



# NAPLAN Online Research and Development

Report 1: Readability and Layout Study – Literature Review

Report 2: Readability and Layout Study – Field Trial

2015

# NAPLAN ONLINE RESEARCH AND DEVELOPMENT

## Readability and Layout Study

The readability and layout study investigated how students perceive and react to a range of different online item configurations and displays. The purpose of the study was to evaluate different item display and navigation options relative to onscreen layouts.

Layout presentations included different options for the location of images in relation to text on screen, navigation functionality within a single stimulus and across multiple stimuli, presentation of audio items for spelling, and presentation of item content (including horizontal or vertical response option layouts; font type, line spacing, and line length).

Students in focus groups were presented with an option on screens and were asked to comment on differences between sample displays, and suggest how they would improve upon the layouts. Students across year levels tended to be consistent in their preferences and recommendations for layout options, with the exception of navigating long reading passages.

The study was conducted in two phases: a literature review and a field trial.

Reports from two contractors have been provided in this document:

Associate Professor Esther Care, Dr Rebekah Luo, Nafisa Awwal, Yasotha V,  
*NAPLAN Online Readability and Layout Study: Literature Review.*  
Assessment Research Centre, University of Melbourne, August 2015.

Laurie Laughlin Davis, Ph.D., Emily Lai, Ph.D., Rusty Brandt, *NAPLAN Online Readability and Layout Study.* Pearson, Melbourne, February 2016.



Assessment

*Teaching*

*Measurement*

*Psychometrics*

## **NAPLAN Online Readability and Layout Study – Literature Review**

Authored for the Australian Curriculum, Assessment and  
Reporting Authority (ACARA)

### **Project Team**

Associate Professor Esther Care

Dr Rebekah Luo

Nafisa Awwal

Yasotha V

**August 2015**

**Assessment Research Centre**

Melbourne Graduate School of Education

The University of Melbourne

Victoria 3010 Australia

Telephone +61 3 8344 2899

Facsimile +61 3 8344 8739

[www.education.unimelb.edu.au/arc/](http://www.education.unimelb.edu.au/arc/)

© Assessment Research Centre, 2013

# Table of Contents

---

1	Introduction.....	6
1.1	Overview of the current literature review.....	7
2	Online Assessment.....	8
2.1	Progressive Achievement Tests (PAT) in Reading Fourth Edition, Australia.....	8
2.2	On Demand Testing, Australia.....	8
2.3	Online Literacy and Numeracy Assessment (OLNA), Western Australia.....	8
2.4	Programme for International Student Assessment (PISA).....	9
2.4.1	Electronic Reading Assessment (ERA) by PISA.....	11
2.5	Online Reading Literacy Assessment (ORLA), Taiwan.....	12
2.6	Partnership for Assessment of Readiness for College and Careers (PARCC), USA ...	13
2.7	Smarter Balanced, USA.....	16
2.8	Programme for the International Assessment of Adult Competencies (PIAAC) .....	18
2.9	Review of online assessment tools and programs.....	19
3	Comparison between Paper and Online Assessments.....	26
3.1	Paper-based Assessments.....	26
3.2	Online Assessments.....	26
3.2.1	Web page reading.....	28
3.3	Factors to Consider for Online Assessment.....	30
4	Critical Analysis of Factors Important for Online Assessments.....	31
4.1	Typeface Characteristics.....	31
4.1.1	Font Type.....	31
4.1.2	Font Size.....	32
4.2	Text Format.....	32
4.2.1	Interlinear Spacing.....	33

4.2.2	Line Lengths .....	34
4.2.3	Length of text or passages .....	35
4.2.4	Columns .....	36
4.2.5	Justification of text.....	37
4.2.6	Display Size.....	38
4.3	Presentation Layout .....	39
4.3.1	Item Placement.....	39
4.3.2	Navigation Format .....	42
4.3.3	Reading tools .....	44
4.4	Device Features .....	45
5	Recommendations for NAPLAN Online .....	49
5.1	Display Format .....	50
5.1.1	Category 1: Text only passages, some with add-on graphics .....	50
5.1.2	Category 2: Text with graphics that aids reading comprehension .....	51
5.1.3	Category 3: Compare and contrast texts and passages.....	56
5.1.4	Other media .....	63
5.2	Item Types .....	69
5.2.1	Drop-down List.....	70
5.2.2	Drag and Drop Items .....	70
5.2.3	Click and Select or Select and Highlight Answers .....	72
5.2.4	Hot Spot Items .....	73
5.3	Navigation options .....	76
5.3.1	Progress Bar .....	76
5.3.2	Scrolling.....	82
5.3.3	Pagination .....	86
5.3.4	Innovative options .....	89

5.4	Reading Tools .....	94
5.4.1	Highlighter.....	94
5.4.2	Line Reader .....	95
5.4.3	Magnifying Tools.....	96
5.4.4	Flag Tool .....	100
5.4.5	Review Tool.....	101
5.4.6	Answer Eliminator and Selection Indicator .....	102
5.4.7	Notepad .....	104
5.5	Summary of recommendations.....	105
6	Cognitive Interviews .....	109
6.1	Concept check .....	110
6.1.1	Task Concept Check Example.....	110
6.2	Cognitive Interviews.....	111
6.2.1	Cognitive Interviews with Students – Concurrent Report.....	112
6.2.2	Cognitive Interviews with Students – Retrospective Report.....	115
6.3	Data Analysis of the Cognitive Interviews.....	116
6.4	Logistics checks .....	117
6.5	Summary of the cognitive interview section .....	118
6.6	Recommendations for the use of cognitive interviews .....	119
7	Summary.....	121
8	References .....	122

# 1 Introduction

---

Over the last decade, the world of education has undergone major transformations due to advances in technology. Unlike before, children in most developed countries are learning with the aid of new technologies such as computers and tablets. In addition, educational assessments are slowly shifting from traditional paper-and-pen test to online- and computer-based assessments, and online assessments are increasingly being developed and used.

Consistent with this trend, some large scale international assessment programs such as the Programme for International Student Assessment (PISA) have introduced computer-based tests for participating students. Likewise, in Australia, several online assessment tools have been developed, such as the Progressive Achievement Tests in Reading Fourth Edition by the Australian Council for Education Research (ACER), and On Demand Testing by the Victorian Curriculum and Assessment Authority (VCAA). Currently, the Australian Government is working with states and territories and the Australian Curriculum, Assessment and Reporting Authority (ACARA) to move the National Assessment Program – Literacy and Numeracy (NAPLAN) tests online, with goals to implement NAPLAN online from 2017 so that students can complete the assessment using a computer or other electronic device such as a tablet (ACARA, 2014).

Introduced in 2008, NAPLAN is an annual assessment for students in Years 3, 5, 7 and 9. The main purpose of NAPLAN is to assess the skills of Australian students in literacy and numeracy against the national standards. The results of NAPLAN tests also provide information to students, parents, teachers and schools about student achievement and progress. It is held that moving NAPLAN into an electronic format will bring about many benefits, including the use of tailored test design and the opportunity to broaden the scope of an assessment (ACARA, 2014). Factors relating to readability and layout of online assessment must therefore be examined and considered. The format of NAPLAN online tests will be influenced by the current paper-based test format. The NAPLAN reading test is currently presented using a reading magazine containing passages, and a test booklet containing multiple choice questions (indicate answer by shading a bubble) or constructed response (write answer in a box or lines provided). Given the fact that these two

components need to be considered for transfer to an online environment, the online reading test could be presented across bilateral sections on a screen.

The transition from paper-and-pen test to computer or online assessment has raised a number of questions pertaining to the differences between paper and computer-based tests, and the effects of test presentation modes on student performance. Studies examining these research questions have reported conflicting findings; some studies report no statistically significant differences in student reading performance across the media (paper or online assessment) (Wang, Jiao, Young, Brooks, & Olson, 2008; Young, 2014) while others have found statistically significant differences in reading speed and reading comprehension when tests were presented in different modes (Dillon, 1992; Kim & Kim, 2013). Research to date has provided limited information on how to best transfer test content from print to digital display. Factors such as typeface readability and characteristics, item layout and presentation, and device type may affect students' performance and engagement on assessments. Further investigation into these factors is needed to improve the validity and reliability of online assessments.

## 1.1 Overview of the current literature review

The purpose of this report is to review current literature about readability and layout of onscreen assessments. This review will consist of five main sections.

First, we present and describe national and international online reading or literacy assessments tools used in large scale studies or settings. If an assessment program or tool covers several learning domains, only information related to reading or literacy assessment is presented and discussed in detail. Screen shots and examples of assessment items are included if they are available online or accessible by the general public. Second, we compare and discuss differences between paper- and online assessments, and identify factors to consider when developing an online assessment. Third, we evaluate and analyse factors that are important for the design and implementation of online assessments. Fourth, we provide recommendations and practical solutions for transitioning NAPLAN reading assessments online. Last, we provide guidelines for use of cognitive interview and guidelines for conducting associated processes such as concept check and logistics checks.

## 2 Online Assessment

---

### 2.1 Progressive Achievement Tests (PAT) in Reading Fourth Edition, Australia

The Progressive Achievement Tests in Reading Fourth Edition (PAT Reading Fourth Edition) is a “thoroughly researched and normed test for measuring and tracking student achievement in reading comprehension, word knowledge and spelling” (ACER, 2015). Developed by the Australian Council for Educational Research (ACER), PAT Reading contains multiple-choice questions in Reading Comprehension, Vocabulary and Spelling. Only the PAT Reading Comprehension and Vocabulary tests are currently available for online administration. Online reading comprehension tests are available for students in Foundation year to Year 10, while the Vocabulary tests are available for students in years 3 to 10.

### 2.2 On Demand Testing, Australia

On Demand Testing is an online tool developed by the Victorian Curriculum and Assessment Authority (VCAA) to assist teachers with assessment needs such as identifying a student’s strengths and weaknesses (VCAA, 2014). On Demand tests are designed to link to the Australian Curriculum and the Victorian Essential Learning Standards (AusVELS), and covers both general ability tests and topic-specific assessments (VCAA, 2014).

On Demand testing program provides a selection of assessments for English and Mathematics, and for Literacy-Reading, Language-Spelling, Number and Algebra, and Measurement & Geometry. Tests are available in two formats: computer adaptive test format or linear test format. Computer adaptive tests provide sets of questions to students according to student ability; linear tests contain a fixed set of questions which are presented to all students in the same order (VCAA, 2013).

### 2.3 Online Literacy and Numeracy Assessment (OLNA), Western Australia

The Online Literacy and Numeracy Assessment (OLNA) provides students with opportunities to demonstrate minimum standards of literacy and numeracy required for the Western

Australian Certificate of Education (WACE), a senior secondary certificate in Australia (School Curriculum and Standards Authority, 2015). From 2016, a minimum standard of literacy and numeracy must be demonstrated by students in order to be eligible for a WACE. To demonstrate this minimum standard, students can either complete the OLN or obtain prequalification through Year 9 NAPLAN. OLN is conducted annually in March and September, and students in Years 10, 11 and 12 are given up to six opportunities (two per year) to demonstrate the minimum standard.

There are three online assessment components – Reading, Writing and Numeracy. Students are allowed 60 minutes to complete each assessment. The reading component of OLN is comprised of 60 multiple-choice questions.

## 2.4 Programme for International Student Assessment (PISA)

The Programme for International Student Assessment (PISA) is a triennial international survey of 15-year old school students in Organisation for Economic Co-operation and Development (OECD) and non-OECD countries. The goal of PISA is to evaluate and measure how well students approaching the end of compulsory schooling are prepared for their future in real-world contexts. Specifically, knowledge and skills in reading literacy, mathematics and scientific literacy are assessed. From 2015, the primary mode of delivery for all PISA assessments including reading literacy will be computer-based, although paper-based assessment instruments will still be made available to participating countries (OECD, 2013).

Sample items from the computer-based reading assessment are shown in Figures 1 and 2. These items were originally designed for use in paper and pencil test, but have been adapted slightly for presentation in computer-based format.

Sample Questions : PISA Reading OECD

[Home >>](#)    [Instructions >>](#)    [Toggle Marking Guide >>](#)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

[<< Previous](#)      **Question 7: A Just Judge**      [Next >>](#)

**JUST JUDGE TEXT**

Refer to the story *A Just Judge* to answer the questions which follow it.

**A JUST JUDGE**

An Algerian king named Bauakas wanted to find out whether or not it was true, as he had been told, that in one of his cities lived a just judge who could instantly discern the truth, and from whom no rogue was ever able to conceal himself. Bauakas exchanged clothes with a merchant and went on horseback to the city where the judge lived.

At the entrance to the city a cripple approached the king and begged alms of him. Bauakas gave him money and was about to continue on his way, but the cripple clung to his clothing.

[Popup Stimulus >>](#)

Near the beginning of the story we are told that Bauakas exchanged clothes with a merchant.

Why didn't Bauakas want to be recognised?

- He wanted to see if he would still be obeyed when he was an "ordinary" person.
- He planned to appear in a case before the judge, disguised as a merchant.
- He enjoyed disguising himself so he could move about freely and play tricks on his subjects.
- He wanted to see the judge at work in his usual way, uninfluenced by the presence of the king.

Sample items Copyright © OECD unless otherwise noted.  
Item delivery Copyright © 1999-2007, The Australian Council for Educational Research

Figure 1. Sample question of the PISA reading assessment. The passage is presented on a scroll page above the question and multiple-choice response options. Screen shot retrieved from <http://pisa-sq.acer.edu.au/>

Sample Questions : PISA Reading OECD

[Home >>](#)    [Instructions >>](#)    [Toggle Marking Guide >>](#)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

[<< Previous](#)      **Question 13: Bullying**      [Next >>](#)

**BULLYING TEXT**

**PARENTS LACK AWARENESS OF BULLYING**

Only one in three parents polled is aware of bullying involving their children, according to an Education Ministry survey released on Wednesday.

The survey, conducted between December 1994 and January 1995, involved some 19,000 parents, teachers and children at primary, junior and senior high schools where bullying has occurred.

The survey, the first of its kind conducted by the Ministry, covered students from the fourth grade up. According to the survey, 22 per cent of the primary school children polled said they face bullying, compared with 13 per cent of junior high school children and 4 per cent of senior high school students.

[Popup Stimulus >>](#)

Why does the article mention the death of Kiyoteru Okouchi?

**Answer**

Sample items Copyright © OECD unless otherwise noted.  
Item delivery Copyright © 1999-2007, The Australian Council for Educational Research

Figure 2. Sample question of the PISA reading assessment. The passage is presented on a scroll page above the question and answer box. Screen shot retrieved from <http://pisa-sq.acer.edu.au/>

### 2.4.1 Electronic Reading Assessment (ERA) by PISA

The Electronic Reading Assessment (ERA) was an assessment of reading in a digital environment (e.g., reading of electronic texts). This assessment was included in the 2009 and 2012 Programme for International Student Assessment (PISA) (OECD, 2009). Students were given a 10-minute practice session and subsequently asked to complete a 40-minute computer-based assessment. Figures 3 and 4 are examples from the ERA.

The screenshot shows a web browser window with the address <http://www.maikasblog.com/index.html>. The page title is "Maika's Blog". The main content is a blog post titled "Life Begins at 16" dated "TUESDAY, JANUARY 1". The post text reads: "Happy New Year! Just a quick post today to share my New Year's resolution with you. I have made up my mind that this is the year for volunteering (seriously). I am going to find a volunteer job. You may remember that last year I did a couple of short term volunteer jobs which were great, but this year I'd like a long-term position for about a year, so I can really make a difference to someone's life. I've found somewhere to start: [www.iwanttohelp.org](http://www.iwanttohelp.org) - has anyone else used this site?" There is a "Comments" link below the post. A second post is partially visible, dated "SUNDAY, JANUARY 6", starting with "I had a heated debate over lunch today, when my friend Reiner started to quiz me on why I am REALLY interested in volunteering. He was adamant that the only way they can recruit people to volunteer these days is by telling them up front what they'll get out". To the right of the main post is a "Site Contents" sidebar with links for "Home", "About", and "Contact". Below that is an "About Me" sidebar with a cartoon avatar and text: "Life begins at 16 is the personal blog of Maika M. [Read my complete profile.](#)". At the bottom of the browser window, a task is displayed: "IWANTTOHELP: Task 1 [E005Q01] Read Maika's blog entry for January 1. What does the entry say about Maika's experience of volunteering?" followed by four multiple-choice options.

[Next item »](#)

iwanttohelp - Maika's Blog - Home - E005P24 - Internet Browser

Address <http://www.maikasblog.com/index.html>

Maika's Blog

## Life Begins at 16

TUESDAY, JANUARY 1

Happy New Year!  
Just a quick post today to share my New Year's resolution with you. I have made up my mind that this is the year for volunteering (seriously).  
I am going to find a volunteer job.  
You may remember that last year I did a couple of short term volunteer jobs which were great, but this year I'd like a long-term position for about a year, so I can really make a difference to someone's life.  
I've found somewhere to start: [www.iwanttohelp.org](http://www.iwanttohelp.org) - has anyone else used this site?

