students in four States and Territories, by advisory committees and other key staff in the States and Territories.

The PMRT set the policy objectives for the National Science Assessment and the policy priorities for the implementation of the assessment program, including endorsing the definition of scientific literacy, the assessment domain, the items and the plans for reporting results.

After receiving advice from ACER, it also approved the more technical aspects of the design, including, for example, the number of assessment booklets, the ratio of multiple-choice to open-ended items in the booklets, and the number of items per domain per test booklet.

Teachers from all States and Territories were involved in reviewing the tasks and items during an exercise to establish proficiency levels along the measurement and reporting scale. In addition, they were involved in marking the tasks.

The emphasis during these reviews was on ensuring that the items and tasks reflected the understandings and skills in the assessment domain and were not biased unduly for or against particular groups of students.

### **4. Administering the assessments to students**

Once the items and tasks had been written, they were trialled with a sample of students in 24 schools selected from the government, Catholic and independent sectors in New South Wales, Victoria, Queensland and Western Australia. Five Northern Territory schools were also included in response to a request from the Northern Territory representative.

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 between 21 and 31 October 2003. The total number of students in the final sample was 14,172 at 650 schools. The student sampling frame is shown as Table A2.3 of Appendix 2.

## **5. Using the measurement model and technical standards to develop the scientific literacy scale**

The Rasch measurement model was used to analyse the results from the sample of students who participated in the National Science Assessment. This model is 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 technical report for the National Science Assessment.

As part of the quality assurance procedures, an expert psychometric panel was established to set technical standards for sampling, data analysis, scale construction, fit to the Rasch model and differential item functioning. The panel included psychometricians from the PMRT's Expert Measurement Advisory Group (EMAG) and sampling and analysis experts from the PISA team.

# More about the assessment program

## The assessment booklets

The National Science Assessment involved the use of four assessment booklets. Two of the booklets assessed practical tasks and the others pencil-and-paper items. Participating students had to complete one booklet from each category. The practical task booklet required students to undertake an activity in small groups and then respond individually to pencil-and-paper items relating to the activity.

The multiple-choice items in the booklets had only single correct answers. The open-ended items required students to construct their own responses. Some of these items had only single correct answers, while others provided for a partial credit marking scheme to accommodate a range of response levels and allow a wider range of skills to be assessed. The open-ended items were further categorised into those that required a single-word or short-sentence response

(referred to as ‘short answer’ items) and those that required more substantial responses (referred to as ‘extended response’ items).

Each of the items and tasks included stimulus material that was followed by a series of questions relating to the material. Both of the pencil-and-paper assessments contained 13 units (each consisting of stimulus material and a number of associated questions). The two practical assessments comprised a practical task, group activities and either four or five items to be answered individually by students.

## The sampling procedures

The sample was selected using procedures similar to those followed in PISA and TIMSS. The distribution of schools from the various sectors in each State and Territory was drawn according to each sector’s proportion of Year 6 enrolments.

In the smaller States and Territories, the sample size required to achieve the same degree of confidence in the results as the larger States and Territories would have meant that almost all Year 6 students would have had to participate in the National Science Assessment. Consequently, the sample numbers in the Australian Capital Territory, Northern Territory and Tasmania were reduced.

This increased the level of uncertainty associated with the results, although they are still well within acceptable limits.

Staff who were responsible for the sampling in the PISA project oversaw the sampling in the National Science Assessment.

Further information about the characteristics of the sample, including details of students who were granted exemptions or were excluded from the sample, and the procedures used to determine the standard errors of estimates, is provided in Chapter 3 and Appendix 2, as well as in the technical report.

## Assessment administration procedures

Students’ regular class teachers administered the National Science Assessment to minimise disruption to the normal class environment.

Standardised administration procedures were developed and brought together in an Assessment Administrators Manual. In all schools in which students were to complete the National Science Assessment, teachers and school administrators were provided with the manual. Detailed instructions were also given in relation to the participation or exclusion of students with disabilities and students from non-English speaking backgrounds.

To help standardise the assessment conditions and familiarise teachers with the procedures and requirements of the practical assessment, training videos were provided.

The teachers were able to review the manual and the training videos before the assessment date and raise questions with the coordinators of the National Science Assessment in their jurisdictions.

As a result, it was expected that standardised administration of the Assessment would be achieved.

A quality-monitoring program was established to gauge the extent to which class teachers followed the specified administration procedures. This involved trained monitors observing the administration of the Assessment in a random sample of classes in 48 of the 650 schools involved. The monitors reported a high degree of compliance with the administration procedures.

## Marking of responses to open-ended items

Approximately two-thirds of the items were open-ended and required marking by trained markers. Most involved single answers or phrases that could be marked objectively.

Marking guides were prepared by ACER, and refined during the trialling process. The marking team included representatives from most jurisdictions and experienced markers employed by ACER.

The markers participated in a five-hour training session conducted by a member of the test construction team. The session involved formal presentations by the trainers, followed by hands-on practice with sample student answer books. In addition, the markers undertook a further two hours of marking in which a pair of markers marked the same student answer books and moderators reconciled differences in discussion with the markers.

Markers were monitored constantly for reliability by having samples of their student answer books re-marked by senior markers. In cases in which there were differences between markers and senior markers, the scoring was reconciled jointly. This procedure, coupled with the intensive training at the beginning of the marking exercise, ensured that markers were applying the criteria consistently.

## Data entry procedures

Scanning software was used to capture the responses of individual students from the multiple-choice questions in the test booklets. A validation of the scanning process was performed that demonstrated 100 per cent accuracy in data capture.

Images of all student responses were collected and stored.

## School reports

Schools that participated in the National Science Assessment were provided with feedback about the performance of their students before the end of the 2003 school year.

The reports showed the results for each student on an item-by-item basis and presented information about the percentages of students in the national sample and at the school responding correctly to the items.



# Chapter 3

## Profile of Student Performance in Scientific Literacy

---

### Introduction

In this chapter summary statistics for the National Science Assessment are shown in terms of students' mean scores and distributions of scores by State and Territory.

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

### Interpreting the results

There are a number of design effects and structural differences among States and Territories that must be kept in mind in interpreting the results from the National Science Assessment.

The sampling frame for the Assessment targeted all schools with Year 6 students. However, schools with fewer than five students in Year 6 were excluded to ensure that at least three students were present on the day of testing to work together as a group on the practical tasks. In addition, several geographically isolated schools were excluded because of the logistical difficulties of delivering and returning assessment materials.

Appendix 2 shows the percentages of exclusions from the various States and Territories.

Structural differences include the length of time that students have spent in formal schooling by the time they are in Year 6 and their age at the time of the Assessment.

Table 3.1 shows the starting ages and the number of years of formal schooling for each of the States and Territories.

Table 3.2 shows the distributions of ages of students in the sample chosen for the National Science Assessment, by State and Territory.

**Table 3.1** Years of schooling by State and Territory

State/Territory	Average Age at Time of Testing	Average Time at School
NSW	11 yrs 11 m	5 yrs 10 m
VIC	12 yrs 1 m	6 yrs 9 m
QLD	11 yrs 6 m	5 yrs 10 m
SA	11 yrs 8 m	6 yrs 5 m
WA	11 yrs 4 m	5 yrs 9m
TAS	12 yrs 2 m	6 yrs 9 m
NT	11 yrs 10 m	6 yrs 5 m
ACT	12 yrs 0 m	6 yrs 8 m

**Table 3.2** Age of participants by State and Territory

State/Territory	Current Age in Years				
	10 (%)	11 (%)	12 (%)	13 (%)	Others* (%)
NSW	0.3	47.7	50.4	1.2	0.4
VIC	0.2	33.6	63.0	2.3	0.9
QLD	8.6	82.1	8.6	0.1	0.6
SA	0.6	61.1	36.7	0.2	0.4
WA	10.0	85.0	4.1	0.1	0.8
TAS	0.2	20.5	77.9	0.6	0.8
NT	1.6	64.7	32.3	0.6	0.8
ACT	0.2	40.6	56.7	1.9	0.4

\* These students did not indicate their age on the test booklet.

It can be seen from Table 3.2 that 95 per cent of the students in the sample from Western Australia and approximately 91 per cent of the sample from Queensland were 11 years old or less, while approximately 41, 48, 34 and 20 per cent respectively of the samples of students from the Australian Capital Territory, New South Wales, Victoria and Tasmania were 11 years old or less.

The influence of these differences in age on the results is not known. However, when a difference is combined with length of time in formal schooling, it would

be reasonable to assume that these factors would account for some of the variations in the results observed in the National Science Assessment.

## Metric for reporting summary performance

As noted in Chapter 2, the PISA definition of scientific literacy formed the basis for the assessment domain.

Students' scores on the Scientific Literacy Scale were analysed and transformed into a scale that had a national mean of 400 and a standard deviation of 100. The choice of these values meant that about two-thirds of the students in the National Science Assessment scored between 300 and 500 on the Scientific Literacy Scale.

## Reading the bar charts

Figure 3.1 is an example of the bar chart used to display the scaled mean scores and distributions for States and Territories.

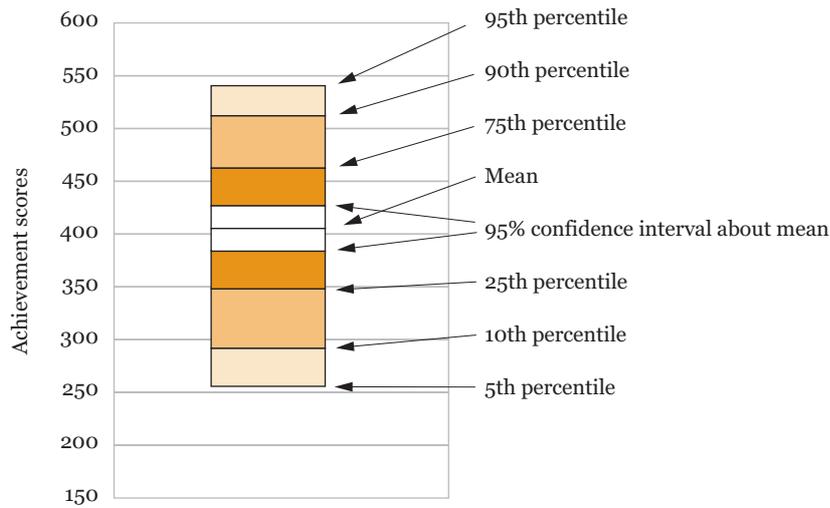
A vertical bar shows the range of student performance.

