AUSTRALIAN CURRICULUM, ASSESSMENT AND REPORTING AUTHORITY

NAP Science Literacy Assessment Framework

March 2018

Contents

Document history

Date	Version	Activity	Author
19 May 2017	0.1	Draft for review	Dr Sofia Kesidou (consultant) Dr Rainer Mittelbach (ACARA) Dr Goran Lazendic (ACARA)
16 August 2017	0.2	Updated draft, incorporating first round of feedback from NAP Science Literacy Expert Panel	ACARA
8 September 2017	0.3.1	Updated draft, incorporating second round of feedback from NAP Science Literacy Expert Panel	ACARA
15 September 2017	0.3.2	Updated draft, incorporating glossary and references	ACARA
22 November 2017	0.4	Draft for ACARA Board, incorporating feedback from jurisdictional review via NADAR	ACARA
7 March 2018	0.5	Final draft for ACARA Board, incorporating feedback from jurisdictional review via NADAR (section 2.4.3 footnote; section 3.3 para 1 moved to 3.1 para 3; section 4.1 constructed response new point 2; section 5.1 new paras 3, 5, 6; reference list includes NAPSL reports).	ACARA
7 March 2018	1.0	Version 0.5 approved by ACARA Board	ACARA

CHAPTER ONE: OVERVIEW

1.1 Background

The National Assessment Program - Science Literacy (NAP-SL) is one of a suite of three national sample assessments — together with civics and citizenship, and information and communication technology (ICT) literacy — which are conducted in three-year cycles with stratified random samples of students. The results contribute to an understanding of student progress towards the achievement of the Educational Goals for Young Australians specified in the Melbourne Declaration. The first science literacy assessment was conducted in 2003. The assessment has been repeated with a new sample of Year 6 students every three years to identify trends over time.

In July 2016, the Education Council decided to extend the NAP-SL to Year 10 students from 2018. The purpose of this decision was to reinforce the need to assess the science literacy progress of Australian students using assessments that are closely aligned with the Australian Curriculum, in addition to using outcomes of the international assessments and surveys. Until now, the Programme for International Student Achievement (PISA) has been the primary national measure of performance for science literacy among secondary school students. Australian students also participate in the Trends in International Mathematics and Science Study (TIMSS) which includes assessment of Year 8 students' knowledge of science curriculum.

The NAP-SL is based on an assessment framework that predates the Australian Curriculum. Additional developments — the introduction of the Year 10 science literacy assessment, the move to online assessment administration in 2015, and the growth in innovative assessment approaches often coupled with the use of computer technologies — make it necessary to create a new framework for assessing science literacy in the 2018 assessment cycle.

This document (*Assessment Framework NAP-Science Literacy 2018*, herein referred to as the framework) contains specifications for both the Year 6 and the Year 10 science literacy assessments. Drawing on the developments discussed in the previous paragraph, the recommendations extend the aspects of science literacy and the depth with which it is assessed, while maintaining the underlying construct of the assessment to enable effective historical comparison.

The framework stipulates and describes the content to be assessed, the cognitive engagement that is expected of students, and the types of assessment tasks and questions to be included in the assessment. The development of the framework was guided by the Australian Curriculum: Science and was informed by research in science education, research in assessment and measurement, best practices in assessment, and international assessment frameworks. The framework was further refined through feedback from reviews by a panel of state and territory and ACARA experts as well as through feedback from a pilot item writing project that aimed to implement key framework recommendations.

1.2 What does the NAP Science Literacy assessment measure?

The NAP-SL measures science literacy as defined in the Australian Curriculum: Science, that is the ability to: 'use scientific knowledge, understanding, and inquiry skills to identify questions, explain science phenomena, solve problems and draw evidence-based conclusions in making sense of the world, and to recognise how understandings of the nature, development, use and influence of science help us make responsible decisions and shape our interpretations of information' (https://www.australiancurriculum.edu.au/f-10-curriculum/science/glossary/?letter=S).

The construct of science literacy measured in the NAP-SL is also informed by the specific aims of the Australian Curriculum: Science:

- the understanding of important science concepts and processes, the practices used to develop scientific knowledge, science's contribution to society, and society's influence on science from a range of cultures;
- the ability to think and act in a scientific way;
- the ability to make informed decisions about local, national and global issues (https://www.australiancurriculum.edu.au/f-10-curriculum/science/rationale/).

The definition of science literacy in the NAP-SL is consistent with recent definitions of science literacy internationally. For example, PISA 2015 defines science literacy as 'the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen' (OECD, 2016[1]). PISA's definition includes being able to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

Similarly, a recent report from the US National Academies of Sciences, Engineering, and Medicine describes the following aspects of science literacy[2]:

- content knowledge (e.g. knowledge of basic facts, concepts, and vocabulary);
- the understanding of scientific practices (e.g. formulation and testing of hypotheses, probability/risk, and causation versus correlation);
- the understanding of science as a social process (e.g. the criteria for the assignment of expertise, the role of peer review, the accumulation of accepted findings, the existence of venues for discussion and critique, and the nature of funding and conflicts of interest).

---

[1] OECD (2016). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*. Paris: OECD Publishing.
[2] National Academies of Sciences, Engineering, and Medicine. (2016). *Science Literacy: Concepts, Contexts, and Consequences.* Washington, DC: The National Academies Press.

In 2009, a student survey was introduced into NAP–SL that served to gather information about Year 6 students' attitudes to, and interests in, science and their science experiences in school. This addition to the assessment program will be continued in 2018. The survey will be expanded to include questions that are of particular relevance for Year 10 students with the aim to gather information about students' perception of the importance of science for their future career paths and their potential intentions of pursuing careers in science, technology, engineering and mathematics (STEM)-related fields.

1.3 Organisation of the framework

The framework includes the following chapters:

Chapter 1: **Overview** provides background information on the Assessment Framework NAP-SL 2018.

Chapter 2: **Content Dimension** describes the content domains — the specific subject matter to be covered in the assessment.

Chapter 3: **Cognitive Dimension** describes the targeted thinking skills and intellectual processes as students respond to the assessment tasks.

Chapter 4: **Item Types** describes the types of assessment items and response formats that would be required to capture the variability and different levels of complexity of performance discussed in chapters 2 and 3.

Chapter 5: **Assessment Structure and Reporting** outlines the design of the assessment and implications of the new assessment framework for reporting NAP-SL results.

CHAPTER TWO: CONTENT DIMENSION

The content dimension defines the content domains — the specific subject matter covered in the assessment.

The content dimension of the NAP-SL 2018 assessment framework is organised according to the strands of the Australian Curriculum: Science. The content of these strands guides the definition of the content to be covered in the NAP-SL 2018 assessment.

The following section, *Content in the Australian Curriculum: Science*, describes each of the strands in the Australian Curriculum. They give an overview of the topic areas and concepts that will be covered in the 2018 assessments.

Following a general description of the strand content is an outline of the proposed organisation of the content for the NAP-SL Year 6 and Year 10 assessments in content domains and the target percentages of items to be allocated to each domain.

The development of content sequences is then discussed as a means to specify the content in detail.

2.1 Content in the Australian Curriculum: Science

The Australian Curriculum: Science includes three strands: *Science Understanding*, *Science as a Human Endeavour* and *Science Inquiry Skills*[3].

2.1.1 Science Understanding

Science Understanding includes the facts, concepts, principles, laws, theories and models that have been established by scientists over time.

The *Science Understanding* strand comprises four sub-strands.

Biological sciences

This sub-strand is concerned with understandings related to living things. Key concepts include: a diverse range of living things have evolved on Earth over hundreds of millions of years; living things are interdependent and interact with each other and their environment; the form and features of living things are related to the functions that their body systems perform.

Chemical sciences

This sub-strand is concerned with understandings related to the composition and behaviour of substances. Key concepts include: the chemical and physical properties of substances are determined by their structure at an atomic scale; substances

[3] The following paragraphs describing the strands, sub-strands and their contents are largely based on descriptions in the Australian Curriculum: Science (see https://www.australiancurriculum.edu.au/f-10-curriculum/science/structure/).

change and new substances are produced by rearranging atoms through atomic interactions and energy transfer.

Earth and space sciences

This sub-strand is concerned with Earth's dynamic structure and its place in the cosmos. Key concepts include: Earth is part of a solar system that is part of a larger universe; Earth is subject to change within and on its surface over a range of timescales as a result of natural processes; human use of resources has an impact on Earth's systems on a local and global level.

Physical sciences

This sub-strand is concerned with understanding the nature of forces and motion, and matter and energy. Key concepts include: forces affect the behaviour of objects; energy can be transferred and transformed from one form to another.

2.1.2 Science as a Human Endeavour

Science as a Human Endeavour includes understandings about the development of science as a unique way of knowing and doing, and the importance of science in contemporary decision-making and problem-solving.