[Comments](#)

SUNDAY, JANUARY 6

I had a heated debate over lunch today, when my friend Reiner started to quiz me on why I am REALLY interested in volunteering. He was adamant that the only way they can recruit people to volunteer these days is by telling them up front what they'll get out

**Site Contents**  
[Home](#)  
[About](#)  
[Contact](#)

**About Me**  
Life begins at 16 is the personal blog of Maika M.  
[Read my complete profile.](#)

**IWANTTOHELP: Task 1** [E005Q01]  
Read Maika's blog entry for January 1. What does the entry say about Maika's experience of volunteering?

- She has been a volunteer for many years.
- She only volunteers in order to be with her friends.
- She has done a little volunteering but would like to do more.
- She has tried volunteering but does not think it is worthwhile.

Figure 3. Example of an Electronic Reading Assessment test from 2009 PISA program. The passage is presented on a scroll page above the question and multiple-choice response options. Screen shot taken from <http://erasq.acer.edu.au/index.php?cmd=home>.

Figure 4. Example of an Electronic Reading Assessment from 2012 PISA programme. The stimulus and question sections are presented on separate scroll sections. Screen shot taken from <http://erasq.acer.edu.au/index.php?cmd=home>

## 2.5 Online Reading Literacy Assessment (ORLA), Taiwan

The Online Reading Literacy Assessment (ORLA) was designed based on PISA's ERA to explore the characteristics of Taiwanese eighth grade students' online reading literacy (Hung, Huang, Yeh, & Chang, 2010). A total of 176 eighth grade students sampled from three schools in the Tainan districts participated in the ORLA. The ORLA consisted of two subsets of multiple choice items and open-ended questions. The first subset assessed webpage information reading and retrieval, while the second subset assessed online reading and communicating (Hung et al., 2010). Figure 5 presents an example of a multiple-choice item from ORLA.

**Question 3:** You are looking for the information of the tour about Yushan, your friend has sent you this list of four website addresses. Which link would lead to the most reliable information?

- (A) <http://www.ysnp.gov.tw/>
- (B) <http://www.esunbank.com.tw/>
- (C) <http://www.ysticket.com/>
- (D) <http://www.yses.hcc.edu.tw/>



Figure 5. Example of a multiple-choice item from the ORLA. In this example, the question is presented on the top of the page, and a scroll web page is presented below the question. Image taken from Hung et al. (2010).

## 2.6 Partnership for Assessment of Readiness for College and Careers (PARCC), USA

The Partnership for Assessment of Readiness for College and Careers (PARCC) is a consortium of states in the US working together to develop a set of computer-based K – 12 assessments in Mathematics and English Language Arts/Literacy (PARCC, 2015a). The PARCC assessment provides teachers, schools, students and parents information regarding students' learning progress and whether students are on track for success after high school. Results from PARCC are intended to help teachers customise learning and teaching to meet student needs. Around 5 million students in grades 3-11 from 11 different states and the District of Columbia participated in the 2014-2015 PARCC assessments. Figures 6 to 8 present example items for the English Language Arts/Literacy (ELA) tests.

The screenshot shows a web browser window with a navigation bar at the top containing a back arrow, a forward arrow, a 'Review' button, a 'Flag' button, and a user profile icon. The page title is 'PARCC SAMPLE SET GRADE 3-5 ELA / GRADE 3-5 ELA SAMPLE ITEMS / 1 OF 11'. The main content area is divided into two sections. On the left, a passage titled 'The Cricket and the Cougar' by Katherine Chandler is presented in a scrollable box. The passage contains four numbered paragraphs. On the right, there are two sections of multiple-choice questions. 'Part A' asks for the meaning of the word 'master' as used in paragraphs 5 and 6, with four options: understand, conquer, befriend, and frighten. 'Part B' asks which detail from the story best supports the answer to Part A, with four options: 'Don't you know that I am king of the beasts?', 'Well, little boaster, you have that cousin here to-morrow...', 'Then he felt a stinging. 'Oh, oh!' he roared, 'get out of my ear!', and 'The cricket sat on a log and looked on.'

Read the passage from "The Cricket and the Cougar" and answer the questions.

from "The Cricket and the Cougar"  
by Katherine Chandler

1 One day the cougar was out walking in the woods. As he was stepping near an old rotten log, he heard a tiny voice say, "Oh, please don't step there. That's my house, and with one step more you will destroy it."

2 The cougar looked down and saw a little cricket sitting on the log. He roared, "And is it you, weak little creature, that dares to tell me where to step? Don't you know that I am king of the beasts?"

3 "You may be king of the beasts, but I am king of my house, and I don't want you to break it down, king or no king."

4 The cougar was amazed at such daring. "Don't you know, you ~~weakling that I could kill you and your house and all your relatives~~"

**Part A**  
What is the meaning of the word **master** as it is used in paragraphs 5 and 6?

- A. understand
- B. conquer
- C. befriend
- D. frighten

**Part B**  
Which detail from the story best supports the answer to Part A?

- A. "Don't you know that I am king of the beasts?"
- B. "Well, little boaster, you have that cousin here to-morrow..."
- C. "Then he felt a stinging. 'Oh, oh!' he roared, 'get out of my ear!'"
- D. "The cricket sat on a log and looked on."

Figure 6. An example of PARCC ELA sample test with multiple choice items for Grade 3 to 5. The passage/story is presented on a scrolling page on the left. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

Amelia Earhart is a famous American remembered for her daring and bravery. Today you will read two texts and view a video to learn about Amelia Earhart. When you are finished reading, you will write an essay that analyzes the strength of the arguments the authors make in relation to Amelia Earhart's bravery.

Read the website entry "The Biography of Amelia Earhart." Then answer the questions.

The Biography of Amelia Earhart

1 When 10-year-old Amelia Mary Earhart saw her first plane at a state fair, she was not impressed. "It was a thing of rusty wire and wood and looked not at all interesting," she said. It wasn't until Earhart attended a stunt-flying exhibition, almost a decade later, that she became seriously interested in aviation. A pilot spotted Earhart and her friend, who were watching from an isolated clearing, and dove at them. "I am sure he said to himself, 'Watch me make them scamper,'" she said. Earhart, who felt a mixture of fear and pleasure, stood her ground. As the plane swooped by, something inside her awakened. "I did not understand it at the time," she said, "but I believe that little red airplane said something to me as it swished by." On December 28, 1920, pilot Frank Hawks gave her a ride that would forever change her life. "By the time I had got two or three hundred feet off the ground," she said, "I knew I had to fly."

2 Although Earhart's convictions were strong, challenging

**Part A**

In paragraph 6 of "The Biography of Amelia Earhart," Earhart is quoted as saying, **After scaring most of the cows in the neighborhood...I pulled up in a farmer's back yard.** How does the quotation contribute to the meaning of the paragraph?

- A. It demonstrates Earhart's calm sense of humor when describing a potentially frightening situation.
- B. It shows that Earhart loved taking risks but regretted when her actions put others in danger.
- C. It suggests that Earhart was humble about her accomplishments and able to admit mistakes.
- D. It illustrates Earhart's awareness of her responsibility as a role model for other women.

**Part B**

In which other paragraph in the article does a quotation from Earhart contribute to the reader's understanding of her character in a similar way as does the quotation in Part A?

- A. paragraph 7
- B. paragraph 8
- C. paragraph 9
- D. paragraph 11

Figure 7. An example of PARCC ELA sample test with multiple choice items for Grade 6 to 8.

The passage/story is presented on a scrolling page on the left. Screen shot taken from

<http://parcc.pearson.com/practice-tests/english/>

PARCC SAMPLE SET HS ELA / HS ELA SAMPLE ITEMS / 5 OF 23

1:58:23 | Guest

Read the article "Fields of Fingerprints: DNA Testing for Crops." Then answer the questions.

Fields of Fingerprints: DNA Testing for Crops

1 DNA testing, the technique which has helped solve high-profile murder cases, may now help to solve crop crimes. Several organizations have started offering DNA testing to the North American plant breeding and seed industry. Most often, the test will be used by plant breeders and research scientists to identify important genes. But sometimes, DNA testing will come in handy when police are trying to solve crimes that involve grain theft. While it is very difficult to tell the differences in a crop variety just by looking at the seeds, DNA fingerprinting will make it possible for police investigators or researchers to pinpoint specific plant traits and accurately identify seed varieties. Easy to use DNA test kits for certain crops should be on the market within the next few years. Specialized computer-based analysis programs identify the fingerprint, or specific genes carried in the seed of individual crop varieties.

Choose and correctly sequence the steps of the process of developing a DNA fingerprint, as described in the text. Drag and drop each selected step into the appropriate box. Not all steps will be used.

Sort the DNA fragments into a pattern.

Use scissors to cut the DNA into strands of differing lengths and shapes.

Identify the pattern by locating the sites where the radioactive probes bonded to the DNA fragments.

Place particles in a series of sieves to separate them by size.

Add an extraction solution to a small amount of the sample being studied.

Shift the patterns to a nylon sheet and inject radioactive probes.

Add enzymes to the extracted DNA.

First...

Then...

Then...

Then...

Finally...

Figure 8. An example of PARCC ELA sample test for Grade 6 to 8. The passage/story is presented on a scrolling page on the left. A drag-and-drop question/item is presented on the right. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

## 2.7 Smarter Balanced, USA

Smarter Balanced is an online assessment system “aligned to the Common Core State Standards (CCSS) in English language arts/literacy (ELA/literacy) and mathematics for grades 3 to 8 and grade 11” (SBAC, n.d.). Smarter Balanced is designed by the Smarter Balanced

Assessment Consortium, a public agency in the United States, to measure students' progress and attainment of skills and knowledge important for college and students' future career. The system includes both summative assessments and optional interim assessments. Summative assessments are comprised of computer adaptive items and performance tasks which are administered in the last 12 weeks of the school year. Interim assessments, on the other hand, are optional comprehensive and content-cluster measures that are administered at different intervals throughout the school year (as determined by the school) (SBAC, n.d.).

Sample items of English language arts/Literacy tests are presented in Figures 9 and 10.

The screenshot displays a digital test interface. On the left, a passage titled "Grandma Ruth" is presented in a scrollable format. The passage describes a narrator's discovery of a baseball in their grandmother's closet and the subsequent conversation with their grandmother. On the right, a question prompt asks the user to read sentences from the passage and answer a question. The question text is: "My grandma pulled the ball out, unwrapped it, and held it out for us to see. The ball was **scarred** almost beyond recognition. It had dog bite marks, dirt scuffs, and fraying seams. Right in the middle was a big signature in black ink that I had somehow overlooked. It was smudged now and faded, but it still clearly said 'Babe Ruth.' I began to shake inside." Below the question, the instruction reads: "Click on two phrases from the paragraph that help you understand the meaning of **scarred**." The interface also features a score of 43000, navigation buttons (BACK, NEXT, ITEM SCORE), and a "Feedback and Support" link.

Figure 9. Sample item of English language arts/literacy test. Test passage is presented on the left scrolling page while question is presented on the right. For this item, participants are required to identify and select phrases from the passage on the left to answer the question. Screen shot taken from <http://www.smarterbalanced.org/sample-items-and-performance-tasks/>

English Language Arts / Literacy | Mathematics

43016

The following paragraph is an excerpt from a student's report about plant life in the southern United States. After reading the paragraph, you will identify details that are unnecessary and explain why they should be removed.

**The Invasion of Kudzu**

In 1876, Philadelphia, Pennsylvania, hosted the Centennial Exposition in honor of the country's 100th birthday. The Japanese constructed one of the most popular exhibits, primarily due to an amazing vine called *kudzu*. For centuries, the Japanese used the pea vine for many purposes, including medicine, but what attracted the Americans the most were the sweet-smelling purple flowers that covered the vine. The warm, moist climate of the southern states—from Georgia to Florida and north to the Carolinas—was the most hospitable part of the US for the vine. Temperatures in Georgia can reach into the 90s, and the humidity is often above 90%. Residents all over the southeast began planting kudzu. The vine's success caught the attention of many, including Channing Cope who promoted its use for erosion control and animal feed, especially during the Great Depression. Because of the depression, many homes were left abandoned, so there was not anyone to care for them and keep the plants properly groomed. No one predicted, however, that the vine would end up taking over. The vines slowly engulfed pine forests, telephone poles, and even houses, leading to the destruction of native plant life. Pines are not the only trees in the South, however. There are about 250 species of trees in Georgia alone. As kudzu out-competed the local plants, it deprived them of nutrients and, especially, sunlight. Kudzu now covers over 7 million acres of land, and it continues to expand at the rate of 150,000 acres annually. That is almost one foot per day! Attempts to kill it have proven difficult, as it is immune to most herbicides; thus, kudzu continues to smother the southern states. Researchers continue to search for a solution to "the vine that ate the South," but the answer is nowhere in sight.

In the space below, identify the sentences from the paragraph that are unnecessary, and briefly explain why each one should be removed.

Figure 10. In this sample test, the test passage is presented on top while the question is presented on the bottom. For this item, participants are required to type their answers in the box. Screen shot taken from <http://www.smarterbalanced.org/sample-items-and-performance-tasks/>

## 2.8 Programme for the International Assessment of Adult Competencies (PIAAC)

The Programme for the International Assessment of Adult Competencies (PIAAC) is an international assessment of key cognitive and workplace skills (OECD, n.d). Adults between the ages of 16 and 65 in 33 countries have participated in the PIAAC in previous years.

The three main components of the PIAAC survey are:

- i. Direct assessment of skills in literacy, numeracy, reading components and problem solving in technology-rich environment (PS-TRE).
- ii. Modules on the use of cognitive skills (e.g., reading, writing), interaction and social skills (e.g., planning, communication), physical skills (e.g., use of gross and fine motor skills) and learning skills.

- iii. Background questionnaire to collect participant details such as demographic characteristics, education level, and employment status and income.

The PIAAC literacy test assess adults' ability to read digital texts (e.g., text containing navigation features such as scrolling) as well as traditional print-based texts.

An example of a computer-based version of the literacy test is presented in Figure 11.

Look at the list of preschool rules. Highlight information in the list to answer the question below.

What is the latest time that children should arrive at preschool?

### Preschool Rules

Welcome to our Preschool! We are looking forward to a great year of fun, learning and getting to know each other. Please take a moment to review our preschool rules.

- Please have your child here by 9:00 am.
- Bring a small blanket or pillow and/or a small soft toy for naptime.
- Dress your child comfortably and bring a change of clothing.
- Please no jewelry or candy. If your child has a birthday please talk to your child's teacher about a special snack for the children.
- Please bring your child fully dressed, no pajamas.
- Please sign in with your full signature. This is a licensing regulation. Thank you.
- Breakfast will be served until 7:30 am.
- Medications have to be in original, labeled containers and must be signed into the medication sheet located in each classroom.
- If you have any questions, please talk to your classroom teacher or to Ms. Marlene or Ms. Tree.

Figure 11. An example of the computer-based PIAAC literacy test. The item or passage is presented on the right while instruction and questions are presented on the left. To answer the question, respondents are required to highlight words and phrases in the stimulus on the right using a mouse. Screen shot taken from

<http://www.oecd.org/site/piaac/Literacy%20Sample%20Items.pdf>

## 2.9 Review of online assessment tools and programs

In this section, we review and highlight characteristics and features that are salient in the online assessment tools described above.

Table 1 summarizes the online assessment tools and programs described in Section 2.1 to 2.8. Across all online reading assessment tools and programs, we observed four common features and characteristics.

First, in most online assessment, stimulus and items were presented in either a top-bottom or left-right format or a combination of both. That is, passages or stimuli are presented on the top of page, with items or questions presented below. Alternatively, the passage is presented on the left while questions and items are presented on the right or vice-versa. To our knowledge, there is no evidence that points to a best stimuli-item layout format. Online assessment programs that employ their respective layout formats did not provide reasons for their decision. These layout formats were commonly paired with scrolling features which enable the readers to scroll and move through the text. Additionally, most passages were presented on a white background, and add-on graphics such as pictures were presented within the passage section.

Second, although placed in different parts of the screen, “back”, “forward”, “previous” or “next” buttons were used as navigational methods on all online assessments reviewed above.

Third, a range of reading tools and features were incorporated into the online reading assessments. These tools were included to facilitate the online reading process. Examples of tools include:

- Item panel or progress bar; this was commonly presented on the top of the page and indicated the number of items that were included in the assessment, or the percentage of assessment completed.
- Flag feature; this served as a ‘reminder’ for readers to review items later.
- Highlight tool for reading passages.
- Ability to expand or maximise stimulus window/passage section.
- Pop-up notepad for separate note-taking.