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 fifth percentile – the point below which the lowest-scoring 5 per cent of students are located.

Located in the middle region of each bar is a white band with a thin horizontal black line. This black line denotes the mean score, while the white regions on either side give an indication, through the height of the band, of the level of accuracy with which the mean was measured (the smaller the band, the more accurate the measurement).

**Figure 3.1** Example of a bar chart



In technical terms, the white band represents a region of about two 'standard errors' (SE) of the mean on either side of it.

Each State and Territory's result is an estimate of the total population value, inferred from the result obtained by the sample of students tested. Because it is an estimate, it is subject to uncertainty.

If the mean scores were estimated from different samples drawn from the same population of students, the actual results for the mean would vary a little. However, the reader may be confident that the population mean lies between the value obtained and about two SE (actually 1.96) on either side of it.

According to statistical theory, the estimate of the mean from repeated sampling would be expected to fall within that range for 95 of each one hundred samples that were drawn.

The white bands (confidence intervals) vary in size from one State and Territory to the next. Their width is a function of the State or Territory's sample size and the spread of achievement scores on the test. The sample sizes vary in proportion to population so the States and Territories with the smallest populations have the smallest samples and the widest white bands.

## A comparison of scientific literacy among States and Territories

Figure 3.2 shows students' performances in scientific literacy for each State and Territory.

It can be seen that the Northern Territory had the widest spread of scores achieved by the middle 90 per cent of students (those between the fifth and the 95th percentiles). Table 3.3 shows that the actual difference was 386 points.

The Australian Capital Territory and Tasmania had the next widest spread of scores between the fifth and the 95th percentiles – 322 and 317 respectively.

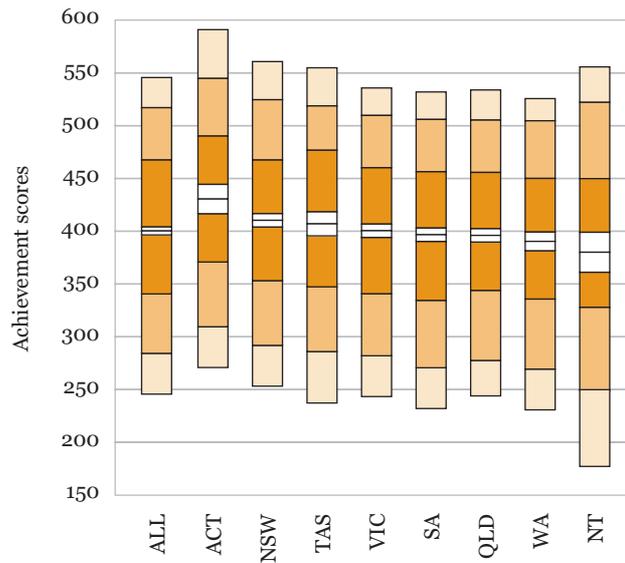
The States and Territories with the least spread of achievement scores for the middle 90 per cent of students were Queensland and Victoria, with 293 and 292 respectively. The spread across Australia was approximately 301.

It can also be seen from Table 3.3 that the Australian Capital Territory had the highest mean score (430), followed by New South Wales (411) and Tasmania (407).

Figure 3.2 shows that approximately 60 per cent of the students in the Australian Capital Territory scored above the national mean and approximately 60 per cent of the students in the Northern Territory and Queensland scored below the national mean.

The bar charts can be used to determine visually whether one State or Territory's mean score is significantly different from that of another. For the means to be significantly different, the white bands on the State and Territory bars should not overlap on the vertical (scores) scale.

**Figure 3.2** Student achievement by State and Territory on the Scientific Literacy Scale



**Table 3.3** Percentile scores by State and Territory

State/ Territory	Mean score	95 per cent confidence interval	Percentile					
			5th	10th	25th	75th	90th	95th
ACT	430	12.4	269	307	371	492	546	591
NSW	411	8.0	252	292	353	468	524	563
TAS	407	12.0	233	286	347	475	522	550
VIC	399	8.2	245	281	341	459	506	537
SA	393	8.0	226	270	333	459	506	537
QLD	392	7.4	239	276	335	453	503	532
WA	390	9.4	230	275	336	449	505	531
NT	379	19.8	173	252	326	448	522	559
ALL	400	3.8	245	282	342	460	512	546

Figure 3.2 shows that the Australian Capital Territory was the only State or Territory with a performance of statistical significance above the national mean (represented by 'ALL'): the bottom of the white band for the Territory was higher than the mean and the white band for Australia.

No States and Territories had the top of its white band below the white band for Australia. This indicates that none performed significantly below the national mean. The figure also shows that the means for all States and Territories, other than that for the Australian Capital Territory, were not significantly different from the mean for Australia.

The highest-achieving students (those at the 95th percentile for their States and Territories) were from the Australian Capital Territory, New South Wales, the Northern Territory and Tasmania. Conversely, the lowest-performing students (those at the fifth percentile for their States and Territories) were students from the Northern Territory, South Australia, Western Australia and Tasmania.

## Multiple comparisons of achievement

Table 3.4 enables a quick comparison of States and Territories to be made. The statistical technique compares the results of the States and Territories on a pairwise-comparison basis. (The Bonferoni adjustment has not been made in Table 3.4.)

**Table 3.4** Multiple comparisons of scientific literacy results by State and Territory

		ACT	NSW	TAS	VIC	SA	QLD	WA	NT	ALL
	Mean	430	411	407	399	393	392	390	379	400
	95 per cent CI	12.4	8.0	12.0	8.2	8.0	7.4	9.4	19.8	3.8
ACT	430	12.4	0	0	1	1	1	1	1	1
NSW	411	8.0	0	0	0	1	1	1	1	0
TAS	407	12.0	0	0	0	0	0	0	0	0
VIC	399	8.2	-1	0	0	0	0	0	0	0
SA	393	8.0	-1	-1	0	0	0	0	0	0
QLD	392	7.4	-1	-1	0	0	0	0	0	0
WA	390	9.4	-1	-1	0	0	0	0	0	0
NT	379	19.8	-1	-1	0	0	0	0	0	0
ALL	400	3.8	-1	0	0	0	0	0	0	0

Note: read across the row to compare a State or Territory's performance with the performance of each State and Territory listed at the top of a column. '1' represents mean performance that is statistically significantly higher than in the comparison State or Territory; '0' indicates that there is no statistically significant difference; and '-1' that mean performance is statistically significantly lower.

It can be seen by reference to Figure 3.2 and Table 3.4 that the students from the Australian Capital Territory achieved a significantly higher mean score than those from all the other States and Territories except New South Wales and Tasmania.

The students from New South Wales achieved a significantly higher mean score than those from all the other States and Territories except the Australian Capital Territory, Tasmania and Victoria.

There was no significant difference in the performance of students from Victoria, Western Australia, South Australia, Queensland and the Northern Territory.

## Results by gender

In this section, gender differences – in terms of mean scores and the distributions of results – are considered across the States and Territories.

There is further discussion of the differences in the proficiency of males and females in Chapter 4.

Figure 3.3 and Table 3.5 show that on the Scientific Literacy Scale, with a mean of 400 and a standard deviation of 100, the mean score for males was 7 points above the mean score for females. This difference, however, was not significant at the 0.05 level.

Table 3.5 indicates that the white bands (confidence intervals) shown in Figure 3.3 are slightly wider for males than for females indicating a greater spread in scores for males than for females. In Figure 3.3, the vertical columns are slightly longer for males than for females.

**Figure 3.3** Proficiency of male and female students on the Scientific Literacy Scale

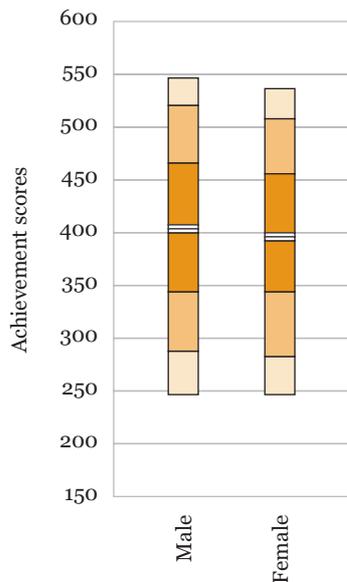


Table 3.5 shows the relative performance of males and females by State and Territory and provides the standard error of measurement (SE) for each State and Territory.

**Table 3.5** Scientific literacy of males and females by State and Territory

State/ Territory	% of Males in Sample	Mean Scores			
		Males	95 per cent CI	Females	95 per cent CI
NSW	49.8	415	10.2	407	8.0
VIC	51.9	403	11.6	397	9.8
QLD	51.5	398	9.4	386	7.6
SA	53.0	393	10.2	392	9.6
WA	51.1	395	10.4	386	11.0
TAS	49.6	411	15.6	402	12.6
NT	48.6	380	23.2	377	20.8
ACT	51.2	431	12.2	430	20.6
ALL	51.1	405	5.0	398	4.2

It can be seen from Table 3.5 that males in the Australian Capital Territory were the highest-performing group (mean 431) followed by females from the same jurisdiction (mean 430). The poorest-performing group of students was females from the Northern Territory (mean 377).

Although the differences were not statistically significant in any particular State and Territory or overall, the tendency for males to perform better than females was consistent in all cases. The highest differences were observed in Queensland (mean difference 12), Western Australia and Tasmania (9), followed by New South Wales (8). The smallest difference was in the Australian Capital Territory (1).

The results were consistent with PISA 2000, which showed no statistically significant difference between the performances of males and females in scientific literacy across States and Territories. However, where PISA did find differences, they tended to be in favour of females:

*The actual best estimates were close to identical for males and females in science, except in Queensland and Tasmania, where the females estimate was higher by about 17 scale points, but this difference was not large enough to be statistically significant at the 0.05 level.*

(Lokan et al., 2000, p. 125)

Table 3.5 also provides information about the proportion of the sample composed of male students – 51.1 per cent. The proportion was highest in South Australia and lowest in the Northern Territory.