The *Science as a Human Endeavour* strand comprises two sub-strands.

Nature and development of science

Key concepts within this sub-strand include: science involves the construction of explanations based on evidence, and science knowledge can be changed as new evidence becomes available; current knowledge has developed over time through the actions of many people; developments in science affect technology, and developments in technology affect science.

Use and influence of science

Key concepts within this sub-strand include: science knowledge and applications affect people's lives; science can be used to inform decisions and actions but, in making decisions about science applications, social implications must also be considered; scientific research is itself influenced by the needs and priorities of society.

2.1.3 Science Inquiry Skills

Science Inquiry Skills is concerned with the practices used to develop scientific knowledge, including questioning, planning and conducting experiments and investigations, collecting and analysing data, drawing critical, evidence-based conclusions, and evaluating and communicating results.

The *Science Inquiry Skills* strand comprises five sub-strands.

Questioning and predicting

Key abilities include identifying and constructing questions that can be investigated scientifically, proposing hypotheses and suggesting possible outcomes.

Planning and conducting

Key abilities include making decisions about how to investigate or solve a problem and carrying out an investigation, including the collection of data.

Processing and analysing data and information

Key abilities include representing data in meaningful and useful ways, identifying trends, patterns and relationships in data and using this evidence to justify conclusions.

Evaluating

Key abilities include considering the quality of available evidence and the merit or significance of a claim, proposition or conclusion with reference to that evidence.

Communicating

Key abilities include conveying information or ideas to others through appropriate representations, text types and modes.

2.2 NAP-SL 2018 content domains and target percentages

2.2.1 Year 6 Assessment

Tables 2.1–2.3 show the proposed content domains, sub-domains and target percentages of assessment items for domains/sub-domains for the Year 6 assessment.

Table 2.1 Content domains and target percentages in the Year 6 assessment

Content domains	Target percentage
Science Understanding	50%
Science as a Human Endeavour	15%
Science Inquiry Skills	35%

Table 2.2 Content domains and sub-domains in the Year 6 assessment

Science Understanding	Science as a Human Endeavour	Science Inquiry Skills
Biological sciences	Nature and development of science	Questioning and predicting
Chemical sciences	Use and influence of science	Planning and conducting
Earth and space sciences		Processing and analysing data and information
Physical sciences		Evaluating
		Communicating

Table 2.3 Target percentages for Year 6 *Science Understanding* sub-domains

Science Understanding	**Target percentage**
Biological sciences	25%
Chemical sciences	25%
Earth and space sciences	25%
Physical sciences	25%

The Year 6 assessment is organised around three content domains — *Science Understanding, Science as a Human Endeavour* and *Science Inquiry Skills* — that align with the organisation and content of the Australian Curriculum: Science (see previous section). The concepts outlined from Foundation to Year 6 in the Australian Curriculum: Science in these domains — further specified in content sequences (see section *Development of Content Sequences*) — will comprise the content space of the Year 6 assessment.

Items written for the content domain *Science as a Human Endeavour* may assess applications of science to students' everyday lives and to society, or they may assess students' understanding of the nature and development of science (e.g. students' understanding of the nature of scientific predictions, tests and evidence). *Science as a Human Endeavour* may also serve as a context for assessment items related to the *Science Understanding* and *Science Inquiry Skills* content domains. However, only those items that explicitly assess students' understanding related to the nature, development and applications of science should be classified within *Science as a Human Endeavour.* Not every item that has something to do with an application of science or otherwise uses *Science as a Human Endeavour* as a context.

The recommended target percentages of assessment items for the content domains in NAP-SL 2018 are broadly consistent with the previous NAP-SL assessments. The *Science Understanding* domain covers similar content as the strand *Applies Conceptual Understanding* in the previous assessment frameworks, which was also allocated 50% of assessment items. The *Science Inquiry Skills* content domain and the nature of science component of *Science as a Human Endeavour* covers similar content as the strands *Experimental design and data gathering* and *Interpreting experimental data* in the previous assessment frameworks, which together were allocated 50% of assessment items. The previous NAP-SL frameworks did not explicitly include content related to the development of science or the use and influence of science components of *Science as a Human Endeavour*.

The recommended target percentages of assessment items for the *Science Understanding* sub-domains are consistent both with the equivalent percentages of content statements in each sub-stand of the Australian Curriculum and the equivalent percentages in the previous NAP-SL assessment frameworks.

The distribution of items across sub-domains of Year 6 *Science as a Human Endeavour* and *Science Inquiry Skills* content domains may vary but will need to provide the coverage of all sub-domains. It is expected that most of the assessment items allocated to *Science Inquiry Skills* will target the *Planning and conducting* and *Processing and analysing data and information* sub-domains.

2.2.2 Year 10 Assessment

Tables 2.4–2.7 show the proposed content domains, sub-domains and relevant target percentages of items for the Year 10 assessment.

Table 2.4 Content domains and target percentages in the Year 10 assessment

Content domains	Target percentage
Science Understanding	50%
Science as a Human Endeavour	15%
Science Inquiry Skills	35%

Table 2.5 Content domains and sub-domains in the Year 10 assessment

Science Understanding	Science as a Human Endeavour	Science Inquiry Skills
Biological sciences	Nature and development of science	Questioning and predicting
Chemical sciences	Use and influence of science	Planning and conducting
Earth and space sciences		Processing and analysing data and information
Physical sciences		Evaluating
		Communicating

Table 2.6 Target percentages for Year 10 *Science Understanding* sub-domains

Science Understanding	Target percentage
Biological sciences	25%
Chemical sciences	25%
Earth and space sciences	25%
Physical sciences	25%

Table 2.7 Target percentages for Year 10 *Science as a Human Endeavour* sub-domains

Science as a Human Endeavour	Target percentage
Nature and development of science	50%
Use and influence of science	50%

The Year 10 assessment is organised around three content domains — *Science Understanding, Science as a Human Endeavour and Science Inquiry Skills* — that align with the organisation and content of the Australian Curriculum: Science (see previous section). The concepts outlined in Years 7–10 in the Australian Curriculum: Science, further specified in content sequences, will comprise the content covered in the Year 10 assessment.

The target percentages of assessment items for the content domains *Science Understanding* and *Science Inquiry Skills* in Table 2.4 may suggest an unbalanced emphasis on science knowledge vs. science practices in the assessment. Table 2.4 lists the recommended targets where each domain is the primary focus of the assessment. In addition to items that focus explicitly on science inquiry, students will be engaged in science practices as they use science knowledge to respond to items and tasks classified under the *Science Understanding* domain, such as those that ask students to use knowledge to make predictions, construct explanations, create and use models, etc. Hence, students will be engaged in science practices in a larger proportion of items than the 35% listed in Table 2.4.

Similarly, the percentage of items in which students will be engaged in thinking about use and influence of science will be greater than the 15% listed in Table 2.4. In addition to items that explicitly assess content on *Use and influence of science*, some items within the *Science Understanding* domain will require students to use science knowledge by applying it to societal issues. For example, an item within the *Science as a Human Endeavour* domain may ask students to consider methods of waste management and how they can affect the environment, while an item within the *Science Understanding* domain may ask students to consider how the properties of materials affect waste management.

The expected target percentages of items for the *Science Understanding* sub-domains in Year 10 reflect the intent of the Australian Curriculum, which places equal emphasis on the science disciplines.

It is expected that most of the assessment items allocated to *Science Inquiry Skills* will target the *Planning and conducting, the Processing and analysing data and information* and *the Evaluating* sub-domains.

2.3 Development of content sequences

The Australian Curriculum: Science includes a series of content descriptions that outline the expectations of what is to be taught to all students from Foundation to

Year 10. The content in the *Science Understanding* strand is described by year level, while the content in the *Science Inquiry Skills* and *Science as a Human Endeavour* strands is described in two-year bands.

Principled assessment design requires specifications of the construct(s) to be assessed in language detailed enough to guide item and test development. In addition, the construct(s) need to be articulated at different levels of sophistication to support the development of items that give all students the opportunity to respond. The Australian Curriculum: Science content descriptions (in particular those at secondary level) are broad statements that include multiple abstract ideas. Assessment items written to target content at this level of sophistication may not give all students the opportunity to show what they know and are able to do. Many of them will be on the way but may not yet have reached these high expectations.

Designing effective assessment for the NAP-SL requires the articulation of content sequences for the science content outlined in Australian Curriculum: Science that describe the essential elements of the construct(s) in sufficient detail to guide item development and illustrate how the construct(s) can be assessed at different levels of sophistication[4]. This is particularly the case with knowledge statements related to the *Science Understanding* domain.