Lastly, the majority of the assessments included multiple choice questions and/or constructed response items (i.e., short answers). The technologies of online and computer assessments also allowed for new and innovative types of questions which were not feasible

on traditional paper-and-pen assessment. Technology-enhanced items such as drag-and-drop questions, point and click, and select or highlight phrases from a passage were included in some assessment tools.

Table 1

Overview of online assessment tools

Online tools/programs	Age range/ Year levels	Domains tested	Types of questions	Adaptive testing	Presentation layout (stimuli-items)	Navigation features	Other characteristics
PAT Reading	Reading comprehension: Foundation to Year 10  Word knowledge: Year 3 to 10	<ul style="list-style-type: none"> <li>• Reading comprehension</li> <li>• Word knowledge (vocabulary)</li> <li>• Spelling - not available for online administration</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple choice</li> </ul>	x	-	-	-
On Demand Testing	Year 1 to 10	<ul style="list-style-type: none"> <li>• English</li> <li>• Mathematics</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple choice</li> <li>• Short answer</li> <li>• Hot spot</li> <li>• Drag and drop</li> <li>• Mathematical calculator</li> </ul>	✓	-	-	-
OLNA	Year 10 to 12	<ul style="list-style-type: none"> <li>• Reading</li> <li>• Writing</li> <li>• Numeracy</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple choice</li> <li>• Constructed response (writing component only, responses of up to 600 words allowed).</li> </ul>	x	-	-	-

Online tools/programs	Age range/ Year levels	Domains tested	Types of questions	Adaptive testing	Presentation layout (stimuli-items)	Navigation features	Other characteristics
PISA	15-year olds	<ul style="list-style-type: none"> <li>• Reading</li> <li>• Mathematics</li> <li>• Scientific literacy</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple choice</li> <li>• Constructed response</li> </ul>	*	<ul style="list-style-type: none"> <li>• Top-bottom</li> </ul> <p>Story presented within a scrolled textbox.</p>	<p>Item panel on top-centre of page.</p> <p>“Previous” button on left hand corner; “Next” button on right hand corner.</p>	<p>Pop-up stimulus with scrolling features; use mouse clicks to move pop-up screen; can also close pop-up window.</p>
PISA-ERA (2009, 2012)	15-year olds	<ul style="list-style-type: none"> <li>• Online reading literacy/ Digital reading</li> <li>• Reading of electronic texts</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple choice</li> <li>• Point and click</li> </ul>	*	<ul style="list-style-type: none"> <li>• Top-bottom</li> </ul> <p>Web-page type stimulus with scrolling feature.</p> <p>For some items, participants would click on correct link/answers which are presented within story/text. There are also short answer items (e.g., reply to a forum posting).</p>	<p>Back and forward buttons on left hand corner.</p>	<p>Dynamic text (hypertext) within stimuli.</p>

Online tools/programs	Age range/ Year levels	Domains tested	Types of questions	Adaptive testing	Presentation layout (stimuli-items)	Navigation features	Other characteristics
ORLA	Eighth grade students (Taiwan)	<ul style="list-style-type: none"> <li>• Online reading literacy</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple choice</li> </ul>	x	<ul style="list-style-type: none"> <li>• Bottom-top</li> </ul>	-	-
PARCC	Kindergarten to Grade 12 (USA)	<ul style="list-style-type: none"> <li>• English language arts/literacy</li> <li>• Mathematics</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple choice</li> <li>• Drag and drop</li> <li>• Constructed response</li> </ul>	x	<p>A mixed of top-bottom and left-right layout formats.</p> <ul style="list-style-type: none"> <li>• Left-right</li> </ul> <p>Scrolled text; response text box with bold, italics, underline and numbering features.</p> <ul style="list-style-type: none"> <li>• Top-bottom</li> </ul> <p>Compare ideas from two stories. Drag sentences and drop them into the Venn diagram.</p>	<p>Back and forward arrow buttons are presented on top left.</p>	<p>Presentation of two passages using tabs.</p> <p>Features/tools:</p> <ul style="list-style-type: none"> <li>- Highlight tool for reading passages</li> <li>- Review (brings you to a different page, with an overview list of questions and whether you have answered or not)</li> <li>- Flag feature (i.e. flag a question to review later)</li> <li>- Pointer</li> <li>- Notepad (a pop-up blank note pad that can be used to make notes, movable with mouse click-drag, auto-saved, can be closed and reopened with</li> </ul>

Online tools/programs	Age range/ Year levels	Domains tested	Types of questions	Adaptive testing	Presentation layout (stimuli-items)	Navigation features	Other characteristics
							content saved from last use - Answer eliminator (big red cross that can be used to cancel out or eliminate wrong answer; a visual reading aid)
Smarter Balanced	Grades 3 to 8, and Grade 11	<ul style="list-style-type: none"> <li>• English language arts/literacy</li> <li>• Mathematics</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple choice</li> <li>• Technology-enhanced items (e.g., selecting phrases from paragraph)</li> <li>• Constructed response items</li> <li>• Performance tasks</li> </ul>	✓	<ul style="list-style-type: none"> <li>• Left-right</li> </ul> Passage presented within a scrolled section.  <ul style="list-style-type: none"> <li>• Top-bottom</li> </ul>	Back and forward arrow buttons are presented on top right.	Features/tools: <ul style="list-style-type: none"> <li>- Ability to maximize or minimize stimulus window (left part of screen)</li> <li>- Flag item</li> <li>- Optional (but immediate) item score for certain items such as multiple choice questions</li> </ul>
PIAAC	Adults	<ul style="list-style-type: none"> <li>• Literacy</li> <li>• Numeracy</li> <li>• Reading components</li> <li>• Problem solving</li> </ul>	-	x	-	-	-

## 3 Comparison between Paper and Online Assessments

---

Improvements in technology have encouraged the delivery of educational assessments through computers (Kim & Kim, 2013) and this has opened new possibilities for the assessment of reading and/or literacy skills. The advantages of online assessments have been reported in terms of their ability to provide immediate scoring and feedback of students' results, efficiency in test administration, flexible administration schedules, reduced costs and the possible inclusion of multimedia item types that would not be feasible for print assessments (Gil, Martinez, & Vidal-Abarca, 2015). However, an important question is whether tests administered in both formats generate similar test outcomes.

### 3.1 Paper-based Assessments

With regards to paper-based assessments, research has identified a number of strategies used by readers to acquire, understand and retrieve information they read (Huang, 2014). Readers have to make decisions about what and how to read the text including skimming text rapidly, scanning text for a specific piece of information, engaging in detailed and deep reading, or re-reading text (Gil et al., 2015). The question then becomes whether reading skills and comprehension processes change if the medium changes, and if achievement results reflect these changes.

### 3.2 Online Assessments

There is a substantial body of research examining computer-based tests and paper-and-pencil tests. However, results regarding performance related to test delivery mode are mixed. Some have found that students performed better on computer-based assessments (Clariana & Wallace, 2002). In this study, freshman business undergraduates [age details not provided] either completed a computer-based or identical paper-based test. Results indicated that the computer-based test group outperformed the paper-based test group.

Other studies have found lower scores in computer-based assessment compared to paper based assessments (Kim & Kim, 2013). The researchers examined differences in reading performance for an electronic test with a scrolling text mode on a LCD monitor versus a

traditional paper test format. Participants were high school students (Grade 11 students in the US) and results showed that student performance was affected by the presentation medium such that students scored significantly higher in reading comprehension on the paper test, and took a longer time to read and answer questions on the electronic tests (Kim & Kim, 2013).

There are studies that have found no differences as well. A meta-analysis consolidating findings on the impact of administration mode for K-12 reading assessments (i.e. computer based tests and paper-pencil tests) identified that administration mode had no statistically significant effect on reading achievement (Wang et al., 2008). Their research focused on published and unpublished studies on computer-based and paper-based administration, examining effects on K-12 student' reading achievement performance. Thirty-six relevant studies were analysed and the authors found that the difference between students' reading achievement from computer-based or paper based tests was not statistically significant (Wang et al., 2008).

Young (2014) investigated the differences in university student readers' [age details not provided] information gathering processes and strategies for comprehension of information when reading online and on print, and retention of information. Participants were provided with print and web edition texts from three publications, and students' ability to read and retrieve core information from both media were assessed. Subsequent focus group discussions probed their strategies used to engage with the texts in both media. Their standardized reading comprehension scores were also collected. Results indicated that participants demonstrated functional equivalency in both media (Young, 2014).

Researchers have discussed various hypotheses to explain such mixed results. A majority of early comparability studies on paper versus computer documents focused on outcome measures of reading such as speed, accuracy and comprehension (Solak, 2014). There is a burgeoning consensus that the inconsistencies in earlier findings could be related to variations in visual quality of the two presentations and developments in display screen and interface technologies could reduce the level of disparity between the presentation qualities of the two media (Wang et al., 2008). Mason, Patry and Berstein (2001) flagged the flexibility of platforms (i.e. computer interfaces in some studies did not allow students to

skip, review and/or modify responses) as a factor while Paek (2005) highlighted challenges related to scrolling and strategies used to organise information (e.g. underlining key phrases).

### 3.2.1 Web page reading

Despite the inconclusive evidence from studies comparing paper-based and online assessment, evidence from eye tracking studies of adult web readers revealed that readers scan and read through web pages differently compared to reading on paper – online readers tend not to read every word on a web page. In a study investigating how people read websites, Nielsen and colleagues (1997) found that 79% of users scanned through the web page and only 16% read word-by-word. Generally, readers scan through a web page until some element of the page catches their eye and attention (Tortorici, n.d.). At this point, the reader will switch from scanning to focusing on the element and read until he/she is no longer engaged or loses motivation to carry on. The reader would then switch back to the “scanning mode” until another element attracts his/her attention. This scan-and-focus pattern of reading repeats itself over and over during web-page reading (Tortorici, n.d.).

Some researchers suggest that readers follow an “F-pattern” when browsing or reading web content (Jones, 2012; Nielsen, 2006) (see Figure 12). The F-pattern describes the most common eye-scanning patterns observed in web page reading, and consists of the following three components (Nielsen, 2006):

- Users first read in a horizontal movement across the first few lines on the top of the screen or across the upper part of the content area. This initial movement forms the F’s top bar.
- Next, users move down the page slowly and read across in a second horizontal movement, this time covering a shorter area than the first horizontal movement. This subsequent movement forms the F’s lower (shorter) bar.
- Finally, users scan the left side of the screen in a vertical movement, looking for keywords or elements of interest in the paragraphs or subsections that follow. This last movement forms the F’s stem.

According to Nielsen (2006), readers employ the “F reading-pattern” when browsing and reading almost all online material, including articles, web sites, search engine results and etc. Users’ reading behaviour such as the F-pattern must be taken into consideration when designing web pages. Guidelines for webpage design indicate that an “F-layout” would work best as it allows users to scan web content naturally (Jones, 2012). Using the F-Layout, the best or most important content of a web page should be placed across the top, and the first two paragraphs should state the most important information (Jones, 2012; Nielsen, 2006). The following subsections, paragraphs or bullet points should start with information-carrying words so that users will notice these sections when scanning down the left side of the screen or content area – this vertical scanning movement is the final component of the F-pattern.



Figure 12. The F-pattern illustrated in heat-maps from user eye tracking of three websites. The areas where users looked the most are coloured in red, the yellow areas indicate fewer views, followed by blue areas which indicate the least-viewed areas. Grey areas indicate areas that did not attract any fixations. Image taken from

<http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/>

It is important to note that these studies were conducted on adult participants, and examined web page reading behaviour and not online test taking. Nevertheless, these

findings may have important implications for the design of online reading assessments for children. Further research is needed to answer the following questions:

- What reading strategies (e.g., “scan-and-focus”, “read word-by-word”, “F-reading pattern”, or other strategies) do students use when completing an online reading assessment?
- How would students’ reading strategies influence the layout and presentation of elements of an online reading assessment (i.e., passage and item placement)?
  - For example, if students tend to utilise the F-pattern reading strategy in an online reading assessment, how should assessments items be presented or arranged on a screen page?
- How do students’ reading strategies affect their ability to answer items in an online assessment?

### 3.3 Factors to Consider for Online Assessment

There is not a strong body of research evidence concerning the impact of testing modes on student performance and the tests' psychometric properties (Gil et al., 2015). With respect to reading assessments, examining the readability of onscreen text is essential to ensure effective student engagement (Hojjati & Muniandy, 2014). Readability focuses on how well a reader is able to comprehend the content of a particular text through reading (Eslami, 2014). There are a number of factors which can impact on the ability to read text on a computer screen. The following section provides a critical analysis of research on readability factors including typeface characteristics, text layout and format, presentation layout, and device properties and features.

## 4 Critical Analysis of Factors Important for Online Assessments

---

### 4.1 Typeface Characteristics

Display or typeface characteristics refer to issues related to fonts such as font type, font size and character spacing – many of which have been subjected to detailed research for both print and screen display. Research findings on typeface characteristics can guide ones understanding of the relative advantages and disadvantages of particular display types, and also highlight the effects of several interacting factors on reading from screen. However, guidelines of typeface characteristics for print may not be directly applicable to onscreen display.

#### 4.1.1 Font Type

In early research, Dillon (1992) in his review looked at differences between reading from print and screen, and found that font type for onscreen texts did not impact on reading rate after taking into account reasonable font type and screen resolution. In his doctorate thesis, Weisenmiller (1999) also examined the readability of four different font typefaces (i.e. Georgia and Verdana which were both optimised for onscreen readability, and Times New Roman and Arial) and the impact of font type on reading rate and comprehension for university students (i.e. 264 participants with median age 21.5 years). He found no significant differences in reading rate and comprehension between any of the typefaces (Weisenmiller, 1999).

However, with the advent of new fonts, recent studies have illustrated differences between Serif and san Serif fonts for onscreen readability. Sans serif typefaces (e.g. Arial and Comic Sans MS) were perceived as easier to read and more desirable for online texts compared to Serif typefaces (e.g. Times New Roman and Courier New), with a significant effect found for reading accuracy (Sans serif fonts required less effective reading compared to the other text types) when effects of common online typefaces effect on reading was explored for elementary school-aged children (i.e. 27 children aged 9 to 11 years old) (Bernard, Chaparro, Mills, & Halcomb, 2002). Hojjati and Muniandy (2014) examined the readability of different font types (serif vs san serif) and the effects of font type on reading performance in 30

postgraduate students (age details not provided). They reported a significant difference in font readability between serif fonts (e.g. Times New Roman) and sans serif fonts (e.g. Verdana), with the latter identified as the better choice for displaying long text on screen.

Josephson (2008) conducted an exploratory eye-tracking study comparing onscreen legibility of sans serif fonts (e.g. Arial and Verdana) to serif font (e.g. Times New Roman and Georgia) and determined that Verdana performed the best, with participants (i.e. 6 university students - average age 24.2 years) reading more quickly, experiencing less backward movements for review or re-reading. Participants also expressed a preference for the Verdana font over other fonts.

#### 4.1.2 Font Size

As expected, research found that bigger font sizes were perceived as significantly easier to read on a computer screen. In his review, Dillon (1992) observed the differences of display fonts between screen and paper reading, and argued that optimum character size for reading from screen is likely reliant on the task. Bernard et al. (2002) explored actual and perceived readability for different typefaces at 12 and 14 point sizes and determined that although there was no significant effect on reading accuracy or speed for point size, participants (i.e. 27 children aged 9 to 11 years old) considered 14 point size more desirable for school materials. Another related study examined readability of Times New Roman and Arial in 10 point and 12 point sizes for university students (i.e. 35 under-graduate and graduate students with mean age 25 years) and found that there was no significant difference in reading accuracy for font size. However, font size impacted students' reading speed (i.e. 10 point text was read slower than all 12 point text) (Bernard, Chaparro, Mills, & Halcomb, 2003). Participants preferred the 12 point text and perceived it to be less difficult to read than the 10 point text. In another review, Dyson (2004) also reported that reading with 12-point font was faster but noted that increasing font size did not always improve participants' perception of text legibility.

## 4.2 Text Format

Presentation of text on a screen can be influenced by several variables such as interlinear spacing, line length, column format and window size; these variables are used to vary the

amount of text displayed on screen. These typographic variables are inter-related and together they can affect the layout and readability of text. These variables are not generally considered in isolation; therefore, the following section will review studies that have considered or manipulated one or more factors related to text format.