# Chapter 4

## Adding Meaning to the Scientific Literacy Results

---

### Introduction

Chapter 3 showed students' scores and the 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.

### Establishing Proficiency Levels

One of the main objectives of the National Science Assessment 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.

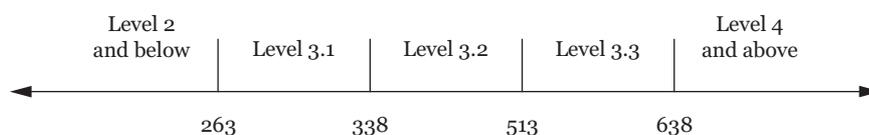
Typically, students whose results are located within a particular Proficiency Level are able to demonstrate the understandings and skills associated with that level and possess the understandings and skills of lower Proficiency Levels.

To establish the Proficiency Levels, a combination of experts' knowledge of the skills required to answer each scientific literacy item and information from the analysis of students' responses was used. Initially, three Proficiency Levels, corresponding with Levels 2, 3 and 4 of the assessment domain, were identified.

However, as 90 per cent of students' scores fell between 116 and 534 on the Scientific Literacy Scale and the majority were in Level 3, three further Proficiency Levels within Level 3 were created, providing five levels for reporting student performance in the Assessment.

The corresponding cut-off scores are shown in Figure 4.1.

**Figure 4.1** Proficiency levels and cut-off scores



It can be seen that a score of 638 or more locates students in Proficiency Level 4 and above.

Similarly, scores in the range of 263 to 637 relate to Proficiency Level 3 on the assessment framework.

Items with difficulties (scores) locating them within each Proficiency Level were judged by subject experts to share similar features and requirements and to differ in recognisable ways from items at other levels.

The cut-offs were set so that approximately equal changes in item difficulty were represented in each level and students whose scaled scores fell within a particular level could be expected to answer at least half of the items at that level correctly.

The difficulty range spanned by each 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 the same 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 the 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 in that level. On average these students would be expected to answer about 50 per cent of the items in that level correctly.

## Describing the Proficiency Levels

Appendix 3 provides the descriptions of knowledge and skills required of students at each Proficiency Level. The descriptions reflect the skills assessed by the full range of scientific literacy items, including the three domains of scientific literacy.

It can be seen from Appendix 3 that the descriptors come from the scientific literacy assessment domain presented in Appendix 1, where Level 3 has been further divided into sub-levels 3.1, 3.2 and 3.3.

## Sample items illustrating Proficiency Levels

This section provides sample items that illustrate the types of understandings and skills that students located at a particular Proficiency Level are likely to be able to display successfully.

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

Only a small number of items has been released in this report – others have been retained for future national science assessments.

Those items chosen for presentation here are recorded in Table A3.1 as ‘illustrative items’. From the descriptors included with the items it is possible to see that those items at the higher Proficiency Levels require more demanding skills and understandings to answer them than do those at the lower levels.

A table of results by State and Territory on the ‘illustrative items’ is provided in a later section of this chapter.

## Sample item illustrating performance at Proficiency Level 3.3

There were no items operating at Proficiency Level 4 and above. However, an item set illustrating performance at Proficiency Level 3.3 is shown in Figure 4.2.

The item set is related to the life cycle of mosquitos. It is assessing Strand C and the context is the Life and Living area. It assesses students’ ability to examine a table of data formed from experimental observation, make a prediction as to the conditions under which a larva changes into a pupa and then draw a conclusion regarding the outcome from the experiment. At a more general level, students who can complete items requiring the same level of scientific literacy as this item would be able to describe key features of a collected set of data and predict outcomes of the next event in a series.

Question 4 of the item set requires students to identify two events that must occur before the pupa evolves from the larva. Students who identify two factors (length and time) as being required are awarded two marks. Those who only indicate that the 'length must be at least 10 mm' are awarded one mark. The item is located at 611 on the scientific literacy scale.

An average Year 6 student would have a 9 per cent chance of obtaining two marks on Question 4 of this item set and a 37 per cent chance of scoring one mark.

**Figure 4.2** Item set, 'Mosquitos', illustrating the type of item at Proficiency Level 3.3.

**MOSQUITOS**  
*Life Cycle of a Mosquito*

Grade 6 students collected mosquito larvae from a pond near the school.

- The larvae were divided into four separate groups.
- Each group was placed in a large jar containing water from the pond.
- A thermometer to measure temperature was placed in each jar.
- Small heaters were used to control the temperature in each jar.

10°C Jar 1      16°C Jar 2      22°C Jar 3      28°C Jar 4

**Q1** The life cycle diagram is shown in a **circle**. What does this mean?

\_\_\_\_\_

\_\_\_\_\_

9

**Figure 4.2** (continued) Item set, 'Mosquitos', illustrating the type of item at Proficiency Level 3.3.

The larvae were examined each day for nine days. Each day three larvae were taken from each jar and their length measured (in millimetres).

Jar	Temperature	Average length of larvae (mm)								
		Day								
		1	2	3	4	5	6	7	8	9
1	10°C	4	4.5	5	5	5.5	5.5	6	6	6
2	16°C	4	5	6	7	8	9	10	pupa	pupa
3	22°C	4	5.5	7	8.5	10	10	pupa	pupa	pupa
4	28°C	4	6	8	10	10	10	pupa	pupa	pupa

**Q2** Look at the pattern in the data. As the temperature went up what happened to the growth of the larvae?

---



---

**Q3** Compare the **results** for Jar 1 with the other jars. Give one reason why the time taken for the pupa to form is different.

---



---

**Q4** Look at the table. What two things must occur before a larva can change into a pupa?

1. \_\_\_\_\_

---

2. \_\_\_\_\_

---

## Sample item illustrating performance at Proficiency Level 3.2

An item illustrative of performance at Proficiency Level 3.2 is shown in Figure 4.3.

The question is from the practical activity undertaken by groups of students.

The activity requires them to drop marbles from different heights into small pans of flour. The students have to measure the width of the crater formed and record it in a table with the height from which the marble was dropped.

When the activity has been completed, each student is required to answer a set of items related to the activity. One item is shown in Figure 4.3.

This extended response type item assesses students' ability to identify and summarise patterns in experimental data. In order to obtain full marks for the item, students had to understand that the width of the crater was related directly to the height from which the marble had been dropped.

The item is located at 486 on the Scientific Literacy Scale. It assesses outcomes reflecting Strand B using concepts related to the Earth and Beyond area.

Year 6 students whose test results corresponded with the national mean score would have a 48 per cent chance of responding correctly to this item.

**Figure 4.3** Item set, 'Craters', illustrating the type of items at Proficiency Level 3.2

<p>Q2 What effect did drop height have on the width of the crater?</p> <hr/> <hr/>
--

## Sample items illustrating performance at Proficiency Levels 3.1 and 3.2

The item set shown in Figure 4.4 comes from a set entitled ‘Bush Pond’.

In this item, students are referred to a picture that contains a variety of plants and animals. The relationship between the animals and the food eaten by them is summarised in a table.

The first item (Question 1) is a multiple-choice item that requires students to identify patterns in scientific data that are presented in a table. The correct answer is Option B. The item assesses Strand C and uses concepts from the Life and Living area.

It is located at 446 on the Scientific Literacy Scale, which means that it is moderately difficult. It is located at Proficiency Level 3.2.

The second item in the set of three requires the students to complete a food web for the pond life. Students are asked to generate a hypothesis from the data provided in the item, draw conclusions and present the result in a structured scientific format.

Question 2 of this item set is located at 351 on the Scientific Literacy Scale, which means that it is easier than the first item in the set and is indicative of items in Proficiency Level 3.1.

The third item of the set (Question 3) is shown to illustrate the range of difficulty of items within Proficiency Level 3.2.

It is an extended response item that requires students to interpret information and then predict what will happen when the conditions change. The item is polytomously scored, so students who provide all four of the following responses – *tadpoles*, *frogs*, *fish* and *snakes* – are awarded two marks, those who provide any three receive one mark and those who do not provide three gain no marks.

The item is located at 457 on the Scientific Literacy Scale. This means that it is the most demanding item of the set. The item assesses Strand C.

Typically, students at Proficiency Level 3.2 are capable of describing the findings of an experiment in simple terms, focusing on one variable.

The location for scoring full marks on the third item is 593 on the Scientific Literacy Scale. This means that Year 6 students whose test results corresponded with the national mean score would have a 23 per cent chance of obtaining full marks on this item and an 83 per cent chance of scoring one mark.

**Figure 4.4** Item set, 'Bush Pond', illustrating the type of item at Proficiency Levels 3.1 and 3.2.

**BUSH POND**

The picture shows animals and plants that live in or around a bush pond. The tiny animals and plants have been shown in close-up.

**Animals**

  
cricket

  
fish

  
frogs

  
mosquito

  
snake

  
tadpole

**Plants**

  
algae  
(tiny plants)

  
rushes

  
water weed

Some facts about the food eaten by the animals in **this** pond.

Food	snake	frog	tadpole	mosquito	cricket	fish
tiny pond plants			✓			
some animals but not insects or tadpoles.	✓					
animal material but animal is not killed				✓		
tadpoles						✓
insects		✓	✓			✓
large pond plants					✓	

15

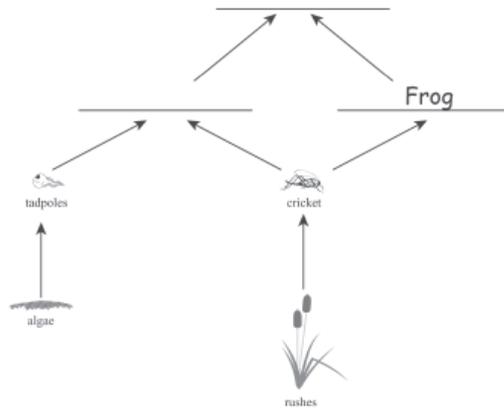
**Figure 4.4** (continued) Item set, 'Bush Pond', illustrating the type of item at Proficiency Levels 3.1 and 3.2.