Proposed content sequences for the three NAP-SL content domains for Years 6 and 10 have been developed, with the purpose of providing information to item developers[5]. The content sequences for the *Science Understanding* domain describe the essential elements for each key concept in the domain and, to the extent possible, articulate increasingly more complex ways of thinking about the concept. Typically, this starts with the more phenomenological (and often intuitive) understandings expected in primary school, progresses through intermediate understandings that make aspects of the concepts that are at odds with students' everyday ideas and misconceptions explicit, before finally advancing to the abstract interconnected, multifaceted understandings that are depicted in the Australian Curriculum: Science content descriptors.

The content descriptions in *Science as a Human Endeavour* and in *Science Inquiry Skills* articulate the intended construct in simple language and sufficient detail to guide item development. Therefore, the content sequences in this domain can be

---

[4] Content sequences are curriculum-based, but they enable assessment guidelines to go beyond specifying a list of key concepts and articulating exactly what about these ideas is important to assess. They are different from learning progressions, which are empirically validated descriptions of successively more sophisticated ways of thinking about an important knowledge or ability domain that follow one another as students learn about a topic over a broad span of time (Wilson, 2012).

[5] Content sequences are not included in this document. The content sequences are intended to break down the broad and sometimes complex Australian Curriculum Content Descriptions into narrower and more specific statements. Where appropriate, they are ordered in a logical sequence of increasing sophistication. Content sequences are not linked to levels of achievement, and are intended to provide guidance for the item developers only. The content sequences will be retained as a separate document and will not be published. They will also need to be revisited once the empirical data from the 2018 study become available.

regarded as the progression of the content descriptions in the Australian Curriculum: Science.

The development of the content sequences for *Science Understanding* broadly followed these steps:

- The conceptual space outlined in each sub-domain was divided into sub-topics, each representing a key concept within the space. Table 2.8 shows the key concepts within *Science Understanding* for which a content sequence was developed.
- Relevant Australian Curriculum: Science content descriptions were unpacked into their constituent aspects.
- Other ideas necessary for understanding the key concept were identified and different levels of sophistication of the concept were articulated. This process was informed by the year level descriptions and achievement standards in the Australian Curriculum: Science. Publications that define content sequences for science literacy, such as the American Association for the Advancement of Science (AAAS) Atlas of Science Literacy[6], were also consulted in this process.
- Research on student learning, learning progressions, and misconceptions was consulted in developing intermediate understandings, especially those that make aspects of the concepts that are at odds with students' everyday, less-sophisticated ideas explicit. This research rarely provides precise guidance about what progression should look like, but it does provide domain-specific as well as domain-general constraints on content sequence design.

Table 2.8 *Science Understanding* content sequences

Biological sciences	Chemical sciences	Earth and space sciences	Physical sciences
Interdependence of life Flow of matter and energy in ecosystems Multi-cellular systems DNA and inherited characteristics Diversity and evolution	Matter – structure, properties and changes	Earth in space Earth structure and processes Earth's resources and geochemical cycles	Forces and motion Energy forms, transfer and conservation

[6] AAAS (2001, 2007). *Atlas of Science Literacy.* Vols. 1&2. Washington DC: AAAS.

The content sequences are to be taken as comprehensive descriptions of the science content to be included in the NAP-SL 2018 assessment. In a few instances, the terms 'such as' or 'for example' are used to denote suggestions. In such cases, the examples to include in the assessment are not necessarily limited to these suggestions.

In order to understand the intent and use of content sequences, the following should also be kept under consideration:

- The statements in the *Science Understanding* and *Science as a Human Endeavour* content sequences are phrased as propositions that express ideas about science. Chapter 3 describes the development of performance expectations[7] as crosses between content statements and cognitive processes, which can more directly guide the development of assessment tasks.
- The statements in each content sequence (or within each constituent aspect in the sequence) are listed as far as possible from lower to higher levels of sophistication.
- Some assessment items will draw on more than one statement or content sequence.
- Retention of knowledge and abilities is assumed from primary to high school. A number of items will be common in both the Year 6 and Year 10 assessments to enable student performance from these year levels to be placed on the same scale. It is recommended that link items should target science ideas/skills from a Foundation to Year 10 content sequence that is already explicit in the Australian Curriculum: Science, and in particular, in the most sophisticated descriptions within the Foundation to Year 6 progression. It is assumed that exposure to the ideas/skills at higher levels of sophistication and greater number of contexts in high school would lead to a larger proportion of students mastering the ideas/skills.
- The content sequences describe in its entirety what is to be assessed in the NAP-SL 2018 assessment. However, they should not be interpreted as a complete description of the science curriculum that should be taught.

[7] Empirical evidence from the 2018 tests will be used to construct the Science Literacy learning progressions that will describe the development of key aspects of curriculum content and cognitive domains outlined in the proposed assessment framework. This work will be completed in conjunction with the extension of the current NAP Science Literacy scale and proficiency levels and standards to Year 10 in the bridging year (2018), and will then be further refined in 2021.

2.4 Key ideas, cross-curriculum priorities and general capabilities

2.4.1 Key ideas

In the Australian Curriculum: Science, there are six key ideas that represent key aspects of the scientific view of the world and transcend disciplinary boundaries (*Patterns, Order and Organisation, Form and Function, Stability and Change, Scale and Measurement, Matter and Energy,* and *Systems*). Year level descriptions and achievement standards help anchor these key ideas in specific content descriptions and can inform the context of questions in the NAP-SL as well as illustrate opportunities for assessing content in more integrated ways.

For example, the Year level description[8] that relates to the content statement 'electrical energy can be transferred and transformed in electrical circuits and can be generated from a range of sources' suggests that students may be asked to view electric circuits as systems through which energy flows and which can be scaled up, linking their understanding of circuits to electricity generation grids.

2.4.2 Cross-curriculum priorities

The Australian Curriculum includes three Cross-Curriculum Priorities (*Aboriginal and Torres Strait Islander Histories and Cultures, Asia and Australia's Engagement with Asia,* and *Sustainability*), with *Sustainability* the most relevant priority for the learning area of Science.

In the NAP-SL assessment, key concepts of *Sustainability* may be assessed in the context of the content sequences *Interdependence of life, Flow of matter and energy in ecosystems* and *Earth's resources and geochemical cycles*. More broadly, *Sustainability* may provide contexts for assessing specific understandings of chemical, biological, physical and Earth and space systems, the interconnectedness of Earth's biosphere, geosphere, hydrosphere and atmosphere, and inquiry skills embedded in investigations of such systems. *Sustainability* may also provide contexts for assessing understandings related to *Science as Human Endeavour – Use and influence of science*, such as that science knowledge and skills can be used to predict possible effects of human activity and to develop management plans that minimise these effects; or, that science provides the basis for decision-making in society, and these decisions can impact on the Earth system.

With regard to the priority *Aboriginal and Torres Strait Islander Histories and Cultures*, Aboriginal and Torres Strait Islander peoples' longstanding scientific knowledge traditions may provide contexts for assessing students' understanding of science concepts within *Science Understanding* or their understanding of *Science as a Human Endeavour*, for example that science knowledge can develop through the contributions of people from a range of cultures.

---

[8] YLD: 'They learn about transfer and transformations of electricity, and continue to develop an understanding of energy flows through systems. They link their experiences of electric circuits as a system at one scale to generation of electricity from a variety of sources at another scale and begin to see links between these systems.'

2.4.3 General capabilities

The general capabilities are a key dimension of the Australian Curriculum. They encompass knowledge, skills, behaviours and dispositions that, together with curriculum content in each learning area and the cross-curriculum priorities, can assist students to live and work successfully in the twenty-first century.

The Australian Curriculum includes seven General Capabilities which are addressed through the content of the learning areas, including Science. The aspects of the general capabilities identified as being most relevant and appropriate to the large-scale assessment of Science, and hence will be reflected in the NAP-SL assessment, include:

Literacy

In the NAP-SL assessment, aspects of the literacy capability will be found within the reading comprehension demands of both the stimuli and the questions and in the requirements of students to compose responses to questions.

Explicit definitions of terms are not part of the assessment, yet vocabulary is important in science communication. It is generally expected that the selection of science terms to be included in the assessment items will be guided by the range of terms that appear in the Australian Curriculum: Science. There is also an expectation that if the intent of an item is to assess understanding of a science idea, then science terminology should not impede a student’s ability to respond to the item.

While literacy plays an important role in science learning and assessment, it is important that the difficulty of items does not derive primarily from the amount and the complexity of the stimulus material and instructions. The NAP-SL stimuli and items should be written to a level appropriate for the students assessed and the literacy demand of items should be monitored by expert review.

Numeracy

Many elements of numeracy are evident in the Australian Curriculum strand *Science Inquiry Skills*. These include practical measurement and the collection, representation, analysis and interpretation of data from investigations. In the NAP-SL assessment, students will be expected to show dispositions and capacities to use appropriate mathematical knowledge and skills as outlined in the Australian Curriculum: Science[9]. In particular, students will be required to use a number of general numeracy skills, including arithmetic skills; measurement skills; data representation and data analysis skills, such as identifying patterns and relationships from numerical data and graphs; statistical analysis of data, including issues related to accuracy and validity; and the use of tools, such as modelling and simulations, to predict values and provide evidence in support of hypotheses.