#### 4.2.1 Interlinear Spacing

Interlinear spacing refers to the amount of space between lines (Dyson, 2004). Control of spacing is a significant part of page design as loose spacing can result in pages 'streaming' and reduce legibility (Hojjati & Muniandy, 2014). There is evidence to suggest that spacing between the lines of a text can impact on speed of on-screen reading. Dyson (2004) consolidated research on variables related to configuration of text on screen and found that double spacing was better for reading than single spacing. Ling and van Schaik (2007) conducted a study on 65 undergraduate students (77% were 25 years and under; 23% aged between 25 and 50 years) to examine the impact of line spacing on students' reading speed and accuracy in visual search of web pages. Participants were presented with a series of web pages which appeared in single-, 1.5- or double spacing. Texts were either left-aligned or justified depending on the experimental condition. Participants were then asked to find a hyperlink in a screen of text. Results revealed that participants performed better with double spacing than 1.5 spacing, and better with 1.5 than single spacing (Ling & van Schaik, 2007).

Hojjati and Muniandy (2014) also tested the effect of single and double spacing on ease of reading, retention and time taken to read with 30 postgraduate students and found a statistically significant main effect of line spacing on ease of reading onscreen. These results are consistent with findings from older studies (Kolers, Duchnicky, & Ferguson, 1981; Kruk & Muter, 1984) who reported slower reading speed when on-screen text was presented with single spacing compared to double spacing. Together, these results seem to suggest that larger interlinear spacing (e.g. double spacing) may be better for reading on-screen texts.

However, some studies comparing spacing options did not report significant difference in reading speed and comprehension among different conditions of line spacing (Ni, Branch, Chen, & Clinton, 2009). A 2002 study tested college students' reading speed for on screen single- and double-spaced passages and determined that there was no statistically

significant difference in performance between the groups although the ‘double-space’ group took a relatively longer time to complete reading (Loh, Branch, Shewanown, & Ali, 2002). This paper did not provide information on the font types on the passages, and there is evidence that different combinations of text formatting could result in outcome differences [e.g. sans serif fonts (Verdana) in double spacing was considered easier to read than serif fonts (Times New Roman)] (Hojjati & Muniandy, 2014).

Nevertheless, it is important to consider interlinear spacing in the context of line length as some researchers have suggested that longer lines require increased or larger spacing to improve readability (Dyson, 2004).

#### 4.2.2 Line Lengths

Line length refers to the number of characters in a line, or the physical length of a line (Dyson, 2004). For text presented in print, line lengths of less than 70 characters per line are recommended. However, the recommended line length for onscreen text is not as clear cut. This is because line lengths of text presented on screen is easily varied by changing font size, window size or adjusting page margins.

To date, studies investigating the optimal use of line length on screen have produced mixed results. Dyson and Kipping (1998) conducted a study to examine the impact of line length on adult’s (aged between 18 and 44 years) reading rate and comprehension. Passages of about 800 words were presented on a computer screen, and six line lengths ranging from 25 to 100 characters per line were compared. This was achieved by changing the size of the right margin while keeping type size the same across all conditions. They found that the number of characters per line affected participants’ reading rate, such that the longest line (i.e., 100 characters per line) was read faster than the shortest line (i.e., 25 characters per line). Similar findings were reported by Duchnicky and Kolers (1983) who found on-screen text with longer line lengths and density was read faster than text with shorter line lengths or density.

Conversely, one study found no difference in reading time or reading efficiency between three line length conditions (45, 76 and 132 characters per line) for children (ranged in age from 9 to 12 years; mean age of 11) and adults (ranged in age from 18 to 61 years; mean age of 29)(Bernard, Fernandez, & Hull, 2002). Their results did, however, indicate that

children preferred the shortest line length (45 characters per line), while adults preferred medium or longer line lengths. Consistent with this finding, some researchers have also recommended medium or shorter line length (around 60 characters per line), stating that readers are more likely to lose their space within the text when reading longer line lengths as they require greater lateral eye movement (Mills & Weldon, 1987).

It is important to note that these studies have employed different methods of manipulating line length; some studies manipulated line length by changing the size of the right margin, while other studies manipulated line length by adjusting screen width (e.g., utilising the full length of a screen or parts of screen). These methods produce different text layout and presentation, which may affect the overall readability of text and also influence reading speed and performance. Evidence from studies investigating optimal line length for online reading remains inconclusive. However, it appears that shorter line length ranging from 45-60 characters per line may be most suitable for children. Further investigation into the optimal line length for online reading assessment is needed.

#### 4.2.3 Length of text or passages

There is limited research about the length of texts on a page although it can be assumed to impact on text complexity. PARCC has set the following grade-level expectations to inform text lengths of literacy or informational passages selected for assessments (PARCC, 2014b):

- Grade 3-5 : 200 – 800 words
- Grade 6-8: 400 – 1000 words
- Grade 9-11 – 500 – 1500 words

However, PARCC also identified that the appropriateness of the text, more than text length, should guide text selection. Length of text in a web page is a balance of the relationship between page and screen size as well as the content of your documents. One web style guide suggests that navigational pages should not have more than two screens worth of information, and navigational links should be integrated at the beginning and the end of the page (Lynch & Horton, 2008). Another web guideline suggests that page lengths should be at most twice the screen height for navigational pages to improve reading and reduce disorientation while scrolling (WARC, 2012).

#### 4.2.4 Columns

Another important variable in text layout is column format, especially in relation to line length. Text on screens can be laid out in single or multiple columns, and depending on device type, readers can move through the text using a vertical or horizontal paging or scrolling function. When text is presented in multiple columns, the length of each individual line decreases (Baker, 2005). This also subsequently reduces lateral eye movement involved in reading long lines of text. Compared to other text layout variables, there has been little research on column formats, especially for online or on-screen environments. Older studies have reported that two-column layouts improved reading performance of text presented on print (Pinelli, Cordle, & McCullough, 1986; Williamson, 1966). However, some findings also suggest that multi-column layouts are too narrow and can reduce reading rate (Burnhill, Hartley, & Young, 1976; Hartley, Burnhill, & Fraser, 1974).

In an earlier study comparing column and navigation method (e.g., paging vs. scrolling), Dyson and Kipping (1997) recruited 18 participants (18 to 44 years) who were each asked to read three passages displayed on screen; two presented in a single column, scrolled and paged, and a three-column paged passage. The single paged column was read faster than the other formats. In addition, when comparing paged movement conditions, text displayed in a single column was read faster than a three column format. However, participants rated the three column format as easier to read compared to the single column format (Dyson & Kipping, 1997).

A more recent study investigated the effects of multi-column displays and justification on online reading performance. Baker (2005) presented participants (mean age of 22.8 years) with a short story displayed in one of six formats (one, two, or three columns, with either full or left-justified format). Results from this study revealed that reading speed was significantly faster for two-column full-justified text than for one-column full-justified text. Reading speed was also significantly faster for one-column left-justified than for one-column full-justified or three column full-justified. They also reported some differences between slow and fast readers. Fast readers performed best when reading texts presented in two-column full-justified format, while slow readers performed best in a single column non-justified layout (Baker, 2005).

Overall, the limited amount of research on column formats has produced mixed findings. Nonetheless, across all studies, three-column texts were read the slowest. Research findings also suggest that students' reading ability and justification of text had an impact on students' reading performance. Therefore, further investigation is still required to examine the optimal way of presenting or choosing column formats for online or on-screen texts.

#### 4.2.5 Justification of text

There are several types of text justification such as left justification (also referred to as no justification) and full justification. Left justification provides equal spacing between words, without hyphenated word-wrapping at the end of sentences. Full justification enables unequal spaces between words and letters to create an even margin on both sides, often including hyphenated breaks in words at the end of sentences (Woody, n.d.).

Left justification produces a ragged right margin which may appear messy. On the other hand, full justification produces neat and vertical margins. However, the uneven variation in spacing between words causes the text to be more difficult to read as the reader's eye has to jump from word to word rather than move through a sentence smoothly along the line (PWS, 2009). With full-justified texts, large gaps between words and sentences may create distracting white patterns through the text – this effect is known as “rivers of white” or “river effect” (PWS, 2009). Full-justified texts are especially difficult to read for readers with disabilities such as dyslexia (Gregory & Poulton, 1970). The uneven spaces between words become visual distractions and causes readers with dyslexia to lose their place repeatedly (Pedley, 2006).

Most web design guidelines suggest that that left justified texts are better for on-screen display as text lines are easier to follow and read. A study by Ling and van Schaik (2007) examined the effects of text presentation on online behaviour and preferences of 65 undergraduate participants. The purpose was to generate design guidelines for web page production. They reported that text presentation had a significant effect on usability of a web page. Specifically, they found that left-aligned text led to better performance in a visual search task of web pages, even though participants preferred full justified text. Similar findings were reported in a study by Baker (2005) where reading speed was fastest under the one-column left-justified and two-column full-justified conditions. Taken together, these

findings seem to suggest that left-justification may be most suitable for web pages and online text. Further studies are still needed, however, to examine the effects of text justification on students' performance in online reading assessments.

#### 4.2.6 Display Size

Window size or height, sometimes referred to as display size, is an important aspect of electronic text presentation across device types. Display size can be varied by changing the number of lines per screen. When comparing display size on screen and paper, differences have been found. Kruk (1984) reported that 40 lines were read faster than 20 lines in either print or on screen, with no differences in reading comprehension scores. The author's findings indicate that large display size may facilitate reading speed.

Other studies compared display sizes for on-screen reading. In Dillon et al.'s (1990) study, 32 subjects (mean age of 29 years) read an onscreen academic journal article for comprehension with no time limit enforced. The article was presented to them on a screen with display size of either 20 lines or 60 lines. There was no significant effect of display size on reading comprehension or reading rate. However, subjects did indicate a preference for the larger screen display than display size. Additionally, participants using the small screen navigated more between pages and changed direction of viewing more often than participants reading from the large screen.

Richardson (1988) conducted a similar study employing an information location task. Participants were asked to search for 10 pieces of information presented in a software reference manual, displayed on a screen with 20 lines or 40 lines. The total amount of time taken to complete the task did not differ significantly for both window size conditions. Again, subjects expressed a preference for the larger window display size. Taken together, these results suggest that although window size did not affect comprehension or performance rates on screen, subjects in both studies indicated a preference for larger display size.

The majority of studies discussed above were conducted in the 1990s or earlier, and as such these studies have focused on computer window size. Over time, with advances in technology, new devices such as tablets and e-readers have been developed. The conclusions and guidelines drawn from computer studies should not be generalised or

applied to other devices as there are inherent differences in display and features of newer devices compared to a computer. Therefore, further investigation exploring display sizes in newer devices such as tablets is warranted.

### 4.3 Presentation Layout

Presentation layout focuses on the arrangement, layout and design of items and stimulus on an online page. Almond, Steinberg and Mislevy (2002) termed this idea a presentation model, which essentially describes how tasks or items will look and feel when presented in the delivery environment. Literature indicates that information content transferred from print to digital display may affect reading comprehension. Research comparing computer-based tests (CBTs) and pen-and-paper tests (PPTs) have reported mixed findings in relation to reading speed, accuracy, comprehension, cognitive process load and cognitive measures (Noyes & Garland, 2008). This has highlighted the need for further investigation into the differences between elements of CBTs and PPTs in order to bridge the gap between these two modes. The following section will discuss and review the various methods that researchers have employed or suggested during or as results of their experiments.

#### 4.3.1 Item Placement

Item placement implies the notion of assembling and positioning multiple elements of an assessment (such as stimulus, question, response options, response format and graphics) on screen. Decisions about how information is laid out or where various elements are positioned should be made in view of other critical factors such as typeface characteristics and text formats.

An on-screen presentation might require a different layout compared to pen-and-paper presentation. Therefore, careful consideration should be given when deciding how elements of an item are placed online, and how much information should be displayed simultaneously on an online page. These are important factors to consider in item placement as they can influence how information is perceived and processed by students. In addition, display size in digital medium is relatively limited, and screen display size tends to be smaller on tablets than on computer. Therefore, item placement and design needs to take into account these factors so that an assessment will appear equivalent across devices.

There is a lack of research evidence on the effects of item placement and organisation on reading speed, accuracy, comprehension or cognitive processes. Despite this, there seems to be a common convention across standardised and large scale assessments in regards to item placement and position on screen. The stimulus is commonly presented on the left with items (question and response options) presented on the right (e.g. PARCC, SBAC, PIACC, PISA). Alternatively, the stimulus is presented on the top followed by items at the bottom (e.g. ERA, ORLA); this layout is adapted from paper-based tests and most likely due to the ease of use and to compliment familiarity with paper tests. In addition, the majority of online assessments have included only one question (e.g., multiple choice, constructed response, drag-and-drop and etc.) on each online page, with the following question displayed on the next page and accessible by clicking the 'Next' or 'Forward' button. Overall, most reading assessment utilised fixed texts or prose (such as narration, exposition and argumentation) which comprised sentences and paragraphs. The amount of text that is immediately visible when displayed on screen encourages readers to approach the text in a particular sequence (OECD, 2009, 2013). Depending on graphic placement, the inclusion of graphical elements within a text may influence the way readers approach and navigate through the text. The following section discusses briefly about graphic placement in online assessments.

### ***Graphic placement***

Graphic placement refers to the arrangement and position of graphics in relation to the page or text. There are different types of graphics, including drawings, diagrams, photographs, tables, charts and graphs. When graphics are included in a passage of text, the surrounding text could wrap around the graphic on one or both sides, or be overprinted behind or in front of the graphic. Alternatively, the graphic could be treated as a separate paragraph or character.

There is very limited information or guidelines regarding graphic placement for the online assessment environment. Based on the review in Section 2, it is observed that some online assessments do not include any graphics such as picture or illustration, while others such as the PARCC assessment, include graphics (e.g. pictures or photos) in their passages; these graphics appear at the beginning of the passage (Figure 13) and/or within the text between paragraphs (Figure 14). The effectiveness and readability of such graphic placement is

unknown and should be tested, especially if graphics are included to supplement the text. For example, when a supporting graphic (e.g., a diagram showing the life cycle of a butterfly) is included to supplement the text, the placement of the graphic is particularly important because the graphic serves as a visual aid and provides students with additional information regarding the text. To our knowledge, there are no rigid guidelines or hard-and-fast rules for presenting graphics in an online reading assessment. It is recommended that different graphic presentation or layout options be trialled and tested with students of different ages to determine the best and most suitable way of presenting graphics in online reading assessments. Further recommendations relating to graphic and item placement for NAPLAN online is discussed in Section 5.1.2.

The screenshot shows a web-based assessment interface. At the top, there are navigation buttons (back, forward), 'Review' and 'Flag' buttons, and a 'Guest' user profile. Below the navigation bar, the page title reads 'GRADE 03 ELA/LITERACY PRACTICE TEST / UNIT 1 / 6 OF 12'. The main content area is divided into two columns. The left column contains a reading passage titled 'from "What Is a Spacewalk?" by NASA'. It features a photograph of an astronaut in space, with the caption 'Ed White was the first American to perform a spacewalk.' and the source 'NASA—Public Domain'. Below the photo is a section titled 'Why Do Astronauts Go on Spacewalks?' with a numbered list item: '1 Astronauts go on spacewalks for many reasons. Spacewalks let'. The right column contains two parts of a question. 'Part A' asks for the meaning of the word 'thrusters' as used in paragraph 8, with four multiple-choice options: A. ropes that connect, B. machines that power, C. spacesuits that protect, and D. bags that hold tools. 'Part B' asks which detail from the passage helps understand the meaning of 'thrusters', with four multiple-choice options: A. "... hooked to the spacewalker." (paragraph 7), B. "... is worn like a backpack." (paragraph 8), C. "... fly back to the spacecraft." (paragraph 8), and D. "... with a small joystick ..." (paragraph 8).

Figure 13. Photograph was placed in the beginning of a passage from the Grade 4 Literacy test for PARCC assessment. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

← → Review Flag

GRADE 04 ELA/LITERACY PRACTICE TEST / UNIT 1 / 7 OF 12

provide protein for their families—quite a plant! In 1896 peanuts were not even recognized as a crop in the United States, but Carver would help change that.



CREDIT: Johnston, Frances Benjamin. Photographs "George Washington Carver, full-length portrait, standing in field, probably at Tuskegee, holding a piece of soil." 1906 Booker T. Washington Collection, Prints and Photograph Division, Library of Congress

**George Washington Carver gathering soil samples.**

6 Carver told farmers to rotate their crops: plant cotton one year, then the next year plant peanuts and other soil-restoring plants, like

**Part A**  
What is a main idea of the article?

- A. George Washington Carver was a bright young man and an excellent teacher.
- B. George Washington Carver taught farmers how to improve crop production.
- C. George Washington Carver learned about caring for plants as a young boy.
- D. George Washington Carver was recognized by Congress for his accomplishment.