**Producers** are living things that make their own food. **Consumers** get their food by eating other living things.

**Q1** Which of the following lists contain **only** consumers?

- tadpole, water weed, fish
- frog, cricket, snake
- algae, fish, mosquito
- snake, frog, water weed

**Q2** Complete this food web for the pond life. Fill in the blank spaces.



**Q3** If mosquitos, crickets and algae disappeared, name all the animals that would have less food.

---



---

## Sample item illustrating performance at Proficiency Level 3.1

The next item (Figure 4.5) is illustrative of items at Proficiency 3.1. It is a multiple-choice item (Question 1 of the item set) that requires students to read a table summarising the results of an experiment and determine from the data which paper clip is the strongest. Students receive one mark in this item for nominating Option B as the correct answer.

It is located at 294 on the Scientific Literacy Scale. This means that Year 6 students whose test results corresponded with the national mean score would have an 87 per cent chance of responding correctly to it.

Students who are located at this Proficiency Level can generally describe the findings of an experiment in which one variable is being considered and interpret the results from the data where an element of comparison is required.

The item assesses Strand C of the scientific literacy domain.

**Figure 4.5** Item set, 'Paper Clips', illustrating the type of item at Proficiency Level 3.1.

**PAPER CLIPS**

Jill and Tracey wanted to make a mobile. They decided to use paper clips to make chains from which to hang decorations. The mobile would look something like this.



Each chain was made of 5 paper clips.

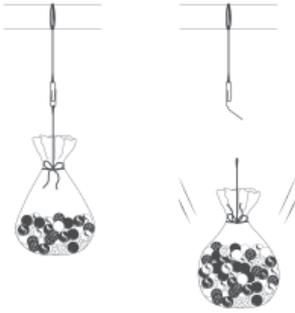


Jill and Tracey had been given four types of paper clip. Each paper clip type was made of a different metal.

Jill and Tracey designed an experiment to test the strength of the paper clips. They hung bags of marbles to each clip.

More marbles were added until the clip bent and the bag dropped.

They did the test three times for each type of paper clip. The results are shown in the table on the next page.



11

**Figure 4.5** (continued) Item set, 'Paper Clips', illustrating the type of item at Proficiency Level 3.1.

Number of marbles when the bag dropped			
Paper clip type	First Test	Second Test	Third Test
Holdall	52	50	49
Clipon	73	72	74
Plastico	41	44	42
Slippet	69	68	69

**Q1** Which paper clip type was the strongest to make the mobile?

- Holdall
- Clipon
- Plastico
- Slippet

**Q2** Jill and Tracey carried out the experiment three times for each type of paper clip.

Why did they do that?

---

---

## Sample item illustrating performance at Proficiency Level 2 and below

The item shown in Figure 4.6 is a short answer type that requires students to use their previous experiences to identify which of two pieces of sandpaper would feel rougher. Students who are at the top of Proficiency Level 2 or below would be expected to answer this and other questions at a similar level correctly more often than not. It can be inferred that such students can make a choice for a situation based on first-hand concrete experience or requiring the application of limited knowledge to this and similar situations.

In order to be awarded the one mark allocated to the item, the students had to provide the answer *'Heavy Duty'*.

**Figure 4.6** Item set, 'Sandpaper', illustrating the type of item in Proficiency Level 2 and below.

**SANDPAPER**

Sandpaper is used to smooth wood surfaces. There are different types of sandpaper. Sandpaper can be 'heavy duty' for very rough surfaces. 'Heavy duty' sandpaper has larger grains on its surface. Sandpaper can also be 'light' for surfaces that need very fine smoothing. 'Light' sandpaper has very small grains.



Heavy duty  
sandpaper - large grain



Light  
sandpaper - small grain

**Q1** Imagine you brushed your hand across the surface of each type of sandpaper. Which would feel rougher?

\_\_\_\_\_

**Q2** You are helping to make a wooden garden seat. It needs to be very smooth. You decide to use the 'heavy duty' sandpaper to start smoothing the wood and you then finish with the 'light' sandpaper. Why is this the best choice?

\_\_\_\_\_

\_\_\_\_\_

This item is located at 226 on the Scientific Literacy Scale. This means that Year 6 students whose test results corresponded with the national mean score would have a 92 per cent chance of responding correctly to it.

## Results on sample items

The percentage correct scores on each of the sample items for each State and Territory are shown in Table 4.1.

**Table 4.1** Selected results (percentage correct scores and mean scores on polytomously-scored items, expressed as percentage) on illustrative science items by Proficiency Level, by State and Territory

Item Figure	Proficiency Level	ALL %	NSW %	VIC %	QLD %	SA %	WA %	TAS %	NT %	ACT %
Q4 Mosquitos	3.3	24	29	25	20	25	22	23	20	30
Q2 Craters	3.2	49	56	46	48	48	46	51	46	56
Q3 Bush Pond	3.2	51	51	52	51	48	50	49	50	56
Q1 Paper Clips	3.1	83	83	84	83	81	81	81	79	85
Q1 Sandpaper	2 and below	88	90	90	88	88	87	87	82	88

It can be seen from Table 4.1 that the percentage correct for the items increases as the Proficiency Levels decrease. For example, 88 per cent of students across Australia responded correctly to the item representing Level 2 and below. The mean score for Question 4 of the Mosquitos item set is 24 per cent.

## Distributions of students within Proficiency Levels

The distributions of students within Proficiency Levels are shown in Figure 4.7 and summarised in Table 4.2.

At the national level, approximately 4.6 per cent (CI =  $\pm 0.4$  per cent) of students performed at Proficiency Level 2 and below. The assessment instruments were constructed with the expectation that most Year 6 students would demonstrate the understandings and skills of Proficiency Level 3 (see Appendix 1).

Roughly 58 per cent of students were proficient at Level 3.2 and above.

Table 4.2 shows that the Australian Capital Territory had the largest proportion of students working at Proficiency Level 4 or above (0.2 per cent), followed by New South Wales and Tasmania with 0.1 per cent.

The Northern Territory was estimated to have 10.7 per cent of its students working at Proficiency Level 2 and below and about 50 per cent of its students working at Proficiency Level 3.2 and above.

**Figure 4.7** Proficiency Levels on the Scientific Literacy Scale by State and Territory

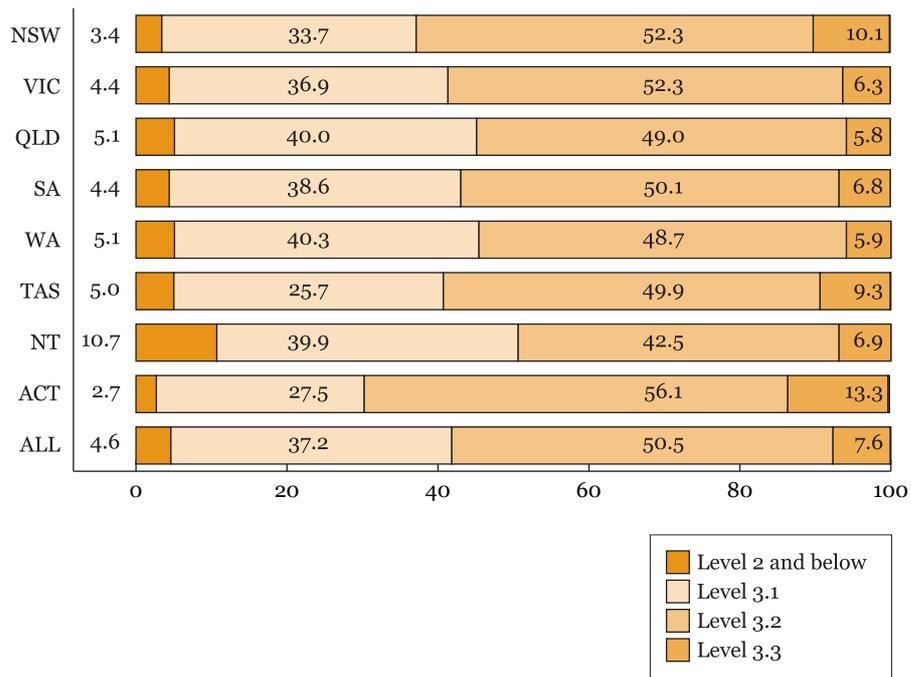


Table 4.2 shows the percentages of students in each of the States and Territories at Proficiency Levels 3.1, 3.2, 3.3 and the highest and lowest Proficiency Levels. It also shows the 95 per cent confidence intervals around the estimates for each Proficiency Level.

**Table 4.2** Percentages of students at each Proficiency Level on the Scientific Literacy Scale by State and Territory

State/Territory	Proficiency Level				
	2 and Below	3.1	3.2	3.3	4 and Above
NSW	3.4 (±0.8)	33.7 (±2.1)	52.6 (±2.4)	10.1 (±1.6)	0.1 (±0.2)
VIC	4.4 (±1.0)	36.9 (±2.7)	52.3 (±2.7)	6.3 (±1.2)	0.0 (±0.1)
QLD	5.1 (±0.9)	40.0 (±2.2)	49.0 (±2.0)	5.8 (±1.1)	0.0 (±0.0)
SA	4.4 (±1.2)	38.6 (±2.5)	50.1 (±2.3)	6.8 (±1.3)	0.0 (±0.1)
WA	5.1 (±1.0)	40.3 (±2.2)	48.7 (±2.3)	5.9 (±1.2)	0.0 (±0.0)
TAS	5.0 (±1.4)	35.7 (±2.9)	49.9 (±2.9)	9.3 (±1.8)	0.1 (±0.3)
NT	10.7 (±3.6)	39.9 (±5.6)	42.5 (±4.8)	6.9 (±2.8)	0.0 (±0.0)
ACT	2.7 (±1.1)	27.5 (±3.9)	56.1 (±4.8)	13.3 (±2.7)	0.2 (±0.5)
ALL	4.6 (±0.4)	37.2 (±0.9)	50.5 (±0.9)	7.6 (±0.5)	0.1 (±0.1)

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

The next section presents results by group (gender, Indigenous students and students from language backgrounds other than English) and results by school location.

## Distributions of student groups across Proficiency Levels

### Females and males

In Chapter 3, the mean scores and distributions for males and females nationally and by State and Territory were compared. There were no statistically significant differences, although males performed slightly better. Gender differences at the State and Territory level are explored further in the item-level results provided in Chapter 5.

Table 4.3 shows the distributions of results across the Proficiency Levels for males and females and confirms that there were no significant differences in performance.