However, items that assess predominantly numeracy skills should be avoided. If numeracy is required for understanding and responding to an item, it should not

---

[9] http://www.acara.edu.au/verve/_resources/Science_-_GC_learning_area.pdf

introduce construct-irrelevant variance to the assessment. The numeracy demand of items should be monitored by expert review to ensure that it is at an appropriate level for the assessment.

Information and Communication Technology (ICT)

Aspects of the ICT capability arise from the online delivery of the NAP-SL assessment, when students use the online system to undertake specific tasks. Students will be expected to use their ICT capability to access information, and collect, record, analyse and represent data. They will also be expected to use digital aids such as animations, simulations and other digital models to explore phenomena and test predictions.

Ethical understanding

Aspects of ethical understanding arise in the context of planning investigations and considering solutions to social and personal issues. Students will be expected to consider the implications of their investigations on the environment and living organisms as well as more broadly the ethical guidelines that apply in their investigations and those of others. They will also be expected to take into account ethical considerations when asked to make decisions about social or environmental issues, in addition to using science information and skills.

Intercultural understanding

Aspects of intercultural understanding arise in the context of considering the development and nature of science. Students will be expected to recognise that diverse cultural groups and perspectives have been contributing to the development of science knowledge and applications, consider how science benefits from participation and collaboration from a diversity of cultures, and recognise that increasingly scientists work in culturally diverse teams.

Personal and social capability

Aspects of the personal and social capability arise when students demonstrate abilities to question, solve problems and communicate their findings, and use scientific knowledge to inform personal and community decisions. Students will be expected to use their scientific knowledge to propose solutions to issues that impact their lives (such as health, nutrition and environmental change), and consider the application of science to meet personal and social needs.

Critical and creative thinking

Aspects of the critical and creative thinking capability arise from important cognitive skills inherent in scientific inquiry and broader scientific ways of thinking. The elements of critical and creating thinking capability from the Australian Curriculum have guided the development of the cognitive dimension of the NAP-SL – the thinking skills and intellectual processes to be engaged by the students as they respond to the assessment tasks. Chapter 3 in this framework provides additional information about how the elements of critical and creative thinking have been reflected in the Cognitive Dimension.

2.5 Assessment item contexts

The NAP-SL assesses student understanding and abilities in specific contexts. Students are expected to demonstrate their knowledge and abilities in a variety of contexts, including school science, personal, national, global, contemporary and historical contexts. Different sub-domains may lend themselves better to some contexts than to others. For example, the history of science provides good contexts for assessing the nature of science, while contemporary issues such as health and disease, natural resources, and environmental quality and hazards[10] provide good contexts for assessing the use and influence of science. Formal specifications are not set for context allocation.

Although assessment items may be embedded in both familiar and less familiar contexts, contexts should not introduce construct-irrelevant variance in the assessment, and they should be age-appropriate and sensitive to cultural differences.

[10] See OECD (2016). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*. Paris: OECD Publishing.

CHAPTER THREE: COGNITIVE DIMENSION

3.1 Cognitive dimension and cognitive processes

An important new feature of the NAP-SL 2018 assessment framework is the explicit definition of a cognitive dimension within the assessment of science literacy and across all three content domains[11]. The purpose of this chapter is to stipulate nationally consistent definitions and an explanation of the cognitive dimension of the science literacy measurement construct assessed in the NAP Science Literacy assessments. The definitions and taxonomy used to develop the cognitive domain are based on a well-known body of empirical evidence and, equally important, are consistent with those used in international science assessments including PISA and TIMSS. The cognitive dimension seeks to make explicit the thinking skills and intellectual processes that will be engaged by the students to respond to the assessment tasks.

The cognitive dimension in this framework is guided by the ways that science knowledge, science inquiry skills, and knowledge about science can be used by students, and the cognitive complexities that are inherent in these uses. It draws on a number of frameworks that define cognitive demand (including the revised Bloom Taxonomy[12]) as well as on the cognitive processes that underpin critical and creative thinking as defined by the General Capability in the Australian Curriculum[13]. It is adapted here to link more explicitly to both conceptual understandings and abilities, applying one dimension to all three content domains of the NAP-SL (*Science Understanding, Science Inquiry Skills* and *Science as a Human Endeavour*). This is consistent with other international frameworks, which incorporate aspects of science inquiry skills into a single cognitive demand rating scale (see, for example, Webb & Wixson, 2002[14]; Mullis & Martin, 2013[15]; OECD, 2016[16]). It is based on the assumption that there is a content in the procedural side of science (the 'knowing how'), which can be described and assessed[17].

It is recommended that the cognitive dimension is used by item writers along with the content descriptions and achievement standards as a guide to designing items, rather than merely as a framework to classify items.

---

[11] This is consistent with many national and international frameworks, such as TIMSS, PISA and NAP-CC.

[12] Anderson, L. W. & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

[13] https://www.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/critical-and-creative-thinking/; http://v7-5.australiancurriculum.edu.au/GeneralCapabilities/Pdf/Critical-and-creative-thinking

[14] Webb, N. & Wixson, K.K (2002). *Depth-of-Knowledge Levels for Four Content Areas*. Wisconsin Center for Education Research.

[15] Mullis, I.V.S. & Martin, M.O. (Eds.). (2013). *TIMSS 2015 frameworks: Science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

[16] OECD (2016). *PISA 2015 Assessment and Analytical Framework*: *Science, Reading, Mathematic and Financial Literacy*. Paris: OECD Publishing.

[17] Gott, R. & Duggan, S. (2003). *Understanding and using scientific evidence*. London: Sage Publications.

In this framework, the cognitive dimension — the cognitive processes that underpin what students are required to do in a task — comprises three areas:

- Knowing and using procedures
- Reasoning, analysing and evaluating
- Synthesising and creating

Knowing and using procedures requires knowledge of facts and definitions as well as the ability to illustrate concepts and generalisations with examples, and to relate science concepts to phenomena and observations. It also includes knowing simple procedures as well as the ability to perform simple science processes or procedures (e.g. make simple observations, measure, use a scale or units).

In the context of *Science Inquiry Skills* specifically, this area encompasses the knowledge and use of (practical) skills and procedures. These include the use of measuring instruments and the mechanical aspects of constructing tables and graphs, which are necessary but not in themselves sufficient to carry out most aspects of inquiry.

Reasoning, analysing and evaluating requires students to engage in applying knowledge, skills, processes, equipment and methods (e.g. classify, organise data, collect, compare and display data, use a diagram to illustrate a relationship, give a simple explanation) in contexts likely to be familiar or straightforward. It also requires students to analyse information and evaluate evidence and arguments with respect to quality and sufficiency of data.

In the context of *Science Inquiry Skills* specifically, this area also encompasses the application of procedural understanding related to inquiry processes (e.g. when students need to make decisions about what and when to measure, how many times and over what time period).

Synthesising and creating requires students to consider a number of different factors or concepts, put elements together (e.g. concepts, evidence, procedures, skills) into a coherent whole or compile elements in new ways or into something new and different. Tasks in *Synthesising and creating* are generally more open-ended or unstructured and can involve more than one approach or strategy; they require considerable cognitive effort because there is not likely to be a well-rehearsed method or pathway to approaching the task.

In the context of *Science Inquiry Skills* specifically, this area encompasses creating and using models, planning and designing scientific investigations, and carrying out full-blown extended investigations to solve a problem (from specifying a problem to designing and conducting the investigation, to analysing and evaluating the data [critically interpreting the data and methods of data collection], and forming conclusions drawing on concepts and evidence).

Tables 3.1—3.3 further specify and define the cognitive processes within the three cognitive areas. They focus on cognitive processes that can be used to characterise cognitive tasks that students may perform in science assessments (or in real life) when they use science knowledge and skills. Relevant elements of the

Australian Curriculum general capability *Critical and Creative Thinking* that are assessable in the context of a large-scale assessment have been integrated into these tables.

The examples that are used to illustrate the cognitive processes in Tables 3.1—3.3 are not exhaustive but should be sufficiently comprehensive to clarify the different demands of each area and guide assessment design.