**Part B**  
Which detail from the article supports the answer to Part A?

- A. "When he was a child, he was known as the 'plant doctor.'" (paragraph 1)
- B. "Later, when he was teaching at Tuskegee Institute, he put his plant skills to good use." (paragraph 2)
- C. "Carver knew that certain plants put nutrients back into the soil." (paragraph 4)
- D. "In 1921 the UPAA asked Carver to speak to the U.S. Congress about the many uses for peanuts." (paragraph 8)

Figure 14. A photograph was placed in between paragraph 5 and 6 in the Grade 4 Literacy test for PARCC assessment. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

### 4.3.2 Navigation Format

The two most common techniques of navigating or progressing through an electronic text are scrolling and pagination. For scrolling, users can use a mouse and scroll bar to move contents up or down a screen. For pagination, users move through a text in discrete pages or sections; this method is parallel to page-turning with paper texts. Scrolling can be applied horizontally or vertically across a variety of devices, including computers and tablets. Depending on length of texts and device types, scrolling and pagination are useful features that help readers navigate through a text efficiently. The way in which readers navigate through texts can vary based on medium or type of extracts.

A recent study investigated the interactions between scrolling and pagination with and without callouts; callouts contained quotes or conclusions from the main text and were

placed in the margins (Klyszejko et al., 2014). There were four main experimental conditions with a 2 (scrolling vs. pagination) by 2 (with callouts vs. without callouts) design. Participants (n=42; mean age = 34 years) were randomly assigned to one of the four groups and were requested to read text presented on the screen before completing a text recall test. Participants' eye movement during both reading and recall task was recorded. In terms of processing time, reading with scrolling facility was more demanding than reading a paginated text. In addition, recall test scores were significantly lower in the scrolling condition compared to the pagination condition (Klyszejko et al., 2014). Participants did, however, make more transitions between the main body of text and callouts in the scrolling condition than pagination condition. The authors conclude that callouts may be beneficial for reading comprehension when less efficient formats such as scrolling are used. Overall, these findings seem to support the use of pagination over scrolling, especially for complex and lengthy online text.

Kim and Kim (2013) examined the differences in reading performance (i.e. reading comprehension score) of teenagers (108 senior high school students; 16-17 year olds) on two mode formats (digital with scrolling vs paper mode) and indicated that navigation format could influence reading performance. They explained that scrolling presentation of text may have disrupted readers' structural representation of text, and negatively impacted students reading performance (i.e. reading comprehension scores) in passage based assessments. Although most of the students in the study were comfortable using scrolling to navigate through passages, the researchers argued that reading with scrolling may be cognitively more demanding than reading whole pages in a page-by-page (pagination) arrangement. This idea was supported by Piolat, Roussey and Thunin (1997)(1997)(1997), who also reported that scrolling text presentation interferes with students' ability to remember the physical structure of text compared to paging. In their study, students (54 university students) [age details not provided] were presented with passages in either scrolling or paging interface and asked to locate specific information. Results revealed relatively poor spatial memory for the passage that was read using the scrolled version. To address issues related to scrolling text interfaces, several other studies (Higgins, Russell, & Hoffmann, 2005; Pommerich, 2004) compared scrolling interface to a pagination interface for reading passages on computers and found that students generally performed worse

under the scrolling condition than the whole page condition – although findings were not statistically significant. In a more recent attempt, Kerr and Symons (2006) noticed that children’s (60 grade 5 students) reading rate (read more slowly) and reading comprehension (less efficient in comprehending information and higher order reading skills) were affected by the scrolling presentation of text on computer monitors.

Evidence from older studies suggests that there is no significant difference in reading performance between scrolling and paging (Mills & Weldon, 1987), however most researchers agree that lengthy text should be presented vertically (Dillon, 1992; Dillon et al., 1990). There may be an inclination for novice computer users to prefer paging, perhaps due to its close manipulation adherence to reading from paper (Schwartz, Beldie, & Pastoor, 1983). However, Dillon (1992) in his review noted that experienced users prefer scrolling to manipulate text possibly due to its speed – although there is evidence doubting its advantage over paging. The evidence surrounding the argument for the disadvantage of using a scrolling text interface is far from conclusive.

### 4.3.3 Reading tools

Reading on screen could be enhanced by including reading tools such as a highlighter. Some application software (e.g. Acrobat Reader) and web pages have already incorporated reading tools such as a highlighter and sticky notes (i.e., similar to Post-it notes) into their programs or webpage. The inclusion of reading tools in an online assessment is less common, and the advantages and disadvantages of including these tools are still being explored.

A few studies examining students’ use of reading tools in an online reading environment reported that annotation tools helped students processed information and enhanced students’ understanding of the online text (Lu & Deng, 2013; Nor, Azman, & Hamat, 2013; Nor, Hamat, Azman, Noor, & Bakar, 2011). In Nor et al.’s (2011) study, the authors explained that annotation tools were developed to help readers make notes when they read – this can range from writing short notes in bulleted points, to highlighting or underlining sections of online text. These notes help readers remember and understand what they have read. Eighty-one university students were recruited for this study. Students were given three online texts to read. Each text was accompanied by a questionnaire which was designed to

investigate students' perception towards the technological tools and their usage of the tools. They reported that the tools helped students to process information in a systematic way; for example, it allowed the reader to insert prompts and notes within the text. Similar findings were reported in the other two studies; both studies involved tertiary students.

The above findings relate to reading online materials. It is important to note that reading online materials and taking an online reading assessment are different in several ways despite their commonalities. For instance, online reading test often include time restrictions whereas reading online materials is usually considered a leisure activity. Given this, the usefulness of reading tools in an online assessment environment must be examined to determine whether such tools help students in test-taking. To date, most online assessments do not have reading tools. In the PARCC and Smarter Balanced assessments, however, tools such as a highlighter, flag or review tools and line readers were incorporated. Further discussion on reading tools included in currently available online assessment programs is presented in Section 5.4.

#### 4.4 Device Features

The use of digital devices for instruction and assessment (both formative and summative) purposes are continually growing with the advancement in technology. Students are increasingly interacting with their peers and teacher on academic content to demonstrate their expertise, ability, knowledge and skills. Online testing is regarded as an effective way for evaluating such competencies in students. The range of input and output devices used by students to access test contents delivered online is dependent on the availability of new technologies and the perceived utility of such technologies exercised across classrooms and schools. Moreover, at present, online tests encompass a wide variety of different innovative or technologically enhanced questions (TEQ) in addition to traditional multiple choice items (MCQ). Hence, there is a need for valid and reliable online test designs to ensure student test performance (test scores) is not negatively impacted when multiple devices are employed.

For assessment purposes, devices that are typically considered for testing or research studies include desktop computer with monitor, external keyboard and pointing device (such as mouse); portable laptop computers (including quick-start, hybrid and convertible

with touch functionalities) with keyboard (onscreen or external), monitor (relatively smaller screen than computers) and various pointing device options; portable tablet device (with even smaller screen than other devices), keyboard (onscreen or external), using touch/fingers as the pointing device and peripherals (e.g. styluses, stands, or external keyboards). A key element to note in terms of devices is the varied physical display size (monitor/screen size of device measured diagonally), resolution (clarity of the text and images displayed on screen) and input style (i.e. touch and keyboard functioning) across and within the same type of devices. Variation in these features can have major impact on how information is presented or accessed. For example, devices from laptop to tablets all have varying sizes of output screens from as small as 7-inch to as large as any number (perhaps not bigger than 32-inch for online testing which is what schools are most likely to have).

Professional educational and psychological testing standards indicate that neither mode of delivery nor the device used to access test content should influence students' performance and assessment outcomes (AERA, APA, & NCME, 2014). The increasing number of diverse technology in classrooms and schools poses greater difficulty in creating equivalence when assessing students under online conditions. Parshall, Spray, Kalohn and Davey (2002) described equivalence in terms of item characteristics, examinee scores, test constructs, and examinee behaviour between the two test versions that has been designed to be used as parallel version of the other as evidence to achieve fairness and validity. Some researchers (King, Kong, & Bleil, 2011; Sandene et al., 2005) suggest that student performance across desktop and laptop computers are relatively comparable (resultant student scores produced when tested on different devices) as they have similar form factors (i.e. the way in which student uses the device to access and manipulate digital content).

However, several studies have pointed out significant differences between tablets, computers and touch screen devices such as screen size, limit of information visibility, keyboard differences; these differences can influence students' test-taking experience and subsequently their test scores (Bridgeman, Lennon, & Jackenthal, 2003; King et al., 2011; Sandene et al., 2005). These researchers explained that differences in students' test scores between devices may be due to factors such as screen size (e.g., 7-inch versus 15-inch) as smaller screen size may limit on-screen information visibility; or input style and method (e.g., touch screen or external keyboard).

There is limited quantitative research evidence on students' interactions with tablets, however more qualitatively conducted research using cognitive labs and usability studies are available. A few researchers have concluded that there are no significant effects of device type (tablet vs. computer) for primary or high school students. Davis and her colleagues (2013) found no device effect (i.e. 10-inch iPad and 7-inch Google Nexus vs. computer) during their investigation using a "think-aloud" protocol (concurrent visualisation) for either 5th grade (16 males and 15 females) or high school (15 males and 17 females) students, in which students were asked to write short essays. These results are consistent with another study where K to 12 grade students were assessed on reading, writing and maths MCQ items (Olson, 2014). Similar results were reported in a more recent larger study with high school students (258 males and 224 females), which showed no observable performance differences in student test scores across devices for any content area (such as reading, maths and science) or item type (such as MCQ, TEIs, short type written response) (Davis, Kong, & McBride, 2015). However, some reading and essay items demonstrated somewhat significant differences in students' scores between devices used in the studies (Davis et al., 2015; Sandene et al., 2005).

Most of the device comparability studies included short surveys to collate some qualitative information about students' use and experience of different devices. Across most studies, there is a consistent pattern found in students' use of devices – most students had some form of prior experience, either in test taking or some other uses (both inside and outside classrooms), with computers or tablets. In terms of preference, most frequently endorsed options are paper and computer, with students indicating that they prefer to use these devices as they are either more familiar or have more experience with such devices.

Qualitative results (Davis et al., 2015; Davis et al., 2013; Strain-Seymour, Craft, Davis, & Elbom, 2013; Wang et al., 2008) suggest that despite the absence of evidence for statistically significant differences in reading across different devices, reading may be an area where students may prefer to work with tablet interfaces – where it is natural to scroll with fingers instead of manipulating with other intermediary devices (e.g., mouse, stylus). Although preference of device for testing is consistent across studies, further research is required with a focus on student reading interaction (with the physical or hardware aspects) and manipulation (with the interface and test content) on touch interface (such as tablets)

to observe student performance and preference outcomes, especially while working on longer passages/reading sections. In addition, most of the research considerations were directed around 10-inch tablets to conform to current practices (PARCC, 2015b; SBAC, 2014b) of allowable device sizes for large scale assessment programs, as smaller screen sizes may interfere with performance in high stake testing situations. As schools may have already employed smaller tablets (less than 10-inch e.g. 7-inch etc.) for classroom activities, future research may need to consider student performances across a variety of tablet and touch screen devices and test these devices for usability issues.

## 5 Recommendations for NAPLAN Online

---

The critical analysis of factors important for online assessment has revealed four major points. First, sans serif fonts such as Arial and Verdana are recommended over serif fonts (e.g., Times New Roman). Most studies have reported better performance in reading when online texts were presented in sans-serif fonts than serif fonts. It is important to note that some sans serif fonts (e.g., Verdana) were designed specifically for use on computers. As for font size, most studies of children and adults have indicated that 12- or 14-point size was good for onscreen reading; however this is dependent on readers' age and body of text. We recommend 12- to 14-point size fonts for Years 3 and 5, and 10- to 12-point fonts for Years 7 and 9.

Second, text layout (e.g. spacing and line length) of online assessment varied across studies. Given this, it is difficult to provide definite recommendations. However, findings from previous studies suggest that, for online reading assessments, 1.5- or double-spacing may be more suitable than single-spacing, depending on the type of text. In addition, for children, shorter line length ranging from 45 to 60 characters per line is also recommended. Single and double-column formatting is preferred; three-column or multiple column texts were read the slowest and are therefore not recommended.

Third, stimulus and items for online assessments were generally presented in a left-right or top-bottom format. That is, the stimulus was presented on the left and items on the right, and vice versa; or stimulus on top of page followed by items on the bottom. There was no evidence to support the superiority of either format, and majority of online assessment programs have used both formats. Paging and scrolling were common navigation techniques used in online reading, however, research evidence surrounding the advantages and disadvantages of each technique is not conclusive.

Lastly, the literature has highlighted several distinct differences between devices (such as desktop computers and tablets) that are commonly used for assessment purposes; some of the differences include display size and input style. Variation in device features may impact how information is presented and accessed on screen which can in turn influence students'

user experience and performance. More research is still needed to examine the usability of different devices for online assessment of reading.

## 5.1 Display Format

Through the review of example NAPLAN paper-tests, three main types of passages or texts are identified:

- i. Plain texts, with no or few add-on graphics
- ii. Text with supplementing graphics that aid reading comprehension, such as photos or figures about food chain or life cycle
- iii. Compare and contrast texts which are presented side by side

The recommendations and display format options for NAPLAN online will be presented based on these categories. Methods of presenting multiple-media content (e.g., a passage and a video clip or audio file) in reading assessments are also reviewed. Considerations for different age groups and device types will be discussed in each respective section/recommendation.

### 5.1.1 Category 1: Text only passages, some with add-on graphics

A vast variety of reading content do not require any graphics to support the text. However, it is evident from different form of texts (such as narrative, expository, document, argumentative etc.) that graphics are often added to make the content more appealing to the readers or to serve as decorative visual cues. The goal of including such graphics is to engage and motivate readers with the reading task, and to make subject matter attractive and exciting in general. Below are options for presenting text only passages online:

#### ***Option 1 – Direct transfer***

Paper-based passages with add-on graphics can be directly transferred into an online environment. With this approach, the ideas of paper representation is directly imitated online – whether it be with or without background images or other form of additional graphics placed at different parts of the text. It is important to note that text presented on screen should not lie directly on images or have an image as the screen background as this may affect the readability of the text displayed. A challenge with this option is that, if paper-

based passages containing large images in the beginning of the text is transferred directly online, then there is a chance that some of the text would be hidden (page content truncated) on an online page and students would need to navigate more (page or scroll through) to view the reading contents. Moreover, direct transfer may limit viewing ability across devices and graphic colours may be difficult to convert from one form to another. This method is likely to impact reading ability when two or more passages need to be published in the same space together.

### ***Option 2 – Enhanced translation***

A key factor to consider when designing items (e.g. stimulus, question, and option) for online environment is the limited screen space across different devices. Complementary graphics should be utilised and placed in the correct location so that it encourages readers' engagement rather than hinder their reading flow. For this reason, graphics format such as size, colour, contrast and resolution may need to be modified from how they appear on paper format when presented online. Graphics (plain add-ons or visual aids) should always convey the passage content/information and be placed in context. Web design guidelines should be followed for effective use of background or other graphics. Careful considerations need to be given to the types of devices used for administering online tests, and the impact that image quality or location may have on young children. Moreover, graphics format and location is also dependent on navigational layout (e.g. scrolling or paging).

#### **5.1.2 Category 2: Text with graphics that aids reading comprehension**

Passages are often presented with supplementing graphics or figures with meaningful purposes (e.g. figure illustrating life cycle of an animal or food chain) to aid in reading of different reading materials such as posters, ads, and flyers. In this context, the images or graphics act as visual cues to support delivery of important information or even highlight meaningful messages from texts. The following options illustrate how such graphics or multimedia elements can be presented in an online environment.

### ***Option 1 – Direct transfer***

To imitate reading on paper, texts displayed online will incorporate graphics, figures and diagrams in the same order and location as they would appear in the paper version of the test. Similar to paper versions, online graphics can be positioned at either the top or bottom

of text, in-line (e.g. square, tight, through) with text, and before or after (e.g. left, right, centred) a section or paragraph of the text where graphics are deemed necessary to illustrate ideas or details from the passage. Passages with graphics can be displayed in a single column or two-column text format, using either the top-down or left-right presentation format. Many studies have utilised this method to present graphical elements in a reading assessment; this may be due to its similarity to paper-based reading assessment and to students' familiarity with this layout format. This display styles can be applied to either types of devices (i.e., computer and tablets) and for students across all age groups.

### ***Option 2 – Graphics as hyperlinks***

When displaying texts and graphics online, graphics can be referenced with hyperlinks (Figure 15). Students can view the images by clicking on the hyperlinks. Using hyperlinks, graphics can be placed anywhere in the passage. However, it is preferable that graphics are presented at the bottom of the page so that more text can be displayed in the screen view; this would reduce obstruction to the natural flow of reading. One advantage of this option is that the passage will not be covered or block by the image but will instead be displayed at the bottom of the page. This option is similar to what is found in books, where graphics are usually presented on the next page or in the top or bottom half of a page, rather than in between paragraphs. Alternately, diagrams or other related graphics can be presented in the text (i.e. in the paragraph for which it is relevant) as hyperlink, however, graphics appear in a pop-up window when the hyperlink is clicked (Figure 16). It should to be noted that if graphics are represented as hyperlinks, the hyperlinks should stand out or look different to the normal reading text. For example, the hyperlinks could appear in a different font, colour, size and etc.