**Table 4.3** Percentages of students at each Proficiency Level on the Scientific Literacy Scale by gender, State and Territory

State/Territory	Gender	Proficiency Level					Total
		2 and Below	3.1	3.2	3.3	4 and Above	
NSW	Male	3.3 (±1.2)	32.4 (±3.2)	52.9 (±3.2)	11.2 (±2.1)	0.2 (±0.3)	100
	Female	3.6 (±1.3)	35.1 (±3.2)	52.4 (±3.8)	8.8 (±2.3)	0.1 (±0.2)	100
VIC	Male	4.3 (±1.6)	35.9 (±3.6)	52.6 (±3.4)	7.1 (±1.7)	0.0 (±0.1)	100
	Female	4.3 (±1.5)	37.6 (±3.6)	52.4 (±3.7)	5.6 (±1.5)	0.0 (±0.1)	100
QLD	Male	4.3 (±1.2)	39.0 (±3.0)	49.1 (±3.2)	7.0 (±1.7)	0.0 (±0.1)	100
	Female	5.2 (±1.6)	41.2 (±3.5)	49.1 (±3.4)	4.6 (±1.5)	0.0 (±0.0)	100
SA	Male	4.4 (±1.4)	38.3 (±3.4)	49.5 (±4.2)	7.7 (±2.5)	0.0 (±0.1)	100
	Female	4.4 (±1.7)	38.8 (±3.8)	50.8 (±4.7)	5.9 (±2.1)	0.0 (±0.1)	100
WA	Male	4.4 (±1.4)	40.2 (±3.3)	48.4 (±3.1)	6.9 (±1.7)	0.0 (±0.1)	100
	Female	5.7 (±1.4)	40.5 (±3.0)	48.9 (±3.2)	4.9 (±1.4)	0.0 (±0.0)	100
TAS	Male	5.2 (±2.0)	33.8 (±4.4)	51.1 (±4.6)	9.9 (±2.9)	0.2 (±0.4)	100
	Female	4.8 (±1.9)	37.6 (±4.0)	48.9 (±4.5)	8.5 (±2.8)	0.2 (±0.4)	100
NT	Male	10.4 (±4.9)	42.3 (±6.9)	41.5 (±6.7)	5.8 (±3.7)	0.0 (±0.0)	100
	Female	10.7 (±5.3)	37.7 (±8.0)	43.7 (±6.8)	7.9 (±3.8)	0.0 (±0.0)	100
ACT	Male	2.3 (±1.5)	27.5 (±5.2)	57.9 (±6.1)	12.1 (±3.9)	0.2 (±0.6)	100
	Female	3.1 (±1.8)	27.4 (±6.0)	54.3 (±6.4)	14.9 (±3.8)	0.2 (±0.7)	100
ALL	Male	4.4 (±0.6)	36.5 (±1.3)	50.7 (±1.3)	8.3 (±0.7)	0.1 (±0.1)	100
	Female	4.8 (±0.6)	37.8 (±1.2)	50.5 (±1.2)	6.9 (±0.6)	0.0 (±0.1)	100

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

## Indigenous students

Sampling of schools was undertaken to enable reliable estimates of achievement by Indigenous students to be made at the national level.

Indigenous students' results relative to non-Indigenous results are shown in Table 4.4.

**Table 4.4** Mean scores and confidence intervals for Indigenous and non-Indigenous students on scientific literacy

Student Group	Mean Score	95 per cent confidence interval
Indigenous	338	±12.6
Non-Indigenous	404	±3.8
ALL	400	±3.8

Indigenous students, with a mean score of 338, did not perform as well as non-Indigenous students, with a mean score of 404. The differences are statistically significant at the 0.05 level. They are also educationally significant, as can be seen from the differences in scientific literacy proficiency shown in Table 4.5.

Nevertheless, approximately 2 per cent of Indigenous students achieved Proficiency Level 3.3. In addition, 30 per cent of the students were working at Proficiency Level 3.2 and above.

Table 4.5 shows that approximately 18 per cent of the Indigenous students were working at Level 2 or below compared with about 4 per cent of the non-Indigenous students.

**Table 4.5** Percentages of Indigenous and non-Indigenous students at each Proficiency Level on the Scientific Literacy Scale

Student Group	Percentage of Students in Proficiency Level					Totals
	2 and Below	3.1	3.2	3.3	4 and Above	
Indigenous	18.2 (±3.9)	51.9 (±5.9)	28.3 (±4.3)	1.7 (±1.3)	0.0 (±0.0)	100
Non-Indigenous	4.0 (±0.4)	36.3 (±1.0)	51.7 (±0.9)	7.9 (±0.5)	0.1 (±0.1)	100
Totals	4.6 (±0.4)	37.2 (±0.9)	50.5 (±0.9)	7.6 (±0.5)	0.1 (±0.1)	100

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

When the percentages of Indigenous and non-Indigenous students answering correctly each of the items in the assessment were examined, it appeared that Indigenous students performed relatively better in Strand A, which was concerned with the scientific process (that is, formulating or identifying questions and hypotheses, planning investigations and collecting evidence) than in Strands B and C.

The results also suggested that Indigenous students did worse (3.1 per cent) on items that required extended written responses (mean difference from non-Indigenous students of about 14 per cent) than on multiple-choice items (mean difference about 11 per cent).

## Students from language backgrounds other than English

About 12 per cent of the students in the sample (1,662) indicated that they had language backgrounds other than English. In Table 4.6, their results are compared with those of students whose language background was English.

**Table 4.6** Mean score and confidence intervals of scientific literacy scores by main language spoken by students at home

Student Group	Mean Score	95 per cent confidence interval
Home language not English	379	±9.0
Home language English	405	±3.8
Totals	400	±3.8

The difference in mean scores between the students whose home language was English and those whose home language was not English, is statistically significant and indicates that the former have a higher level of scientific literacy.

Most of the items were presented in contexts that required students to read some material (see examples earlier in this chapter) and were measures of reading skills as well as science skills.

Although a conscious effort was made during the construction of the assessment materials to minimise the reading load, some reading was necessary because the items were contextualised and this undoubtedly contributed to the difference in performance between the two groups.

The distributions of Proficiency Levels for these two groups is shown in Table 4.7. The difference in relative performance becomes more apparent when a comparison is made between the percentages of students achieving Level 3.2 and above. Approximately 60 per cent of those whose home language was English did so, compared with about 48 per cent of those whose home language was not English.

**Table 4.7** Percentage distribution of scientific literacy Proficiency Levels by main language spoken at home

Student Group	Percentage of Students in Proficiency Level					Totals
	Level 2 and Below	Level 3.1	Level 3.2	Level 3.3	Level 4 and Above	
Home language not English	7.9 (±1.5)	44.0 (±3.1)	43.6 (±2.9)	4.5 (±1.6)	0.0 (±0.0)	100
Home language English	4.1 (±0.4)	36.2 (±1.0)	51.6 (±1.0)	8.0 (±0.7)	0.1 (±0.1)	100
Totals	4.6 (±0.4)	37.2 (±0.9)	50.5 (±0.9)	7.6 (±0.5)	0.1 (±0.1)	100

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

## School location

Data on the geographic location of the schools attended by the sampled students have been used to report results in five categories determined by the MCEETYA Schools Geographic Location Classification. A similar system was used to report Australian results for PISA 2000.

Table 4.8 shows the distributions of mean scaled scores and 95 per cent confidence intervals.

It can be seen that the differences among four of the location categories are not statistically significant. However, the 3 per cent of students in the 'Remote Zone' category performed significantly worse in scientific literacy than any other location. Students attending schools in 'Major Urban Statistical Districts' achieved the highest mean scaled scores.

**Table 4.8** Distributions of mean scores by MCEETYA geolocation categories

MCEETYA Geolocation Category	% of Students	Mean Score	95 per cent confidence interval
Mainland State Capital City Regions	51	401	±5.4
Major Urban Statistical Districts	16	408	±10.6
Provincial City Statistical Districts	11	401	±13.6
Other Regional Areas	19	400	±10.6
Remote Zone	3	363	±19.0
Totals	100	400	±3.8

Table 4.9 shows the distribution of results across the Proficiency Levels for the five location categories. A comparison of mean scores indicates that schools in major urban areas had the highest percentages of students in the top levels.

**Table 4.9** Percentage distribution of scientific literacy Proficiency Levels by MCEETYA geolocation category

MCEETYA Geolocation Category	Percentage of Students in Proficiency Levels					Total
	Level 2 and Below	Level 3.1	Level 3.2	Level 3.3	Level 4 and Above	
Mainland State Capital City Regions	4.3 (±0.6)	37.6 (±1.5)	50.9 (±1.4)	7.1 (±0.9)	0.0 (±0.1)	100
Major Urban Statistical Districts	3.8 (±0.9)	33.4 (±2.0)	52.4 (±2.1)	10.3 (±1.3)	0.1 (±0.2)	100
Provincial City Statistical Districts	5.5 (±1.2)	39.1 (±2.6)	48.4 (±2.8)	7.0 (±1.5)	0.1 (±0.2)	100
Other Regional Areas	4.8 (±1.1)	37.8 (±2.4)	50.5 (±2.1)	7.0 (±1.3)	0.0 (±0.1)	100
Remote Zone	11.5 (±3.6)	40.2 (±5.8)	41.3 (±5.5)	7.3 (±3.1)	0.0 (±0.0)	100
Totals	4.6 (±0.4)	37.2 (±0.9)	50.5 (±0.9)	7.6 (±0.5)	0.1 (±0.1)	100

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



# Chapter 5

## The Standard for Year 6 Scientific Literacy

---

### Introduction

This chapter identifies the standard for Year 6 science literacy and describes students' proficiency in relation to the standard.

### The standard

A standard for science literacy has been established as part of the first cycle of national assessment to provide parents, educators and the community with a clear picture of the proficiency 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 expert group used their classroom experience and knowledge of the science curriculum in the various States and Territories to examine the test items from the national assessment. The crucial science 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. There are also Year 6 students who exceed the proficient standard (those who perform at Level 3.3 and above) and they show 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 benchmarks in literacy and numeracy have not been set for scientific literacy. Such benchmarks, defined as the critical level [of skill and understanding] without which a student will have difficulty making sufficient progress at school, are more suited to foundational areas such as reading, writing and numeracy where deficiencies will have a significant effect on students’ future learning and functioning in society.