Table 3.1 Knowing and using procedures

Knowing and using procedures	
Recognise	Make or identify accurate statements about science phenomena, concepts, relationships, procedures or statements about the scientific endeavour; recognise an instance of a concept/entity/ generalisation (e.g. producers or decomposers in a food web)
Define	Identify statements that define particular concepts and content[18]
Describe	Make straightforward observations of features/objects; identify and extract information from simple data sources or diagrams; describe factual information, processes and relationships about science or the scientific endeavour
Illustrate with examples	Identify or provide examples that support/clarify statements about the scientific endeavour or statements about particular science concepts, relationships and theories
Relate	Relate science concepts to phenomena and observations
Use tools and procedures	Demonstrate skills in the use of science equipment, tools, measurement devices/scales, mechanical aspects of constructing and reading graphs and tables

Table 3.2 Reasoning, analysing and evaluating

Reasoning, analysing and evaluating	
Compare/ Contrast/Classify	Identify similarities and differences between objects, processes or ideas; organise and process information; classify objects or processes based on characteristics/properties
Represent	Make representations (e.g. diagrams) to describe and illustrate aspects of concepts, structures, relationships, processes; make or use representations to communicate or find solutions to problems
Collect, analyse and interpret data	Make decisions about variables to be investigated and controlled, measurements to conduct; represent data in tables and graphs using appropriate labelling and scales; identify and summarise patterns in the data; interpolate/extrapolate from data
Make inferences	Make inferences from data, information given and/or own knowledge; give reasons/evidence to support an inference
Predict/Explain	Make predictions based on evidence and concepts; give reasons to support predictions; construct and defend explanations based on evidence and/or concepts; transfer knowledge into new

[18] This refers to identifying statements that correctly define a concept, not coming up with a complete and correct definition.

	contexts by making predictions and constructing explanations in new situations
Analyse information, evidence and arguments	Pose questions to probe assumptions, to identify gaps in information, evidence or arguments, or to investigate ideas or issues; prioritise information/evidence required to draw a conclusion or to make a decision; identify facts, observations or data that can be used as evidence to support an explanation, conclusion or argument; identify whether there is sufficient evidence to justify a claim, explanation or conclusion; identify evidence needed to decide among competing claims, explanations or solutions; integrate new information or evidence into ideas
Evaluate information, evidence, procedures and arguments	Evaluate information and evidence according to criteria such as relevance, bias, validity and/or reliability; evaluate claims, conclusions and arguments with respect to quality of evidence and reasoning supporting them; recognise flaws (e.g. gaps) in reasoning; consider and evaluate alternative explanations, processes and solutions; evaluate steps of investigations

Table 3.3 Synthesising and creating

Synthesising and creating	
Generate hypotheses	Generate hypotheses based on background knowledge, preliminary observations, and logic; generate and test alternative hypotheses; identify and justify the thinking processes behind such hypotheses
Construct arguments and draw conclusions	Construct sound and valid arguments supported by evidence and logical reasoning; draw conclusions that address questions/hypotheses and are supported with evidence; draw or support conclusions using evidence and scientific understanding; adapt conclusions as new evidence becomes available; draw general conclusions that go beyond the experimental or given conditions
Create and use models	Create models to explain a phenomenon or make a prediction (using imagery or analogies, as relevant); adapt models as new evidence becomes available; use computer simulations to test models under different conditions
Plan and design investigations	Plan and design whole investigations appropriate for answering scientific questions or solving problems
Make connections	Make connections between different concepts and areas of science; make decisions considering both scientific and social factors and trade-offs; synthesise complex information to inform a course of action
Solve problems	Seek and provide solutions to problems that require consideration of different factors and/or concepts; identify, assess and test options, implications and consequences when seeking solutions; propose alternative solutions to problems; justify the reasons behind choosing particular options/solutions; design solutions to problems of social significance, using science knowledge, considering a range of perspectives and trade-offs, and assessing risks

3.2 Cognitive processes in a balanced assessment

NAP-SL items in all three strands will be classified according to the cognitive areas and cognitive processes they seek to engage in students. This will ensure that the NAP-SL assessment will include items that cover a range of complexity and that students will be asked to use knowledge and skills in a variety of ways, some associated with higher and some with lower demand. This means that the assessment will be able to provide information on how students across the ability range can deal with tasks at different levels of demand.

It should be noted that the difficulty of an item should be distinguished from its complexity or cognitive demand. Item difficulty relates to the proportion of students answering a given item correctly while item demand relates to the cognitive processes necessary for successful completion of a task[19]. An item may have low demand because it only involves simple recognition of information but be difficult because it is assessing an unfamiliar content area. An assessment may be well targeted, e.g. include a sufficient number of difficult items for students of higher ability, yet not include items that will engage students with the higher level cognitive processes. It may also include items that assess a narrow spectrum of knowledge use (e.g. recognise, make inferences, draw conclusions) rather than include important uses of knowledge such as relating science concepts to phenomena or evaluating alternative explanations, processes and solutions.

3.3 Developing performance expectations

Science content statements and the processes listed in Tables 3.1–3.3 can serve as the basis of the development of 'performance expectations' which articulate the types of tasks that provide evidence of student understanding and proficiency with the constructs assessed.

Statements in content sequences specify exactly what knowledge or skills we want students to have, while the processes in Tables 3.1–3.3 specify how we may want them to engage with the knowledge and skills. By crossing a content statement with a cognitive process, one can develop performance expectations for each content statement. The pool of performance expectations for content statements will further articulate the assessment domain, reflect the expectations stated in the Australian Curriculum: Science achievement standards, and support more 'principled' assessment design, and, as a result, higher assessment construct validity.

For example, consider the specific content expectation 'Whenever an object A exerts a force on another object B, object B exerts a force equal in magnitude and opposite in direction back on object A' from the content progression *Forces and motion*. Using Tables 3.1–3.3, one can consider how this concept can be

---

[19] Item difficulty may depend on a variety of factors, such as reading load, linguistic demands, content complexity, and item response type. Item response type is sometimes used along with cognitive complexity of the task to determine the cognitive demand of an item. Empirical studies that have attempted to predict item difficulty based on existing models have had limited success (El Masri et al., 2017).

embedded in tasks of different demands that demonstrate different uses of knowledge (e.g. identifying instances of this concept in real life, giving an example of a design that takes this concept into account, analysing a physical situation in terms of the action/reaction forces acting on the bodies, etc.). This is a sample of performance expectations that result from this process and can guide the development of specific tasks that assess this concept:

- Identify action/reaction pairs in sports photographs.
- Give an example of how engineers take advantage of reaction forces in their designs.
- Analyse the relative magnitudes of forces in interactions with different effects on the interacting objects (e.g. different damage, different movement).
- Predict and explain the relative speeds of two objects following an interaction when the objects are dissimilar in some property (e.g. mass, size).
- Propose a valid investigation for testing whether objects of different mass dent by different amounts during a collision.

Sample performance expectations have been drafted[20] for content sequences about *Science as a Human Endeavour, Forces and Motion* and *Cellular Systems (Basic Functions)*. They demonstrate how different expectations with a range of cognitive demands can be generated from the same content description.

[20] Sample performance expectations are not included in this document. Empirical evidence from the 2018 tests will be used to construct the Science Literacy learning progressions that will describe the development of key aspects of curriculum content and cognitive domains outlined in the proposed assessment framework. This work will be completed in conjunction with the extension of the current NAP Science Literacy scale and proficiency levels and standards to Year 10 in the bridging year (2018), and will then be further refined in 2021.

CHAPTER FOUR: ITEM TYPES

4.1 Response formats and item types

The cognitive dimension and ensuing performance expectations discussed in Chapter 3 suggest that to capture the variability and different levels of complexity of performance, different types of assessment items and response formats would be required. To take an extreme case, an assessment that only consists of multiple-choice questions would not be representative of the construct(s) nor capture the range of cognitive demands as defined in Chapters 2 and 3 of this framework.

Three types of response formats will be used in the NAP-SL to assess the understandings and abilities identified in the framework: selected response, constructed response, and technology-enhanced intermediate constrained items.

In **selected-response** formats, students respond to a question by selecting the answer(s) they believe is/are more justifiable from a given set of alternatives. With computer-based testing, there is a great variety of selected-response formats to use (see below)[21]. However, a greater variety of formats in the assessment does not necessarily make a better test. Items that use ‘drag and drop’ utility can often be completed more efficiently using a multiple-choice format. The uses of ‘drag and drop’ listed below do assess categorising, ranking and sequencing more efficiently than multiple-choice or their paper-and-pencil equivalents. In general, the type of performance expectation(s) identified for development should guide the response format(s) used, not the other way around.

- **Conventional multiple choice**: Options may be words, graphical, pictorial and may incorporate new media. In the NAP-SL assessment, whenever possible, and especially when assessing use of knowledge to predict and explain phenomena or understanding of the nature of evidence, students’ misconceptions, mental models and alternate ways of thinking about the natural world should guide the development of distractors.
- **Selection/Multiple answer**: Select one or more options (including ‘all that apply’).
- **Two-tier multiple choice**: Select an option for a prediction, explanation etc. and then select from a different set of options to justify reasoning. This format appears to offer an efficient way of assessing higher cognitive demand items related to making and justifying hypotheses, predictions, explanations and arguments. The sequential responses will need to be integrated in a way that avoids interdependence of items. This is a new format for NAP-SL that is recommended to be included in the NAP-SL 2018 assessment.