This option can be applied to various devices without any issues. However, further testing is needed to investigate how students of different ages (e.g. primary vs. secondary school students) will engage with the task or whether students' experience and familiarity would impact performance.



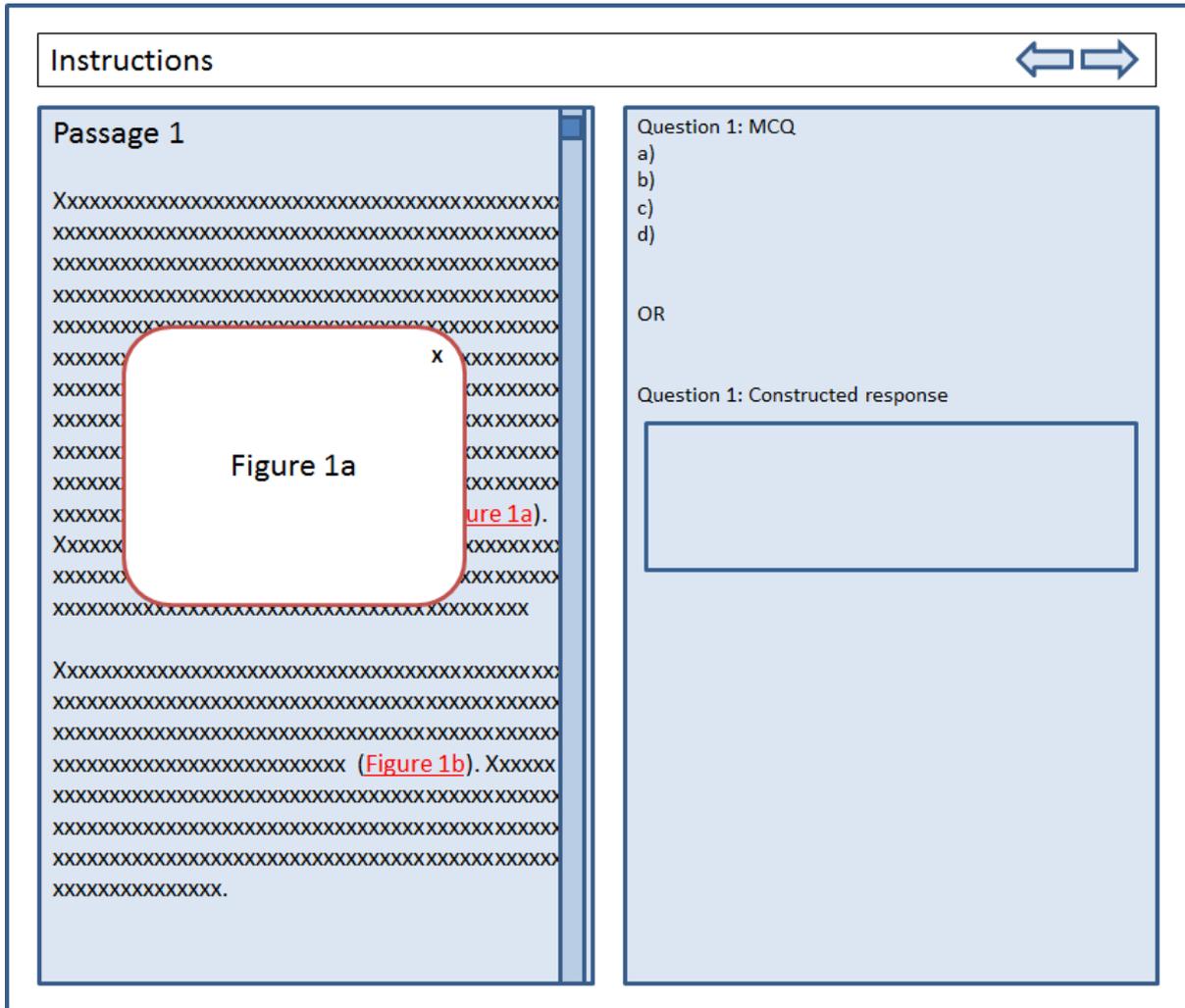


Figure 16. Graphics and figures are presented as hyperlink in text. When clicked, the graphic or image will pop-up and displayed over the text. Readers can close the pop-up image by clicking on the 'x' on the top right hand corner of the pop-up window.

**Option 3 - Graphics as thumbnails**

A common way of rendering images on webpages is to show thumbnail preview of the images. This is a common practice when designing website as it helps reduce page loading time when dealing with multimedia contents (such as graphics, audio etc.). Reading passages can also benefit from such ideas of displaying images in an online environment. This convention is perhaps a better extension of the previous option (discussed in option 2). The images related to the text will appear in the same way as they might if presented on paper – with the difference that they would be a miniature form of the actual image (Figure 17). This will give students an idea of what the actual image may look like and they are still

able to view the image in its actual context – just in a scaled down version. Once clicked, the full version of the graphic pops up with enhanced view (Figure 18). When designing reading assessments for different devices, quality of graphics will not need to be compromised because of available space – specifically the ones with smaller screen (e.g. tablet, netbooks). These characteristics can be applied to various devices and browser types and will particularly be more useful for small screen devices. Students are less likely to be affected as the view of graphics is always visible on screen. There may need to be some indication (such as using a magnifier or plus sign) that the thumbnails can be clicked to get enlarged views.

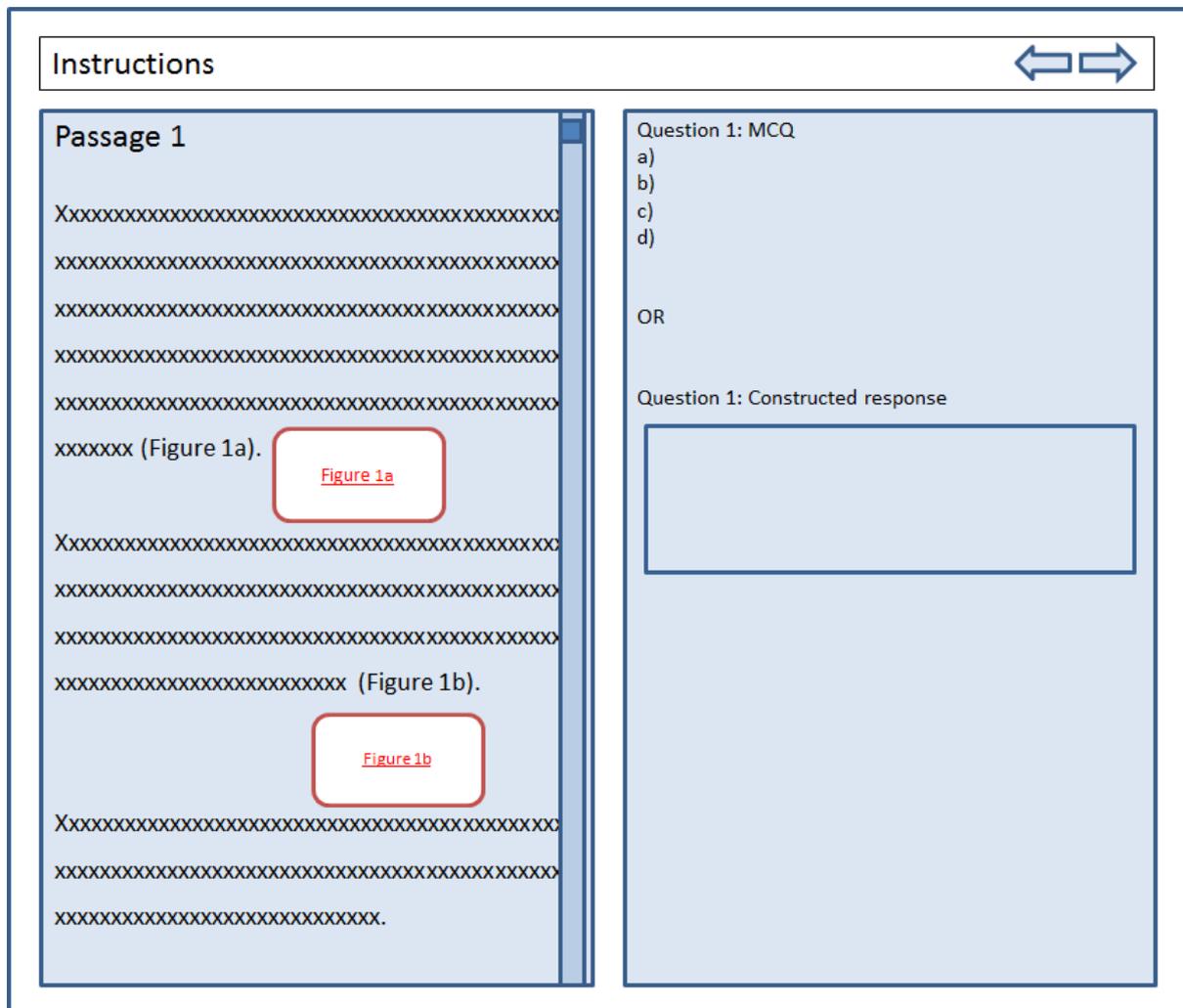


Figure 17. Graphics and figures are presented in text as miniature thumbnails.

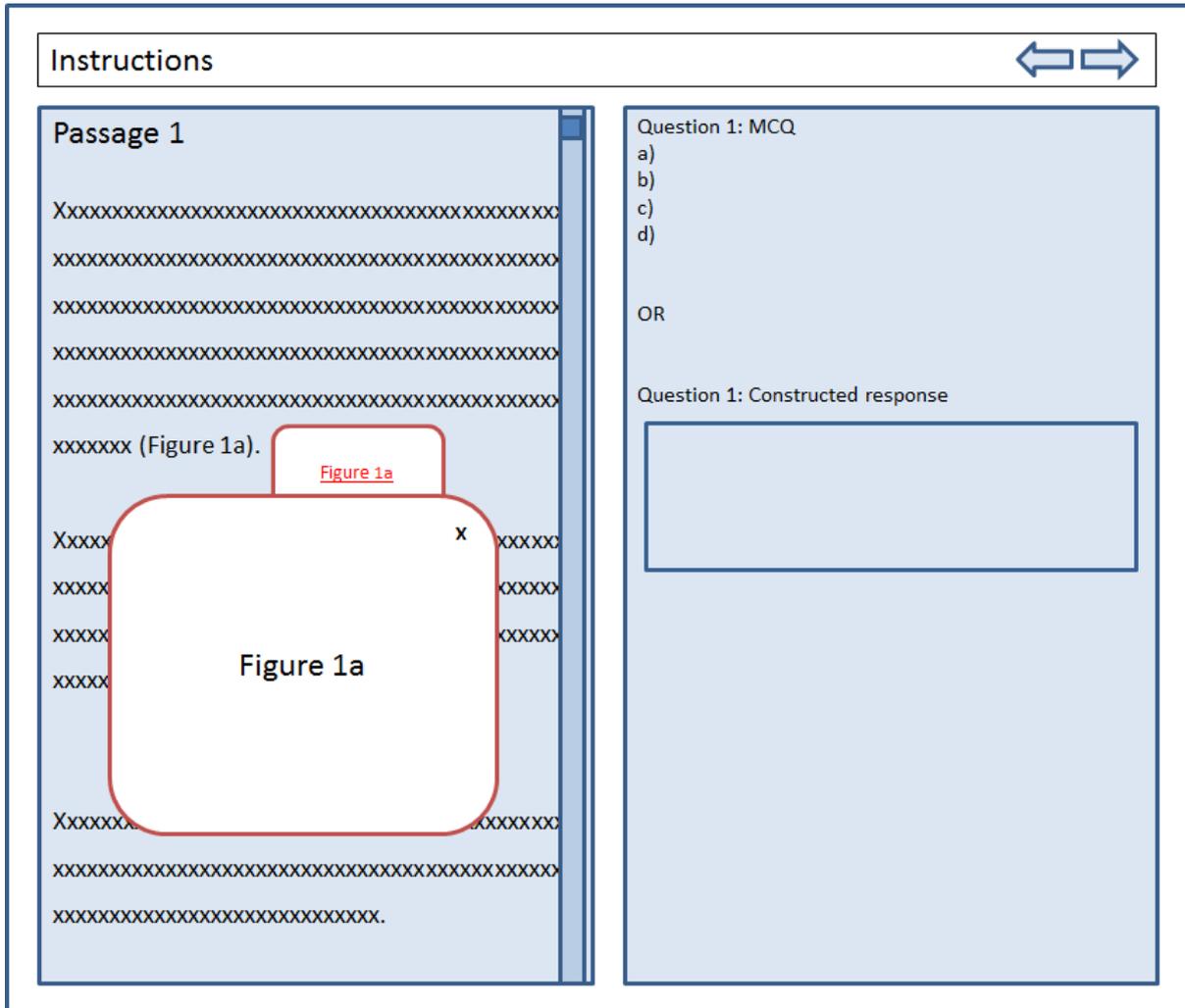


Figure 18. Graphics and figures are presented in text as miniature thumbnails. When a thumbnail is clicked, an enlarged version of the graphic or image will be displayed over the text. Readers can close the enlarged image by clicking on the 'x' on the top right hand corner of the pop-up window.

### 5.1.3 Category 3: Compare and contrast texts and passages

The following options are recommended for presenting 'compare and contrast' texts and passages.

#### **Option 1 - Plain text transfer**

'Compare and contrast' passages are taken from paper-based formats and applied directly to a computer screen. To simulate reading on paper, passages can be presented side-by-side, with questions presented below the passages (Figure 19). Alternatively, passages can

be presented vertically one after another on a scrolled page, with questions presented on the right side of the page (Figure 20). This option is feasible on both computers and tablets. Given that tablets have smaller screen sizes compared to computers, passages and questions may appear in smaller. However, students may use the zoom feature available on most tablet devices to enlarge or magnify the text. Further discussion about features and tools of each device type is presented in section 5.4.

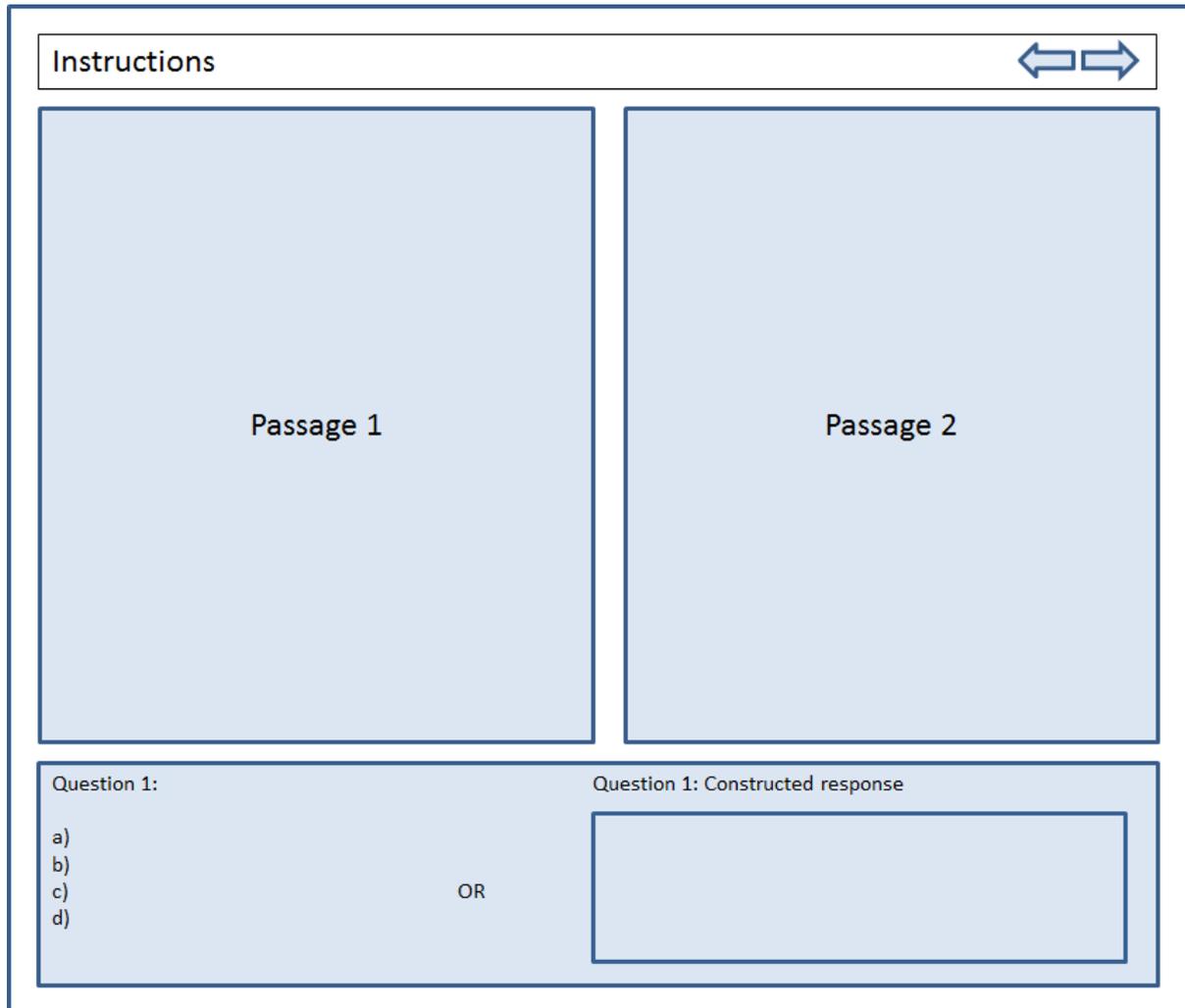


Figure 19. Plain transfer of ‘compare-and-contrast’ paper-based text to computer using a top-bottom display format.