The proficient standard will be the main reference point for monitoring science literacy in Australian primary schools over time. Every three years a new national Year 6 science assessment will be conducted to gauge whether student proficiency has improved.

Information about students’ performances in relation to the Year 6 standard from the first national science assessment is reported in the following section.

## Student groups achieving the Year 6 standard

The information in this section of the report draws on the distribution of students’ performances across Proficiency Levels as shown in Chapter 4. However, in this section attention is given to the percentage of students in Australian schools who have demonstrated proficiency by reaching Level 3.2 or better, i.e. the percentage of students who have demonstrated the Year 6 standard for science literacy.

Table 5.1 shows that approximately 58 per cent of students achieved the proficient standard. Approximately 70 per cent of students from the Australian Capital Territory reached the standard while about 49 per cent of students from the Northern Territory did so. It should be noted that, as with the mean scores in Chapter 4, when confidence intervals are taken into account, it is unlikely that there will be a significant difference between ACT, NSW or Tasmania in terms of the proportion of students achieving the proficient standard.

**Table 5.1** Percentages of students achieving the proficient standard or better, by State and Territory

State/Territory	Proficient Standard or Better
NSW	62.8 (±2.1)
VIC	58.7 (±2.5)
QLD	54.9 (±2.1)
SA	57.0 (±2.4)
WA	54.6 (±2.2)
TAS	59.3 (±2.9)
NT	49.4 (±5.5)
ACT	69.8 (±3.6)
ALL	58.2 (±0.9)

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

Table 5.2 shows the percentages of students achieving the proficient standard or better across various groups.

**Table 5.2** Percentages of students achieving the proficient standard or better across male/female; Indigenous/non-Indigenous; and Home language English/Home language not English groups

Group	Proficient Standard or Better
Male	59.1 (±1.3)
Female	57.4 (±1.2)
Home language English	59.7 (±1.0)
Home language not English	48.1 (±3.0)
Non-Indigenous	59.7 (±0.9)
Indigenous	29.8 (±4.5)
ALL	58.2 (±0.9)

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

The results in Table 5.3 are an extension of Table 5.1 and show the percentage of students who achieved or bettered each of the levels. The highest proficiency levels (Levels 3.3 & 4 and above) were achieved by approximately 7.7 percent of students.

**Table 5.3** Percentages of students at or above each Proficiency Level on the Scientific Literacy scale by State and Territory

State/Territory	Proficiency Level			
	3.1 or Above	Proficient 3.2 or Above	3.3 or Above	4 or Above
NSW	96.6 (±0.8)	62.8 (±2.1)	10.2 (±1.7)	0.1 (±0.2)
VIC	95.6 (±1.0)	58.7 (±2.5)	6.4 (±1.2)	0.0 (±0.1)
QLD	94.9 (±0.9)	54.9 (±2.1)	5.9 (±1.1)	0.0 (±0.0)
SA	95.6 (±1.2)	57.0 (±2.4)	6.9 (±1.3)	0.0 (±0.1)
WA	94.9 (±1.0)	54.6 (±2.2)	6.0 (±1.2)	0.0 (±0.0)
TAS	95.0 (±1.4)	59.3 (±2.9)	9.4 (±1.8)	0.1 (±0.3)
NT	89.3 (±3.6)	49.4 (±5.5)	6.9 (±2.8)	0.0 (±0.0)
ACT	97.3 (±1.1)	69.8 (±3.9)	13.6 (±2.8)	0.2 (±0.5)
ALL	95.4 (±0.4)	58.2 (±0.9)	7.7 (±0.5)	0.1 (±0.1)

Note: figures in parentheses refer to 95 percent confidence intervals

# Chapter 6

## Further Consideration of Results on Scientific Literacy Items

---

### Introduction

This chapter analyses the items in more detail and provides information about the strengths and weaknesses shown by students in the different States and Territories and across the strands and major concept areas of scientific literacy.

### Coverage of scientific literacy

The distribution of items across the assessment domain for scientific literacy (strands of the domain and major concept areas) is shown in Table 6.1. There were 70 items in the two pencil-and-paper tests and two practical tasks. Each student had to sit for one pencil-and-paper and one practical task.

**Table 6.1** Distribution of assessment items across the assessment domain for scientific literacy

Domain Aspect	Item Type and Number of Items			Totals
	Multiple-choice	Short Answer	Extended Response	
Distribution of items by strand				
Strand A	14	5	13	32
Strand B	4	2	10	16
Strand C	6	3	13	22
Totals	21	12	37	70
Distribution of items by major science conceptual area				
Life and Living	6	5	10	21
Earth and Beyond	5	1	8	14
Natural and Processed Materials	10	1	7	18
Energy and Change	3	3	11	17
Totals	24	10	36	70

## Major science concepts

The scientific literacy domain specified concepts in terms of major thematic areas rather than within traditional subject boundaries such as physics, chemistry or biology.

The thematic areas of scientific literacy are articulated in Table 6.1. They were considered to be more relevant to students at primary school and, according to PISA 2000, ‘to all people in their lives beyond school than the more traditional subject areas...’ (Lokan et al., 2001, p. 97).

It can be seen from Table 6.1 that the items were distributed relatively evenly across conceptual areas.

## Types of assessment items

The strands and major concepts of scientific literacy were assessed through a range of item types (Table 6.1). Thirty-six of the items were classified as being extended response, 24 as multiple-choice and ten as requiring short answers.

Almost all of the items were presented in item sets or units, with two or more items relating to each stimulus text and/or diagram.

The National Science Assessment markers commented on the relatively large number of students who did not respond to items that required extended answers.

Table 6.2 shows the percentages of students omitting responses by item type.

**Table 6.2** Percentages of students omitting responses by item type by State and Territory

State/Territory	Gender	Item Type and Percentage of Students Omitting Responses		
		Multiple-choice	Short Answer	Extended Response
NSW	Males	2.1	3.4	4.5
	Females	3.4	2.7	3.7
VIC	Males	3.1	4.5	5.9
	Females	3.1	4.3	6.2
QLD	Males	3.6	5.6	7.1
	Females	4.2	4.8	6.8
SA	Males	3.7	6.3	7.0
	Females	3.2	4.7	6.4
WA	Males	3.3	4.9	6.7
	Females	3.1	4.7	6.4
TAS	Males	2.7	6.0	6.9
	Females	2.9	4.6	5.6
NT	Males	7.3	8.2	11.1
	Females	6.4	7.6	7.8
ACT	Males	2.8	3.5	5.0
	Females	3.5	4.4	5.9
ALL	Males	3.2	5.1	6.4
	Females	3.3	4.4	6.0

In nearly all cases, the proportions of students omitting responses to extended response type items were double those omitting responses to multiple-choice type items.

The percentages omitting responses to short answer type items were generally higher than those omitting responses to multiple-choice items, but not as high as those omitting responses to extended response type items.

There is no evidence to suggest that gender was associated with these patterns, but it appears that there was a systematic effect throughout the scientific literacy scale. This raises the issue of the literacy demands created by the extended response item types and whether they affected the level of student engagement with the test items.

## The strands

The scientific literacy domain specified concepts in terms of three strands: A, B and C.

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

**Strand B** involved interpreting experimental data and required 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** involved using scientific understandings for describing and explaining natural phenomena, and for interpreting reports about phenomena.

Table 6.1 shows that 48 of the items assessed the process strands (32 Strand A and 16 Strand B) and 22 assessed the conceptual understanding strand (Strand C).

## Analysis of scientific literacy performance at the item level

Further analyses of the National Science Assessment data were undertaken at the level of the items to provide more diagnostic information for the States and Territories.

Table 6.3 shows the numbers of items for each State and Territory for which the percentage of students answering the item correctly was the highest in the sample.

The table also shows how close each State and Territory came to the top-scoring State and Territory, by providing information about the number of items for which the difference in the percentage of students scoring correctly was within 5 per cent of the top jurisdiction, between 5 and 10 per cent and greater than 10 per cent.

**Table 6.3** Numbers of items by State and Territory for which the results were the highest and within the given range from the State or Territory with the highest result

State/ Territory	Number of Items For Which State and Territory Had Highest Result	Number of Items Within Given Range From Highest Result			Totals
		Less Than 5%	Between 5-10%	Over 10%	
NSW	14	40	15	1	70
VIC	6	30	29	5	70
QLD	0	30	28	12	70
SA	0	34	32	4	70
WA	0	27	31	12	70
TAS	5	35	26	4	70
NT	1	15	32	22	70
ACT	44	23	3	0	70

There were 44 items for which students in the Australian Capital Territory had the highest percentage correct (or highest mean score on polytomously-scored items); 23 items for which the percentage correct was within 5 per cent of the State or Territory with the highest percentage correct; and three items for which the percentage correct was between 5 per cent and 10 per cent of the State or Territory with the highest percentage correct. There were no items for which the percentage correct was more than 10 per cent from the State or Territory with the highest percentage correct.

A useful way of interpreting the information in Table 6.3 would be to use the differences from the highest score as an indicator of relative strength. For example, items that are within 5 per cent of the highest percentage correct score could be considered to be a relative strength for a State or Territory. Items that are between 5 per cent and 10 per cent away are neither strength nor a weakness.

Items that are more than 10 per cent away from the top scoring State or Territory could indicate a relative weakness.

Of course, performance on one item is not a very reliable indicator, so it is important to look for trends in the results to get more substantive information about potential strengths and weaknesses in the curriculum.



## References

---

- Ball, S., Rae, I. & Tognolini, J. (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.
- Bigg, J.B. & Collins, K.F. (1982). Evaluating the quality of learning: The SOLO taxonomy. New York: Academic Press.
- Goodrum, D., Hackling, M. & Rennie, L. (2001). The status and quality of teaching and learning of science in Australian schools. Canberra: Department of Education, Training and Youth Affairs.
- Hackling, M.W., Goodrum, D. & Rennie, L. (2001). The state of science in Australian secondary schools. *Australian Science Teachers Journal*, 47(4), 6-17.
- Hackling, M.W. (2002). Review of curriculum frameworks; assessment frameworks; reporting policies and assessment policies. Unpublished correspondence with ACER.
- Laugksch, R.C. (2000). Scientific literacy: A conceptual overview. *Science Education*, 84(1), 71-94.
- Lokan, J., Greenwood, L. and Cresswell, J. (2001). 15-Up and Counting, Reading, Writing, Reasoning... How literate are Australia's students? Melbourne: Australian Council for Educational Research
- OECD Programme for International Student Assessment (1999). Measuring student knowledge and skills: a new framework for assessment. Paris: OECD.