---

[21] Scalise (2009) has developed a comprehensive taxonomy of computer-based items (http://pages.uoregon.edu/kscalise/taxonomy/taxonomy.html).

- **Categorising**: Select, drag and drop words, graphical or pictorial elements for classification purposes.
- **Ranking and sequencing**: Select, drag and drop items to place them in order.
- **Interlinear**: Select from multiple words to insert at various points in a sentence or passage.
- **Hotspot**: Select one or more predefined areas on an image.
- **Match**: Select which source objects match which destination categories by clicking a grid of radio buttons or checking checkboxes.

In **constructed-response** formats, students respond to a question by generating a response (rather than selecting a response from a given set of alternatives). Constructed-response items include short-constructed response and extended-constructed response items.

- **Short-constructed**: One or two words, a phrase or numerical response is required as a response to an item. Short-constructed response that can be completed with multiple-choice format should be avoided. The short-constructed format might be more appropriate when recall rather than recognition of information is important or greater depth of understanding is required than what can be probed with a multiple-choice question. Supplying titles for tables and graphs, graph labels and table headings would also be classified as short-response items.
- **Single numerical:** Enter a single numerical answer in a text box, including setting values for input variables in simulations.
- **Extended-constructed** (**extended text**): One sentence up to a couple of paragraphs are required as a response to an item. This format would be utilised to respond to a question that requests students to apply or integrate concepts, probe students' deeper understanding, and/or probe students' ability to communicate. It is particularly useful for tasks targeting the Synthesising and creating cognitive dimension.
- **Extended-constructed** items in the past NAP-SL assessments had two or three scoring categories (0, 1 or 0, 1, 2). It is recommended that for NAP-SL 2018, a small number of items with four scoring categories (0, 1, 2, 3) are trialled and possibly included in the final assessment. These items would tap into the more multifaceted content descriptions and advanced cognitive dimensions (in particular those that require integration/synthesis of concepts or ideas/evidence from different sources)[22].

---

[22] For example, students might be asked to integrate ideas about the role that the biosphere, lithosphere and atmosphere play in the carbon cycle; use a diagram of rock layers in an area to describe a sequence of events in the area's history that can be determined from the rock layers and use evidence to support their description.

In **technology-enhanced intermediate constrained** formats, students respond to questions while having a larger portion of the outcome space available to them. This means that these formats are less constrained than multiple-choice items (which are fully constrained) but more constrained than the open-ended constructed-response formats. These formats resolve two main concerns with multiple-choice items: when a limited set of choices is provided, students can back-solve (rather than directly solve a problem/answer a question, by testing each of the provided options) or student thinking may be prompted by the option (students 'recognise' the answer). An additional advantage of these formats is that they are scorable by computer. NAP-SL assessments items should take advantage of this type of response format for appropriate targeted content and cognitive demand.

- 'Sore finger': Identify an item that is incorrect/out of place with the rest (e.g. identify part of a model that is incorrect).
- Limited figural drawing: Construct limited drawings (e.g. use dial controls to show shadow's length and orientation, move shadows to correct location, draw force arrows or arrows in food webs). Advantage over multiple choice is that there is no limited set of options and advantage over equivalent paper and pencil is that students do not need to draw from scratch.
- Assembling proof/explanation: Identify and (if required) construct a proof/explanation or piece together evidence by highlighting or dragging and dropping parts from a text.
- Concept map: Create a concept map by arranging words and placing arrows.
- Open-ended graphic: Drag points on a graphic to create a graph.

4.2 Additional technological enhancements

Beyond enlarging the number of item response formats available, it is recommended that the NAP-SL 2018 assessments take advantage of additional technology-based enhancements to items, compared with previous NAP-SL administrations.

Computer-based enhancements can broaden the stimulus material that can be presented to students, and as a result the content that can be readily assessed as well as the cognitive complexity of the required responses:

- Students may observe a video or animation describing a phenomenon, experiment or investigation (instead of reading a stimulus text). Several phenomena, processes, experiments, etc. have been excluded from previous assessments as stimulus material because they are difficult to describe or make accessible to students, and/or their description results in high reading load. This includes phenomena and processes that happen over time, too quickly, too slowly; are on too small or too large a scale to observe directly[23].

[23] This also includes simulations of such phenomena. For example, students may be presented with a simulation of a physical, chemical or biological system or process, which they can explore by changing input parameters and observe how the system reacts (visually

- Students may view data from various external sources, multiple sources of information or media presentations as stimulus material for assessing interactions between science and society.
- Students may be asked to respond to a Predict-Observe-Explain situation, in which they make a prediction about an event, observe video or animation that is likely to surprise them, and ask them to add to or change their ideas about what happened.

Other enhancements, such as adaptivity and capability to collect process data, will be presented in the next section on inquiry tasks[24].

4.3 Science inquiry tasks

4.3.1 Inquiry tasks in previous NAP-SL cycles

In addition to an objective test, the first three cycles of the NAP-SL (2006, 2009 and 2012) included a 45-minute practical component. Its purpose was to provide students with an opportunity to experience practical aspects of science within a formal assessment and test the conventions of science literacy in more depth than was possible in the objective test.[25] The practical component comprised a two-stage structure. In the first stage, students carried out a science practical group work task. Students then answered individually a range of questions related to data representation, drawing conclusions and evaluating their work. Students' group work was carefully structured to enable all students to reach a similar end point. Two practical tasks were administered in each cycle; each student participated in one task. The items in the practical component were part of the same scale as the items in the objective test.

In NAP-SL 2015, a 45-minute online inquiry task was introduced that targeted similar content as the previous NAP-SL practicals. The approach was developed within the capabilities of the 2015 NAP-SL item-authoring and test-delivery systems and took into account the available technology in schools. The tasks presented were closed. Students were placed in an observer's role rather than being active participants in the tasks. Students were not directly engaged in a practical activity but were tested on a range of relevant science inquiry skills based on their observations from the video stimulus. The approach was enhanced with video stimuli.

---

or in the form of graphs of characteristic output parameters). Examples include the movement of gas particles in a closed container as a function of pressure and temperature and populations of predator and prey in dependence of available resources.

[24] There are obviously additional non-subject matter-specific enhancements, such as new strategies for assessing students who are developing their language skills or who have other special needs. Strategies include the provision of glossaries, audio reading of passages, varied size of text and volume, etc.

[25] This section, *Inquiry tasks in previous NAP-SL cycles,* is based on ACARA (2015). *National Assessment Program - Science Literacy. Assessment framework 2015*. Sydney: ACARA.

A key limitation of previous NAP-SL inquiry tasks is their limited degree of openness[26]. Students followed step-by-step instructions to collect the data and were guided with structured questions through the steps of interpreting and evaluating the data. Assessment developers needed to consider the differences in students' science curricula, programs and experiences, particularly in the absence of a unifying science curriculum in previous NAP-SL cycles. In addition, there was concern that absence of structure might produce unanticipated responses (especially in hands-on tasks) or make the large-scale administration of the tasks unmanageable.

Future technology-enhanced inquiry tasks should be more open-ended. This includes giving students an opportunity to determine for themselves the procedures that will yield robust evidence that can be used for justifying their conclusions or solutions to a problem.

4.3.2 Inquiry tasks in NAP-SL 2018

It is recommended that two inquiry tasks be administered in the NAP-SL 2018 assessment — two 40-minute tasks in Year 6 and two 50-minute tasks in Year 10[27]. This is consistent with the number and duration of tasks in previous NAP-SL cycles. Other approaches, such as the administration of two shorter tasks, may also be explored. A key consideration is whether they lead to less efficient use of assessment time as students would need to familiarise themselves with two pieces of stimuli and sets of tools rather than one.

The inquiry tasks should primarily target abilities from the *Science Inquiry Skills* content sequences, with many of the items targeting the cognitive domain *Evaluating and synthesising*. The tasks should be related to science concepts within the content sub-domain *Science Understanding*; however, the inquiry skills rather than the concepts should be in the foreground of the assessment[28]. It is recommended that the tasks be in the middle of the content-lean, content-rich continuum[29] and that the concepts embedded in the tasks are well within grasp of most students. If substantive knowledge of science concepts is required to carry out the tasks, students who do not have this knowledge will not have the opportunity to demonstrate their inquiry skills[30].

---

[26] The degree of openness of a task relates to who defines the problem to be studied, who chooses the method and how many solutions are available. The first two considerations can be placed in a continuum from closely defined to not defined, while the last consideration ranges from one solution to many solutions.

[27] Four tasks would be developed for trial at each year level.

[28] It is recommended that a small number of items assess students' ability to use specific science concepts to design an investigation. In this case, the science content (rather than the inquiry skills) would be in the foreground and focus of the assessment; these items would be classified within the *Science Understanding* (rather than the *Inquiry Skills*) domain. (Students would not need to carry out these investigations, only design them.)