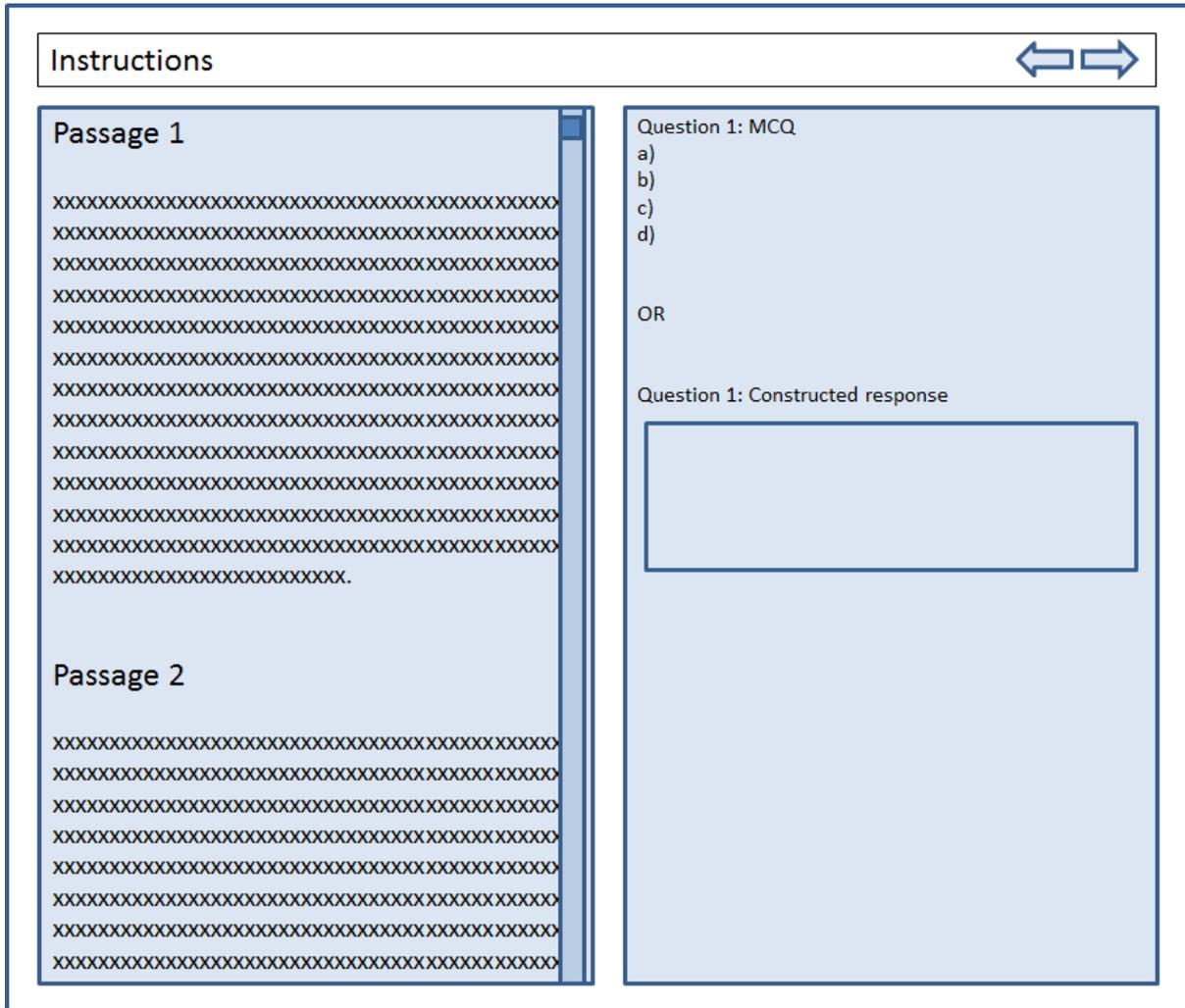


Figure 20. Plain transfer of ‘compare-and-contrast’ paper-based text to computer using a left-right display format.

**Option 2 – Page tabs**

Instead of presenting two texts side by side or vertically on a scrolled page, page tabs can be used to show, hide or customise presentation of multiple texts on a page. This method was employed by the PARCC assessment (Figure 21). In the PARCC assessment, passages were presented on a scroll page on the left using page tabs, while questions were presented on the right side of the screen. It is important to note that page scrolling was only applied to the passage section and not the whole page. Students were able to switch between the passages by simply clicking on the corresponding tab. Questions on the right were always visible irrespective of which text was being read. The PARCC assessment used a left-right

presentation format using tabs to display more than one passage. Page tabs can also be used to present items in a top-bottom format (Figure 22). In this way, passages are presented on the top of a page, with tabs appearing along the right side of a page. Questions will be presented below the passages. A scroll page can be used to accommodate or include more questions. The page tab option would allow for multiple texts to be presented (i.e., not limited to two texts only) in either left-right or top-bottom formats. Although this option can be applied to a computer or tablet, younger children (e.g., Year 3 and 5 students) may not be familiar with the page tab feature. Therefore, this needs to be first tested with children across all year groups.

GRADE 9 ELA/LITERACY / SECTION 1 / 6 OF 23

Today you will read and analyze a short story and a passage from another short story. As you analyze these texts, you will gather information and answer questions about each text and its relationship to the other so that you can craft a written response.

Departure from "Up the Coolly"

Read the story "Departure," about a young man leaving home, by United States writer Sherwood Anderson (1876–1941). Then answer the questions.

Departure  
by Sherwood Anderson

1 Young George Willard got out of bed at four in the morning. It was April and the young tree leaves were just coming out of their buds. The trees along the residence streets in Winesburg are maple and the seeds are winged. When the wind blows they whirl crazily about, filling the air and making a carpet underfoot.

2 George came downstairs into the hotel office carrying a brown leather bag. His trunk was packed for departure. Since two o'clock he had been awake thinking of the journey he was about to take and wondering what he would find at the end of his journey. The boy who

**Part A**

Which is a theme reflected in **both** the passage from "Departure" and the passage from "Up the Coolly"?

- A. Nature can provide a peaceful place for reflection.
- B. Sometimes it is comforting to be left alone.
- C. There is often a vast difference between memory and reality.
- D. One's upbringing can greatly impact relationships with others.

**Part B**

Choose **two** quotations, **one** from **each** passage, that **best** support the answer in Part A.

- A. "On the April morning he wanted to go there again, to walk again in the silence." (from "Departure")
- B. "When the train started Tom Little punched his ticket, grinned and, although he knew George well and knew on what adventure he was just setting out, made no comment." (from "Departure")
- C. "The young man, going out of his town to meet the adventure of life, began to think but he did not think of anything very big or dramatic." (from "Departure")
- D. "It called back to Howard the days when he and Grant, his younger brother, had fished in this little brook for trout . . . ." (from "Up the Coolly")

Figure 21. Grade 9 English language arts/Literacy practice tests from PARCC. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

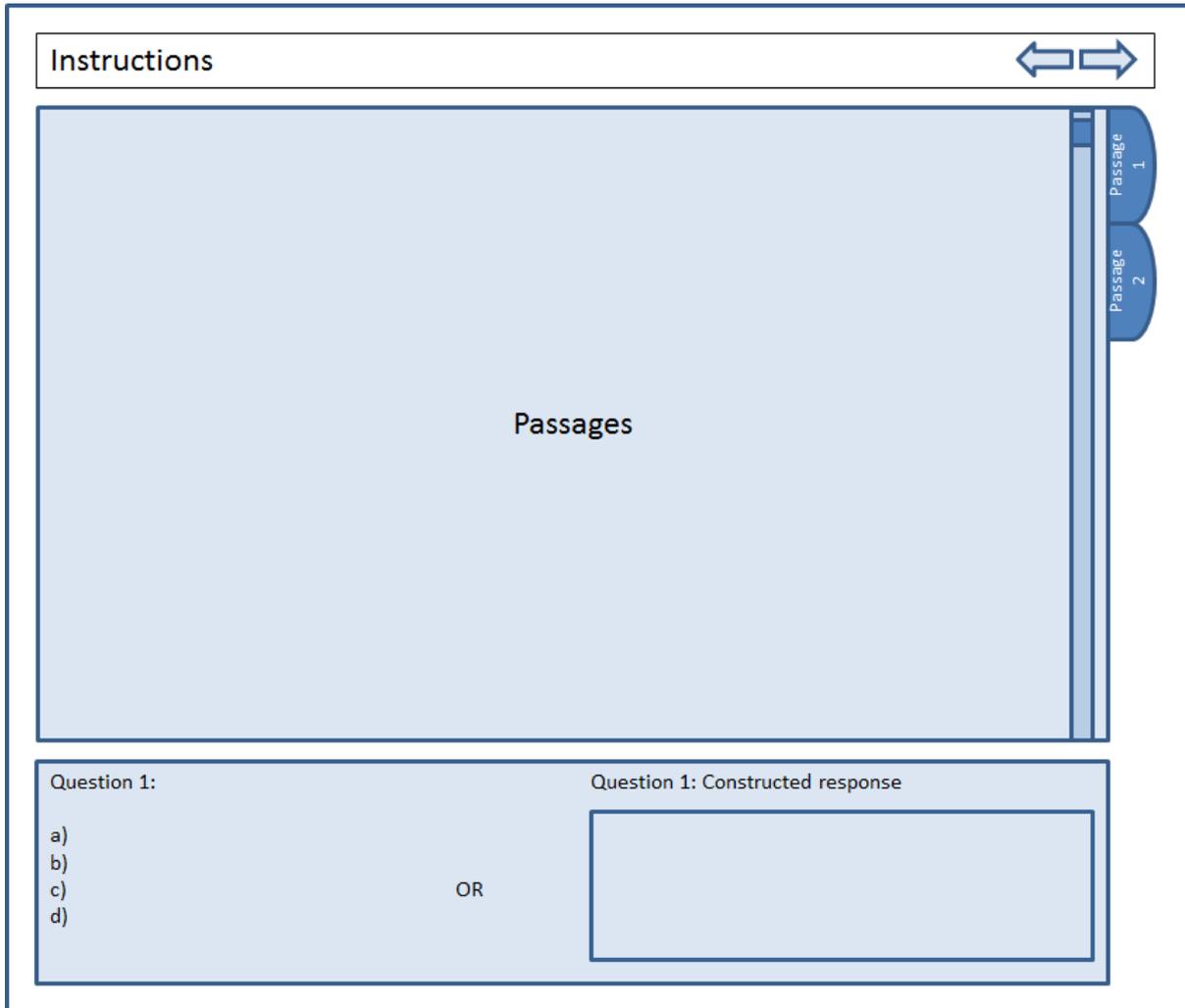
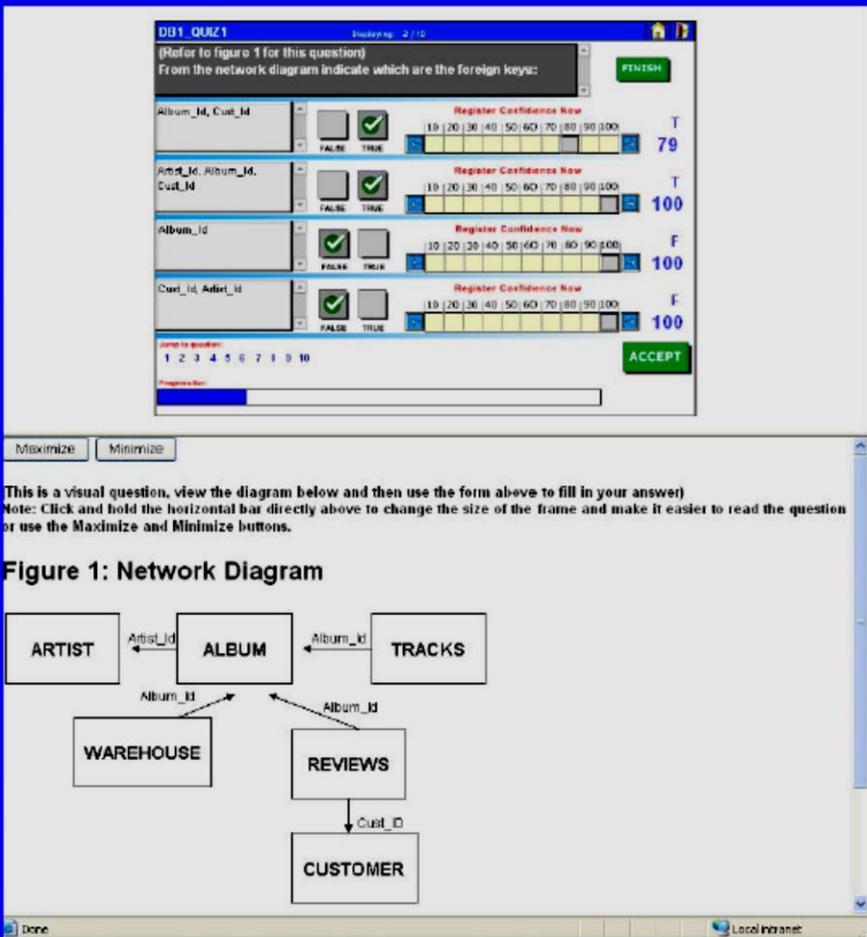


Figure 22. Multiple passages presented using page tabs displayed in a top-bottom format.

**Option 3 – Bifocal display**

The bifocal display technique was discussed in a study by Farrell and Farrell (2011) where questions were presented in a readable text, and supporting graphics/diagrams were presented in a compressed/distorted format which could be maximised or minimised (see Figure 23). This technique can be adapted to display ‘compare and contrast’ passages for NAPLAN online reading assessments. Using the top-bottom (Figure 24) or left-right format (Figure 25), passages can be displayed using the bifocal display technique. Students can switch or toggle between passages to determine which passage is in focus, while the other passage(s) is still displayed on screen. This option can be applied to computers and tablets; however the user interface for this display technique will differ slightly across device (e.g., using mouse to toggle between passages vs. touch screen swiping).