# Appendix 1

## National Year 6 Primary Science Assessment Domain

---



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

Scientific literacy is important as 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 defined scientific literacy as:

*the capacity to use scientific knowledge, to identify questions (investigate)<sup>1</sup> 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 was adopted for the National Science Assessment.

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 Science Assessment as the sample testing methodology to be used allowed for assessments of students' ability to conduct investigations.

## Scientific literacy: the assessment domain

A scientific literacy assessment domain was developed based on the construct of scientific literacy and on an analysis of State and Territory curriculum and assessment frameworks. The assessment domain included a hierarchy of scientific literacy across three strands that were inclusive of the five elements of the PISA 2000 definition:

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 were clustered into three more holistic strands: Elements 2 and 3 and conducting investigations as Strand A, Elements 4 and 5 as Strand B and Element 1 as Strand C.

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

This process strand included posing questions or hypotheses for investigation or recognising scientifically investigable questions; planning investigations by identifying variables and devising procedures in which variables were 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, critiquing the trustworthiness of evidence and claims made by others, and communicating findings.*

This process strand included identifying, describing and explaining the patterns and relationships between variables in scientific data; drawing conclusions that were 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:** *demonstrating science understandings by describing and explaining natural phenomena, making sense of reports and making decisions.*

This conceptual strand included demonstrating conceptual understandings by being able to describe, explain and make sense of natural phenomena; understanding and interpreting reports such as television documentaries, newspaper or magazine articles or conversations related to scientific matters; and making decisions about scientific matters in students' own lives that might 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 assessment domain included a hierarchy describing increasingly complex scientific literacy skills and understandings at six levels, numbered from 1 to 6. The hierarchy was described in terms of three aspects:

- increasing complexity, from explanations that involve one aspect to several aspects and then 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 that which can be observed directly and involve abstract scientific concepts (abstract); and
- increasing complexity in descriptions of 'what' happened in terms of the objects and events, in explanations of 'how' it happened in terms of processes, and in explanations of 'why' it happened in terms of science concepts.

The level descriptions of the domain were linked to the SOLO (Structure of Observed Learning Outcomes) taxonomy (Biggs & Collis, 1982), which 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; and
- 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;
- ikonik - the world is represented as internal images;
- concrete - writing and other symbols are used to represent and describe the experienced world; and
- formal - the world is represented and explained using abstract conceptual systems.

The level descriptions of the domain therefore make 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.

Table A1.1 describes the components of the scientific literacy assessment domain.

**Table A1.1** Scientific literacy assessment domain

	SOLO taxonomy	Strands of scientific literacy		
		<b>Strand A</b> Formulating or identifying investigable questions and hypotheses, planning investigations and collecting evidence.	<b>Strand B</b> Interpreting evidence and drawing conclusions, critiquing the trustworthiness of evidence and claims made by others, and communicating findings.	<b>Strand C</b> Demonstrating science understandings by describing and explaining natural phenomena, making sense of reports and making decisions.
6	Abstract relational	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 types 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, and consistency between data and claims.	Explains complex interactions, systems or relationships using several abstract scientific concepts or principles and the relationships between them.
5	Abstract multi-structural	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.
4	Abstract unistrucltural	Identifies the variable to be changed, the variable to be measured and several variables to be controlled.  Uses repeated trials or replicates.	Calculates averages from repeat trials or replicates, plots line graphs where appropriate.  Conclusions summarise and explain the patterns in the data.  Is 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.
3	Concrete relational	Formulates scientific questions for testing and makes predictions.  Demonstrates awareness of the need for fair testing.  Makes simple standard measurements.  Records data as tables, diagrams or descriptions.	Displays data as tables or bar graphs, identifies and summarises patterns in science data.  Applies the rule by extrapolating or predicting.	Explains the relationships between individual events that have been experienced or reported and can generalise and apply the rule by predicting future events.
2	Concrete multi-structural	Given a question in a familiar context, identifies a variable to be considered.  Observes and describes or makes non-standard measurements and limited records of data.	Makes comparisons between objects or events observed.	Describes changes to, differences between or properties of objects or events that have been experienced or reported.
1	Concrete unistrucltural	Responds to the teacher's questions, observes and describes.	Describes what happened.	Describes an aspect or property of an individual object or event that has been experienced or reported.

# Scientific concepts in the National Science Assessment

A summary of the major scientific concepts found most widely in the various State and Territory curriculum documents was developed as part of the assessment domain (see Table A1.2).

These major concepts are broad statements of scientific understandings 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 hierarchy, which describes the typical developmental stages for scientific literacy, these conceptual statements were used as a guide for the development of assessment items.

It should be noted that because the National Science Assessment test instruments were constructed within the constraints of test length, it might not be practical to include all the listed concepts in instruments constructed for a specific testing cycle.

**Table A1.2** Major scientific concepts

Concepts	Examples
<p><b>Earth and Beyond</b></p> <p>Earth, sky and people: our lives depend on air, water and materials from the ground; the ways in which we live depend on landscape, weather and climate.</p> <p>The changing earth: the earth is composed of materials that are altered by forces within and upon its surface.</p> <p>Our place in space: the earth and life on it are part of an immense system called the universe.</p>	<ul style="list-style-type: none"> <li>• Features of weather, soil and sky and effects on me.</li> <li>• Changes in weather, weather data, seasons, soil landscape and sky (for example, moon phases), weathering and erosion, movement of the sun and shadows, bushfires, land clearing.</li> <li>• People use resources from the earth; need to use them wisely.</li> <li>• Rotation of the earth and night/day, spatial relationships between the sun, earth and the moon.</li> <li>• Planets of our solar system and their characteristics.</li> </ul>
<p><b>Energy and Change</b></p> <p>Energy and us: energy is vital to our existence and our quality of life as individuals and as a society.</p> <p>Transferring energy: interaction and change involve energy transfers; control of energy transfer enables particular changes to be achieved.</p> <p>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.</p>	<ul style="list-style-type: none"> <li>• Uses of energy, patterns of energy use and variations with time of day and season.</li> <li>• Sources, transfers, carriers and receivers of energy, energy and change.</li> <li>• Types of energy, energy of motion – toys and other simple machines – light, sound.</li> <li>• Forces as pushes and pulls, magnetic attraction and repulsion.</li> </ul>
<p><b>Life and Living</b></p> <p>Living together: organisms in a particular environment are interdependent.</p> <p>Structure and function: living things can be understood in terms of functional units and systems.</p> <p>Biodiversity, change and continuity: life on the earth has a history of change and disruption, yet continues from generation to generation.</p>	<ul style="list-style-type: none"> <li>• Living versus non-living.</li> <li>• Plant versus animal and major groups.</li> <li>• Major structures and systems and their functions.</li> <li>• Dependence on the environment: survival needs – food, space and shelter.</li> <li>• Change over lifetime, reproductions and lifecycles.</li> <li>• Interactions between organisms and interdependence, e.g. simple food chains.</li> <li>• Adaptation to physical environment.</li> </ul>
<p><b>Natural and Processed Materials</b></p> <p>Materials and their uses: the properties of materials determine their uses; properties can be modified.</p> <p>Structure and properties: the sub-structure of materials determines their behaviour and properties.</p> <p>Reactions and change: patterns of interaction of materials enable us to understand and control those interactions.</p>	<ul style="list-style-type: none"> <li>• Materials have different properties and uses</li> <li>• The properties of materials can be explained in terms of their visible sub-structure such as fibres.</li> <li>• Materials can change their state and properties.</li> <li>• Solids, liquids and gases.</li> </ul>



## Appendix 2 Sampling

---



# Sampling results

Sampling in the National Science Assessment was carried out in two stages. First, schools were selected at random with a probability proportional to the size of the Year 6 enrolment. Then, if there was only one Year 6 class, it was selected for participation in the study. If there was more than one Year 6 class, one of the classes was selected at random from within the school. It was assumed that there were 25 students in each class. ACER's latest available sampling frame was used to estimate the various enrolment sizes.

Stratification variables in selecting the sample were State/Territory (New South Wales, Victoria, Queensland, South Australia, Western Australia, Tasmania, Australian Capital Territory and Northern Territory); school sector (government, Catholic and independent); and school location based on five categories of the MCEETYA School Geographic Location classification.

Table A2.1 lists the number of schools excluded from the sample by category.

**Table A2.1** Schools excluded from the National Science Assessment

State/Territory	Small schools*	Involved in other ACER studies	Special schools‡	Remote schools	Total excluded schools†
NSW	241	52	2	8	303
VIC	137	56	7	0	200
QLD	195	21	4	59	279
SA	40	7	1	19	67
WA	78	28	48	57	211
TAS	16	0	9	2	27
NT	32	0	0	45	77
ACT	0	0	3	0	3
ALL	739	164	74	190	1 167

\* As the assessment design included a practical task that was a group activity requiring at least three students, schools with Year 6 populations of less than five students were excluded from the sample.

‡ Schools catering exclusively for students with special educational needs and hospital schools were excluded.

† A remote school was a school defined in the MCEETYA database as being located in *Remote Zone Very Remote Areas*.

# Sample achieved

## The main sample

Table A2.2 summarises the main sampling frame for the National Science Assessment.

**Table A2.2** School sampling frame for the National Science Assessment

State/ Territory	Schools				
	Final Population*	Target Sample‡	Non- participants†	Sample Achieved ^	School Participation Rate~
NSW	2 050	122	19	103	0.84
VIC	1 647	122	22	100	0.82
QLD	1 103	122	12	110	0.90
SA	566	130	15	115	0.89
WA	666	126	23	103	0.81
TAS	209	64	4	60	0.94
NT	62	32	9	23	0.72
ACT	101	44	8	36	0.82
ALL	6 404	762	112	650	0.85

\* Total number of schools minus the number of schools in the excluded population.