[29] Content-rich tasks require in-depth understanding of subject matter for task execution. Content-lean tasks are not dependent on prior subject-matter knowledge; performance only depends on information given in the assessment situation (Baxter & Glaser, 1998).

[30] An alternative would be to develop several tasks with different concepts embedded, but this is not an option within the time constraints in the NAP-SL.

The focus of the inquiry tasks should be on aspects of inquiry that cannot be effectively or efficiently assessed in shorter tasks/items. In addition to planning and carrying out investigations, this would include the notion of the overall evaluation of an inquiry in terms of the credibility of the evidence gathered and the solution produced[31]. This suggests that test developers should design upfront an overall strategy for the assessment of inquiry abilities in the NAP-SL — which abilities would be assessed through shorter stimuli and secondary data in the first part of the assessment, and which would be assessed through inquiry tasks. For example, to enable in-depth assessment of some aspects of the inquiry tasks, time-consuming aspects of data representation (e.g. graph drawing) may be assessed in item sets in the first part of the assessment.

It is recommended that the inquiry tasks are computer-based rather than administered as hands-on or hybrid tasks[32]. Uniform administration of hands-on tasks poses challenges, and the need for standardisation, the cost of the assessment kits, safety and class-management concerns pose significant constraints on what can be assessed in hands-on tasks in a large-scale assessment program.

As discussed in a previous section *(Additional technological enhancements)*, computer-based tasks significantly broaden the type of inquiry with which students can engage, and as a result the content that can be readily assessed and the cognitive complexity of the required responses. Students may:

- explore phenomena and processes that are too slow or too fast, or not visible to the naked eye
- explore phenomena or processes that would be considered hazardous (e.g. using hot materials) or messy (e.g. using water)
- develop, use and test representations to model the real world
- carry out repetitions/replications of experiments within short assessment times.

Technology gives developers the ability to manipulate the degree of openness of the task (see also adaptivity below) and capture process data. Computer interactive tasks are likely to increase student engagement with the assessment. The key disadvantage of computer-based tasks is authenticity — science is about the real world.

Computer-based inquiry tasks may be embedded in a virtual laboratory or use a simulation. Both approaches provide opportunities to assess a whole investigation, from understanding the problem, planning how to go about the investigation, implementing that plan, collecting the data, drawing conclusions from the data, and evaluating the whole investigation as one integrated process.

Not all inquiry tasks developed have to be experiments. Tasks should also assess other methods of scientific inquiry such as observation, classifying, pattern

---

[31] This is different from the ability to critique isolated claims or explanations.
[32] In hybrid tasks, students carry out a hands-on investigation but they record their answers on the computer.

seeking, and modelling, rather than only fair-testing. Technology opens the possibility for different types of inquiry that are, for example, too difficult to carry out in a practical assessment session or require looking at data over time. Such types of inquiry may be carried out by students at the planning and predicting level only; students may be presented with secondary data to complete the inquiry. A pilot should investigate a variety of approaches before deciding whether all inquiry tasks in the NAP-SL should follow the same structure. The inquiry task is not an end in itself; it is a means to obtain valid information about the level of student abilities related to important aspects of the content domains.

A key challenge in the design of inquiry tasks is the question of how to provide open-ended environments to tap into difficult-to-assess constructs while giving all students the opportunity to demonstrate what they are able to do and, at the same time, preserving the independence of the items. Some form of adaptivity or branching may be required to address this:

- After students complete the first part of the assessment, they may be asked to complete inquiry tasks at different degrees of openness and difficulty, depending on their performance in the first part of the assessment.
- Stages may be designed within tasks in a way that students can proceed with the next stage even if they have not completed the previous one. For example, a branch may be developed that gives students step-by-step instructions to carry out a task if they have not completed the planning phase of an investigation.
- Instead of full branches, prompts may be given to students who may get stuck in one stage to enable them to proceed. Scoring guides may be developed that allocate marks depending on whether students responded to the question with or without prompting and their scores adjusted accordingly.

Drawing on the previous analysis, specific recommendations for development of NAP-SL inquiry tasks include:

- Advantages and disadvantages of including two smaller inquiry tasks vs. one longer task in the NAP-SL should be explored.
- Tasks should be developed that relate to or connect two or more concepts within different *Science Understanding* sub-domains. The concepts selected should be ones that are typically understood reasonably well by students.
- Regardless of the type (investigation, experiment, etc.), each task should be guided by a question and should be contextualised as authentically as possible; the question should engage students in solving a problem rather than requiring them to carry out procedures for no other reason.
- The question/problem that guides each task should be well-defined, but the task should also give students choice(s) of method and solution.
- Access to a range of appropriate resources/tools should be available so that students can select appropriate instruments and make appropriate measurements.

- Choice of response formats should be like those in the first part of the assessment (see section *Response formats and item types*). For tasks that rely on a virtual environment or simulation, students' interaction with the simulation will also produce action-responses to be captured and scored, e.g. what equipment was selected, which variables were selected and which values were used for those variables, whether multiple trials were run, etc.
- Independence of items must be maintained; that is, a correct or incorrect reply to one item should not lead to a correct or incorrect reply to another item.
- Branching options should be considered to give all students the opportunity to respond to the tasks to the best of their ability.

CHAPTER FIVE: ASSESSMENT STRUCTURE AND REPORTING

5.1 Assessment structure

The NAP-SL uses a cluster rotation design similar to that used in other sample-based international assessments. In the rotation design, assessment forms are assembled so that each form is linked through common clusters to other forms. In this way, a broader range of assessment items can be completed and linked to other items.

To achieve the rotation design for the NAP-SL, the items are written in contextual units. Each unit contains between one and five items that are developed around a single theme or stimulus. Clusters are then constructed by grouping units together. Clusters are grouped together into assessment forms.

Clusters that are intended to contain vertical link items, as addressed in section 2.3, should provide a good sampling of the content and cognitive domains of the assessment framework across both year levels. The inquiry tasks will not be used as vertical links owing to the lack of a framework for the construction of such vertical link tasks. This decision will be revisited once the vertical scale is established and more empirical data on the interaction of content and cognitive aspects of science literacy across the two year levels are collected.

Assuming that the NAP-SL 2018 will follow the same assessment design method, it is recommended that:

- each student is assigned an assessment form of one hour and an inquiry task/inquiry tasks that does/do not exceed one hour
- item sets include at least two to three items to reduce reading load
- assessment forms do not include more than 40 marks (typically, assessment clusters include 13–14 items and assessment forms 39–42 items). This is particularly important if items with four score points are introduced (see Chapter 4).

The NAP-SL student survey that was first introduced in 2009 has been improved and enhanced to provide information that is better aligned with the Australian Curriculum: Science and with the revised definition of science literacy (section 1.2), particularly pertaining to the *Science as a Human Endeavour* strand. The survey covers three broad areas:

- science as a human endeavour,
- student engagement with science, and
- teaching and learning science.

Each of these areas can be stratified further to form a total of 12 survey item clusters. Further details are provided in the 2015 NAP Science Literacy technical report (ACARA 2017b, page 108). The survey outcomes will be reported across these 12 item clusters including their correlation with students' overall achievement in science literacy at the national as well as state and territory levels. The survey responses will be scaled to provide construct indicators of students' perception and engagement with science using the same methodology as outlined in the 2015 NAP Science Literacy reports (ACARA 2017a, 2017b).

5.2 Reporting proficiency in science literacy

One of the main objectives of the NAP-SL is to monitor trends in science literacy performance over time. To enable effective historical comparison, it is important that the underlying construct of the NAP-SL assessment is maintained. ACARA (2015)[33] includes comparisons between the previous NAP-SL content domains and the Australian Curriculum: Science, which show that there is a high degree of alignment between them. This means there is also a high degree of alignment between the previous NAP-SL domains and the Australian Curriculum-based content domains described in this framework. In addition, the specifications for distributions of items across content domains in this framework reflect the item distributions from previous NAP-SL cycles (see Chapter 2, p. 8), further substantiating that the NAP-SL construct is maintained in this framework.

The approach to reporting used by the NAP-SL has been developed in previous assessment cycles and is based on the definition and description of a number of levels of proficiency in science literacy. In previous cycles, descriptions were developed to characterise typical student performance at each proficiency level. The proficiency levels were used to summarise and report on the performance of Year 6 students (across Australia as well as in individual states and territories), to compare performance across subgroups of students and to report on the performance of students over time.

A similar approach will be used to report NAP-SL results for the 2018 cycle. However, in the NAP-SL 2018, assessments will be administered to samples of Year 6 and Year 10 Australian students and the current science literacy scale will be extended to include Year 10 outcomes. The new assessment framework (including the content and cognitive dimensions) will support this further development of the scale.

More specifically, the introduction of the Year 10 assessment and the new assessment framework will support the following advances:

- The science literacy scale will be extended by the definition and description of new proficiency levels to specifically address and provide a description of performance for students at higher levels of ability who demonstrate science literacy commensurate with high school expectations.
- Proficiency levels and descriptions will be revised and enriched using information from the new items and augmented item classifications that incorporate the new framework dimensions developed for this cycle. For example, the assessment will be able to provide information on whether students across the ability range can

---

[33] ACARA (2015). *National Assessment Program - Science Literacy. Assessment framework 2015*. Sydney: ACARA.

use scientific knowledge in a variety of ways and how they deal with tasks at different levels of demand.

- Proficiency level descriptions will be aligned more closely with the expectations of the Australian Curriculum, as the assessment items and their descriptions will be guided by content and cognitive framework dimensions that reflect the knowledge and capabilities articulated in the Australian Curriculum.
- More broadly, stronger alignment between the NAP-SL assessment and the Australian Curriculum and use of technology-enhanced response formats in the assessment means that the results for the 2018 cycle will provide more useful data about Australian students' performance related to the specific knowledge, skills and capabilities included in the Australian Curriculum (including those that are harder to assess with traditional response formats) and will support more in-depth feedback on planning and strategies for future science programs. This includes the identification of opportunities and gaps in how students approach and respond to critical thinking tasks and how they engage with open-ended scenarios that require a deeper level of planning, analysis and synthesis.

GLOSSARY

Term	Definition
Achievement standards (AC[34])	The learning expected of students at each year level or band of years in each learning area of the Australian Curriculum
Cognitive demand	The kind and level of thinking required of students in order to successfully complete a task
Cognitive dimension	One of the two dimensions of the NAP-SL assessment framework, related to the thinking skills and intellectual processes targeted as students respond to assessment tasks
Construct	The underlying cognitive abilities measured by an assessment
Construct validity	The extent to which an assessment adequately measures the construct it purports to measure
Content descriptions (AC)	Statements in each learning area of the Australian Curriculum that outline what is to be taught to students from Foundation to Year 10
Content dimension	One of the two dimensions of the NAP-SL assessment framework, related to the specific subject matter covered in the assessment
Content domains	The knowledge or ability domains that will be assessed
Content sequences	Sequences of ideas or skills that unpack the content descriptions of the Australian Curriculum: Science to guide item development. Content sequences identify the essential elements for each key concept/ability and articulate increasingly more sophisticated ways of thinking about the concept/ability. Content sequences are informed by the existing literature but are not empirically validated in the NAP Science Literacy context
Cross-curriculum priorities (AC)	A dimension of the Australian Curriculum that describes knowledge, understanding and skills relating to: *Aboriginal and Torres Islander histories and cultures; Asia and Australia's engagement with Asia; Sustainability*
General capabilities (AC)	A dimension of the Australian Curriculum that describes knowledge, skills, behaviours and dispositions relating to seven capabilities that are important for life and work in the 21st century
Inquiry task	One of two components in NAP-SL assessments; assesses inquiry skills in the context of a whole investigation
Investigation	A scientific process of answering a question, exploring an idea or solving a problem that requires activities such as planning a course of action, collecting data, interpreting data, reaching a conclusion and communicating these activities
Item	A question included in the assessment which is designed to assess the knowledge or abilities of students
Key ideas (AC: Science)	Key aspects of the scientific view of the world that bridge knowledge and understanding across the different science disciplines
Learning progressions	Empirically validated descriptions of successively more sophisticated ways of thinking about an important knowledge or ability domain that follow one another as students learn about a topic over a broad span of time

[34] Term from the Australian Curriculum (AC)

Objective test	One of two components in NAP-SL assessments; assesses student knowledge and abilities in the context of small item sets framed by a stimulus
Performance expectations	Descriptions of types of tasks that provide evidence of student understanding and proficiency with the construct(s) assessed. Performance expectations result from crossing content sequence statements (content dimension) with cognitive processes (cognitive dimension), and are reflective of the Achievement Standards in the Australian Curriculum: Science
Proficiency levels	Ranges of scores on an assessment scale accompanied by descriptions of performance and skills associated with each level
Response format	The mode in which students respond to an item. Common response formats include selection of the correct response among options and constructed response
Science literacy	The ability to use scientific knowledge, understanding, and inquiry skills to identify questions, acquire new knowledge, explain science phenomena, solve problems and draw evidence-based conclusions in making sense of the world, and to recognise how understandings of the nature, development, use and influence of science help us make responsible decisions and shape our interpretations of information
Science literacy scale	The proficiency scale for the NAP-SL enabling comparisons of results to be made across assessment cycles; describes the development of student proficiency in science literacy along multiple levels
Stimulus	Material used in assessments to provide context for assessing the knowledge and skills of students
Strand (AC)	The largest structural unit within a learning area (subject) in the Australian Curriculum. The learning area Science comprises three interrelated strands: *Science Understanding, Science as a Human Endeavour, Science Inquiry Skills*
Sub-strand (AC)	A subdivision of a strand in the Australian Curriculum
The Australian Curriculum: Science (AC)	Standards for what all young Australians should learn as they progress through schooling in the learning area of science

REFERENCES

AAAS (2001). *Atlas of Science Literacy*. Vol. 1. Washington DC: AAAS.

AAAS (2007). *Atlas of Science Literacy*. Vol. 2. Washington DC: AAAS.

ACARA (2015). *National Assessment Program - Science Literacy. Assessment framework 2015*. Sydney: ACARA.

ACARA (2017a). *NAP Sample Assessment - Science Literacy 2015. Public report*. Sydney: ACARA.

ACARA (2017b). *NAP Sample Assessment - Science Literacy 2015. Technical report*. Sydney: ACARA.

Anderson, L. W. & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Baxter, G. P. & Glaser, R. (1998). Investigating the cognitive complexity of science assessments. *Educational Measurement: Issues and Practice*, *17* (3), 37–45. doi:10.1111/j.1745-3992.1998.tb00627.x

Driver, R., Squires, A., Rushworth, P. & Wood-Robinson, V. (2015). *Making sense of secondary science: Research into children's ideas*. New York: Routledge.

El Masri, Y.H., Ferrara, S., Foltz, P.W. & Baird, J.A. (2017). Predicting item difficulty of science national curriculum tests: the case of key stage 2 assessments. *The Curriculum Journal, 28* (1). doi: 10.1080/09585176.2016.1232201

Gott, R. & Duggan, S. (2003). *Understanding and using scientific evidence*. London: Sage Publications.

Herrmann-Abell, C. F., & DeBoer, G. E. (2008, April). *An analysis of field test results for assessment items aligned to the middle school topic of atoms, molecules, and states of matter*. Paper presented at the National Association for Research in Science Teaching (NARST) Annual Conference, Baltimore, MD.

Johnson, P. (1998). Progression in children's understanding of a "basic" particle theory: A longitudinal study. *International Journal of Science Education*, *20*(4), 393–412. doi: 10.1080/0950069980200402

Johnson, P. (2013). How students' understanding of particle theory develops: A learning progression. In G. Tsaparlis & H. Sevian (Eds.). *Concepts of matter in science education.* (pp. 47–67). Dordrecht: Springer.

Lee, O., Eichinger, D. C., Anderson, C. W., Berkheimer, G. D. & Blakeslee, T. D. (1993). Changing middle school students' conceptions of matter and molecules. *Journal of Research in Science Teaching*, *30* (3), 249–270. doi:10.1002/tea.3660300304

Mohan, L., Chen, J. & Anderson, C. W. (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching*, *46* (6), 675–698. doi: 10.1002/tea.20314

Mullis, I.V.S. & Martin, M.O. (Eds.). (2013). *TIMSS 2015 frameworks: Science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

National Academies of Sciences, Engineering, and Medicine. (2016). *Science Literacy: Concepts, Contexts, and Consequences.* Washington DC: The National Academies Press.

OECD (2016). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*. Paris: OECD Publishing.

Russell, T. & Watt, D. (1990). *Primary SPACE Project Report. Evaporation and condensation*. Liverpool: Liverpool University Press.

Scalise, K. (2009). Computer-based assessment: “Intermediate constraint” questions and tasks for technology platforms. Retrieved September 10, 2017, from http://pages.uoregon.edu/kscalise/taxonomy/taxonomy.html.

Webb, N. & Wixson, K.K. (2002). *Depth-of-knowledge levels for four content areas*. Madison: University of Wisconsin, Wisconsin Center for Education Research.

Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.

Wilson, M. (2012). Responding to a challenge that learning progressions pose to measurement practice: Hypothesized links between dimensions of the outcome progression. In A.C. Alonzo & A.W. Gotwals (Eds.). *Learning progressions in science.* (pp. 317–344). Rotterdam: Sense Publishers.