‡ The number of schools required to achieve the target sample of students. This was premised on the assumption that there would be 25 students per school. If the number of students in the selected school was less than 25, schools were combined to give class sizes of at least 25.

† These schools decided for one reason or another not to participate in the study. Some made it clear before the testing date to State/Territory liaison officers that they were not participating. Others did not return the scripts, even after being contacted by the liaison officers.

^ Sample of schools that returned scripts.

~ Ratio of achieved sample to target groups.

In all, 650 schools participated in the study, representing an unweighted response rate of approximately 85 per cent of selected schools.

Details of the student sampling frame are shown in Table A2.3.

The total number of students selected to participate in the study was 17,124. The actual number that submitted scripts for marking was 14,172, representing an unweighted response rate of 83 per cent.

## School-level student exclusions

Schools excluded 203 students from the assessment, representing approximately one per cent of the desired sample. This figure was consistent with PISA and TIMSS.

Students excluded at the school level fell within one or more of the following classifications:

- **Functional disability:** student has a moderate to severe permanent physical disability such that he or she cannot perform in the assessment situation.
- **Intellectual disability:** student has a mental or emotional disability and is cognitively delayed such that he or she cannot perform in the assessment situation. This includes students who are emotionally or mentally unable to follow even the general instructions of the assessment.
- **Language proficiency:** student is unable to read or speak sufficient English to overcome the language barrier in the testing situation. Typically, students who had received less than one year of instruction in English was excluded.
- **Other:** special situation occurring on the day of the assessment that precluded student from being able to participate in it.

## Population coverage

Table A2.3 describes the student target population of the States and Territories participating in the National Science Assessment. Further information on the target population and the implementation of the study's sampling standards is available in the technical report.

**Table A2.3** Student sampling frame for the National Science Assessment

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9	Column 10	Column 11
State/Territory	Students		Total Available Population	% of School-level Student Exclusions	In Desired Sample	In Target Sample	In Sample Achieved	Within-school Exclusions	% of Within-school Exclusions	Student Participation Rate
	In Total Enrolled Population	Excluded Because School was Excluded								
NSW	89 226	3 392	85 834	3.8	2 800	2 805	2 466	30	1.1	0.88
VIC	65 265	3 007	62 258	4.6	2 800	2 770	2 130	30	1.1	0.77
QLD	53 424	2 497	50 927	4.7	2 800	2 805	2 607	42	1.5	0.93
SA	19 297	629	18 668	3.3	2 800	2 807	2 032	28	1.0	0.72
WA	27 156	3 025	24 131	11.1	2 800	2 812	2 347	22	0.8	0.84
TAS	6 814	114	6 700	1.7	1 400	1 378	1 240	22	1.6	0.90
NT	2 853	658	2 195	23.1	700	706	496	12	1.7	0.70
ACT	4 783	17	4 766	0.4	1 050	1 041	854	17	1.6	0.82
<b>TOTALS</b>	<b>268 818</b>	<b>13 339</b>	<b>255 479</b>	<b>5.0</b>	<b>17 150</b>	<b>17 124</b>	<b>14 172</b>	<b>203</b>	<b>1.2</b>	<b>0.83</b>

Column 1: List of Australian States and Territories.

Column 2: Year 6 students enrolled at Australian schools according to the 2002 census.

Column 3: Students excluded as a consequence of their schools being excluded for one of the reasons specified in Section A2.2.

Column 4: Students available for the sampling frame (Column 2 totals minus Column 3 totals).

Column 5: Percentages of students unavailable because their schools were excluded from the population (Column 3/Column 2 X100 = %).

Column 6: Students needed to achieve the desired requirement of having 112 schools per jurisdiction and an estimated 25 students per school (2,800 students). This requirement was set in order to obtain 95 per cent confidence limits that were within plus or minus 0.1 standard deviations of the mean. The initial desired sample was reduced for the smaller States and Territories (ACT 42 schools, NT 28, TAS 56) so that the assessment program would not be too intrusive. This had the effect of increasing the standard error of the estimates for these States and Territories and increasing the width of the confidence limits around the estimates.

Column 7: Schools selected in the target sample were identified. If they had enrolments of less than 25 students, they were grouped with other schools to get a combined enrolment as close to 25 as possible. The target sample population was the number of students identified by the schools that had agreed to participate in the National Science Assessment as being enrolled in their Year 6 classes.

Column 8: Students who participated in the study: the number of students in the target sample less the number of students who were absent on the day of the assessment, less the number of students who were excluded at the discretion of the schools less the number of students attending schools who had agreed to participate but did not complete the assessment.

Column 9: Students excluded from the study at the discretion of their schools.

Column 10: Percentages of students unavailable because their schools had exempted them from the assessment (Column 9/Column 7 X100 = %).

Column 11: Student participation rate (Column 8/Column 7).

## Appendix 3

### Proficiency Levels, Assessment Domain Descriptors, Illustrative Items and Item Descriptors

---



**Table A3.1** Proficiency levels, assessment domain descriptors, illustrative items and item descriptors

Proficiency Level (Scaled Location)	Assessment Domain Descriptors	Descriptor: a student at this level may display skills like:	Illustrative Items and Item Descriptor
Level 4 and above (scale score >638)	<b>Strand A:</b> 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 scientific concept. Interprets abstract diagrams situated within an unfamiliar context.	No items functioned at this level.
	<b>Strand B:</b> identifies the variable to be changed, the variable to be measured and several variables to be controlled. Uses repeated trials or replicates.	Demonstrates awareness of the need for fair testing by explaining how specific variables must be controlled. Uses repeated trials and replicates in testing.	No items functioned at this level.
	<b>Strand C:</b> calculates averages from repeat trials or replicates, plots line graphs where appropriate. Summarises and explain the patterns in the data. Able to make general suggestions for improving an investigation, e.g. make more measurements.	Critiques investigations noting design flaws. Makes general suggestions for improving an investigation.	No items functioned at this level.
Proficiency Level (Scaled Location)	Assessment Domain Descriptors	Descriptor: a student at this level may display skills like:	Illustrative Items and Item Descriptor
Level 3.3 (scaled score 514-638)	<b>Strand A:</b> explains the relationships between individual events that have been experienced or reported and can generalise and apply the rule by predicting future events.	Applies knowledge of relationships to explain reported phenomenon.	Item 65 [Craters] Extrapolates from experimental evidence to describe a different environment (multiple variables).
	<b>Strand B:</b> formulates scientific questions for testing and makes predictions. Demonstrates awareness of the need for fair testing. Makes simple standard measurements. Records data as tables, diagrams or descriptions.	Demonstrates an awareness of the principles of conducting an experiment and controlling variables. Proposes suitable method for fair collection of data.	Item 68 [Parachute] Explains the aim of an investigation with regard to multiple variables.
	<b>Strand C:</b> displays data as tables or bar graphs, identifies and summarises patterns in science data. Applies the rule by extrapolating or predicting.	Describes key features of a collected set of data and can predict outcome of next event in series. Extrapolates from an observed pattern to describe an expected outcome or event.	Item 53 [School electricity] Identifies and extrapolates patterns in the data provided.

**Table A3.1** Proficiency levels, assessment domain descriptors, illustrative items and item descriptors (continued)

Proficiency Level (Scaled Location)	Assessment Domain Descriptors	Descriptor: a student at this level may display skills like:	Illustrative Items and Item Descriptor
Level 3.2 (scaled score 389-513)	<b>Strand A:</b> explains the relationships between individual events that have been experienced or reported and can generalise and apply the rule by predicting future events.	Interprets information in a contextualised report by application of relevant science knowledge. Uses observed data and personal experience and applies rule to describe expected outcome.	Item 59 [Exercise] Interprets information in a contextualised report by application of relevant scientific knowledge.
	<b>Strand B:</b> formulates scientific questions for testing and makes predictions. Demonstrates awareness of the need for fair testing. 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.	Item 30 [Toy Train] Suggests questions for testing.
	<b>Strand C:</b> displays data as tables or bar graphs, identifies and summarises patterns in science data. Applies the rule by extrapolating or predicting.	Interprets diagrams and graphical data situated in a common or familiar context. Draws conclusions and makes comparisons of scientific data.	Item 18 [Mosquito] Demonstrates an understanding of what is required for fair testing.
Proficiency Level (Scaled Location)	Assessment Domain Descriptors	Descriptor: a student at this level may display skills like:	Illustrative Items and Item Descriptor
Level 3.1 (scaled score 264-388)	<b>Strand A:</b> explains the relationships between individual events that have been experienced or reported and can generalise and apply the rule by predicting future events.	Selects appropriate reason to explain reported observation related to personal experience. Identifies the relationship between events that have been observed or experienced.	Item 16 [Camping Holiday] Selects appropriate reason to explain reported observation related to personal experience.
	<b>Strand B:</b> formulates scientific questions for testing and makes predictions. Demonstrates awareness of the need for fair testing. Makes simple standard measurements. Records data as tables, diagrams or descriptions.	Identifies the variable to be measured or controlled.	Item 54 [Bean Plants] Identifies variables to be measured and/or controlled.
	<b>Strand C:</b> displays data as tables or bar graphs, identifies and summarises patterns in science data. Applies the rule by extrapolating or predicting.	Describes the findings of an experiment in simple terms, focusing on one variable. Interprets simple data set requiring an element of comparison.	Item 19 [Mosquito] Conclusions summarise patterns in the data.

**Table A3.1** Proficiency levels, assessment domain descriptors, illustrative items and item descriptors (continued)

Proficiency Level (Scaled Location)	Assessment Domain Descriptors	Descriptor: A student at this level may display skills like:	Illustrative Items and Item Descriptor
Level 2 and below (scaled score $\leq 263$ )	<b>Strand A:</b> 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, or requiring the application of limited knowledge.  Identifies the difference between properties that have been experienced.	Item 28 [Sandpaper] Identifies the difference between properties that have been experienced.
	<b>Strand B:</b> given a question in a familiar context, identifies a variable to be considered, 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.	Item 67 [Parachutes] Identifies the variable to be considered.
	<b>Strand C:</b> makes comparisons between objects or events observed.	Identifies simple patterns in data and/or interprets a data set containing some interrelated elements.	Item 49 [Erosion] Identifies and summarises pattern in science data.