**DB1 QUIZ1**  
 (Refer to figure 1 for this question)  
 From the network diagram indicate which are the foreign keys:

Register Confidence Now

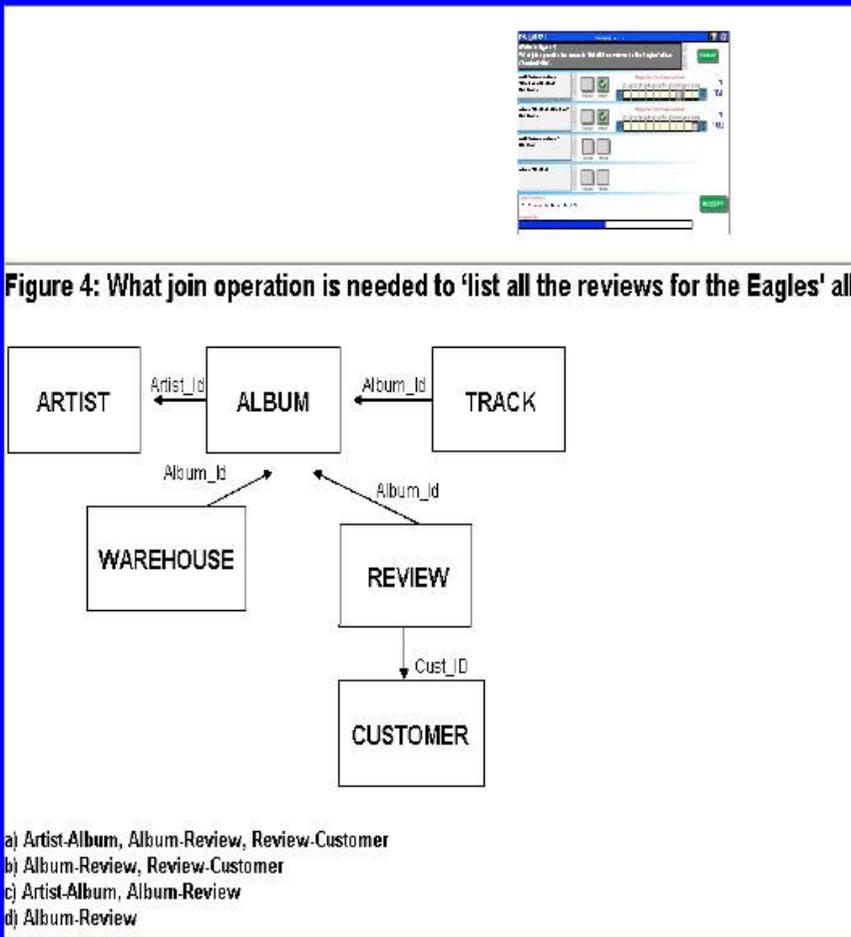
Album_Id, Cust_Id	<input type="checkbox"/> FALSE	<input checked="" type="checkbox"/> TRUE	10 (20 30 40 50 60 70 80 90 100)	79
Arti_Id, Album_Id, Cust_Id	<input type="checkbox"/> FALSE	<input checked="" type="checkbox"/> TRUE	10 (20 30 40 50 60 70 80 90 100)	100
Album_Id	<input checked="" type="checkbox"/> FALSE	<input type="checkbox"/> TRUE	10 (20 30 40 50 60 70 80 90 100)	100
Cust_Id, Artist_Id	<input checked="" type="checkbox"/> FALSE	<input type="checkbox"/> TRUE	10 (20 30 40 50 60 70 80 90 100)	100

Accept to Question: 1 2 3 4 5 6 7 8 9 10

**Figure 1: Network Diagram**

```

        graph TD
            ARTIST -- Artist_Id --> ALBUM
            TRACKS -- Album_Id --> ALBUM
            WAREHOUSE -- Album_Id --> ALBUM
            ALBUM -- Album_Id --> REVIEWS
            REVIEWS -- Cust_Id --> CUSTOMER
            
```



**Figure 4: What join operation is needed to 'list all the reviews for the Eagles' album**

```

        graph TD
            ARTIST -- Artist_Id --> ALBUM
            TRACK -- Album_Id --> ALBUM
            WAREHOUSE -- Album_Id --> ALBUM
            ALBUM -- Album_Id --> REVIEW
            REVIEW -- Cust_Id --> CUSTOMER
            
```

a) Artist-Album, Album-Review, Review-Customer  
 b) Album-Review, Review-Customer  
 c) Artist-Album, Album-Review  
 d) Album-Review

Figure 23. A question displayed using the bifocal technique, where a diagram can be maximised or minimised and the question remains visible on screen although out of focus. Image taken from article by Farrell and Farelle (2011).

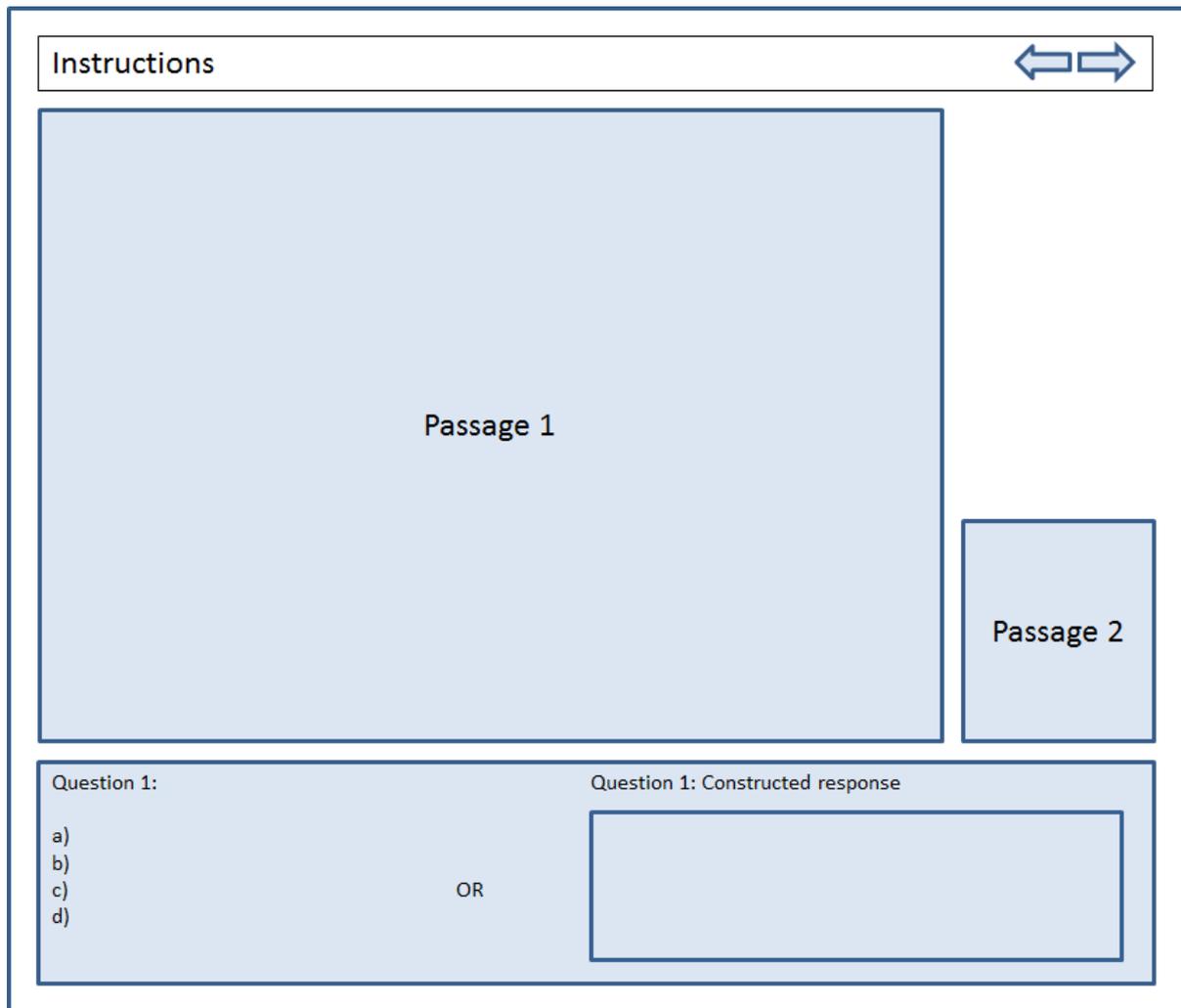


Figure 24. Two passages displayed on a screen using the bifocal display technique. In this example, passages and questions are presented in a top-bottom format. Passage 1 is in focus while passage 2 is still in view but in the background/not in focus.

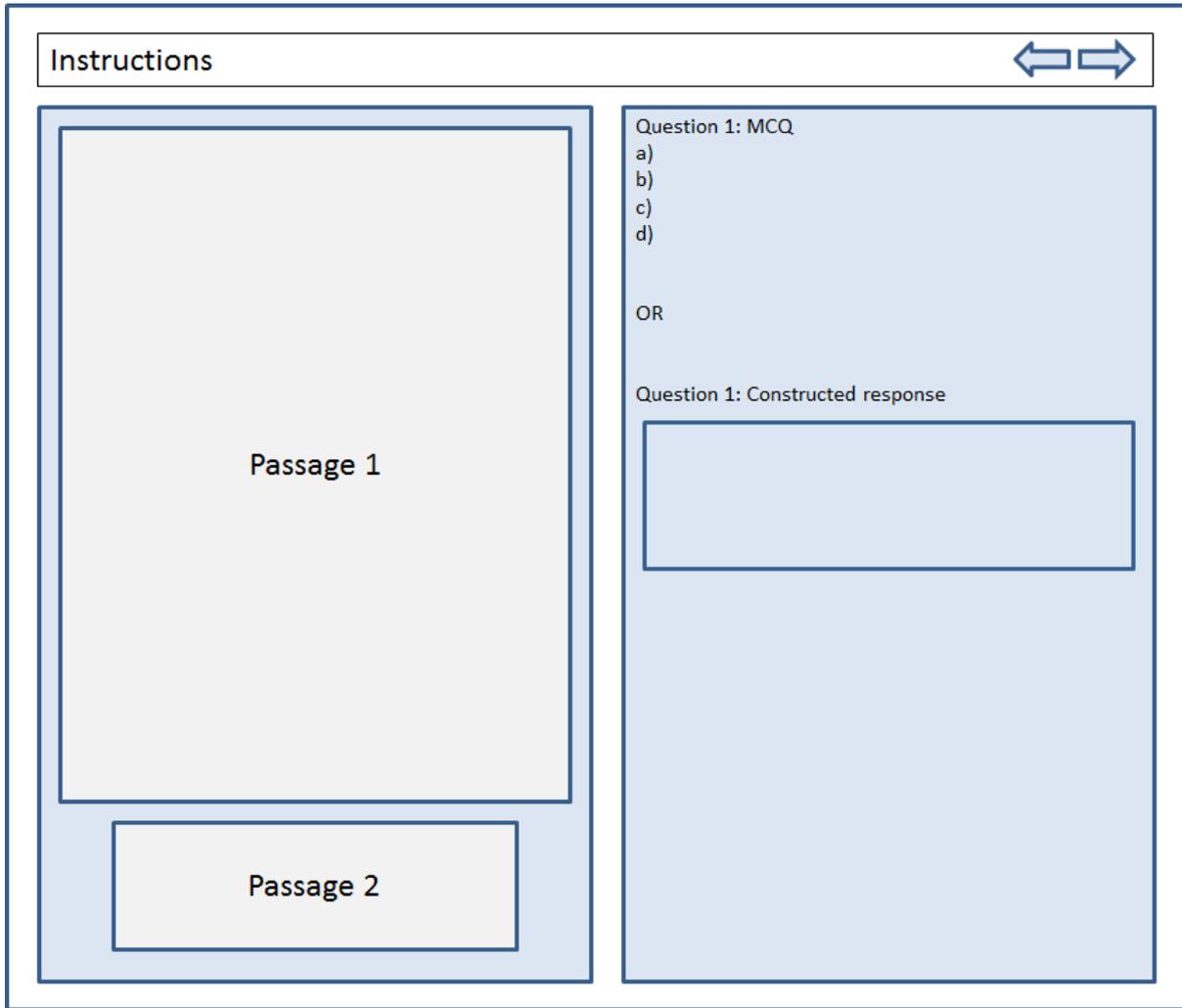


Figure 25. Two passages displayed on a screen using the bifocal display technique. In this example, passages and questions are presented in a left-right format. Passage 1 is in focus while passage 2 is still in view but in the background/not in focus.

#### 5.1.4 Other media

With the advancement of technology, multimedia contents can now be included as part of reading assessments. In a computer-based reading task, students may be presented with an audio recording and/or video footage with or without a text passage. Students would be expected to answer questions related to the information and content presented in those media.

Video clips were included as part of an English Language Arts/Literacy assessment in the Smarter Balanced and PARCC program (see Figures 26 and 27). In these tasks, the video was

presented on the left side of the screen, and the question (either MCQ or constructed response) was presented on the right. Students had to click play and watch the video in order to answer the corresponding question. Literacy tasks that incorporate other forms of media content (e.g., video clips) are assessing more than basic reading skills. Students need to have some level of information and communications technology (ICT) skills to successfully navigate through the online page and answer corresponding questions. Skills and content related to ICT is not within the scope of this review, and therefore will not be discussed in details.

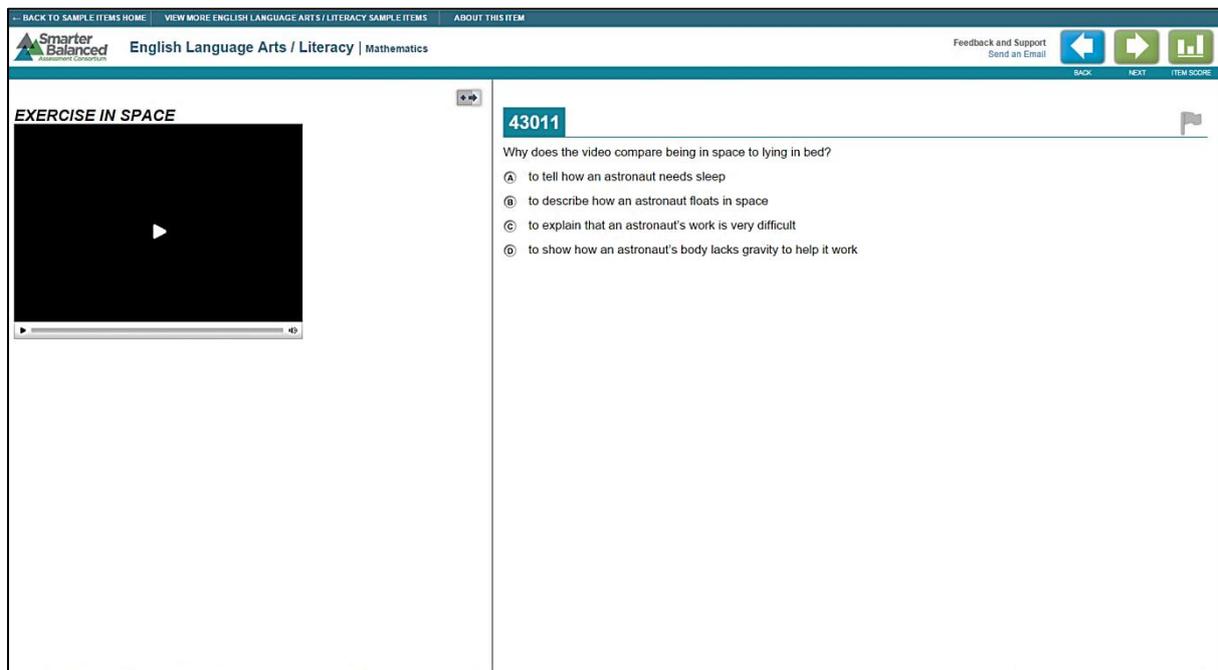


Figure 26. A sample test from the Smarter Balanced which included a video footage. No text passage was included. Screen shot taken from

<http://sampleitems.smarterbalanced.org/itempreview/sbac/ELA.htm>

The screenshot shows a web-based assessment interface. At the top, there is a navigation bar with a 'Review' button, a 'Flag' button, and a 'Guest' user profile. Below the navigation bar, the page title reads 'GRADE 7 ELA/LITERACY / SESSION 1 / 11 OF 23'. The main content area is divided into two columns. The left column contains a text passage and a video player. The right column contains two sets of multiple-choice questions.

**Text Passage:**

Today you will research electricity and consider some of the methods used in science texts and videos to support different purposes. First, you will read an article that explains some general principles of electricity. Next, you will watch a video about fun ways to learn about electricity circuits. Finally, you will read an article that explains how different materials conduct electricity. As you review these sources, think about the purpose of each and the role that explanations, demonstrations, and/or descriptions of experiments play in communicating that purpose. At the end of the task, you will be asked to write an essay.

**Video Player:**

Watch the video *Hands-On Science with Squishy Circuits*. Then answer the questions.

Hands-On Science with Squishy Circuits  
by AnnMarie Thomas

The video player shows a black screen with a white play button in the center.

**Part A:**

In segment 1:14–1:17 of *Hands-On Science with Squishy Circuits*, what does the word **commercial** mean?

- A. used at home
- B. created with others
- C. produced for sale
- D. made widely known

**Part B:**

Which detail from *Hands-On Science with Squishy Circuits* **best** supports the answer in Part A?

- A. "... we spent a summer looking at different play-dough recipes . . . " (0:39–0:42)
- B. "... these recipes probably look really familiar . . . " (0:43–0:45)
- C. "... pretty standard ingredients you probably have in your kitchen . . ." (0:47–0:49)
- D. "... regular play-dough that you buy at the store . . ." (1:07–1:09)

At the bottom of the video player, there is a copyright notice: "Hands-On Science With Squishy Circuits" by AnnMarie Thomas. Copyright © 2011 by TED Conferences, LLC. Published by TED

Figure 27. A sample test from the PARCC assessment which included a video footage. A short instructional text passage was presented above the video clip. Screen shot taken from <http://epat-parcc.testnav.com/client/index.html#getitem/7908>

The examples illustrated in Figures 26 and 27 depend on only one type of media content – video footage. In some cases such as in the PARCC assessment for Grade 7 students, multiple media contents (i.e., text passages and video/audio clips) were used. Students were required to read multiple articles and view a video clip about a common topic, and then compare and contrast information from the multiple sources.

In the sample PARCC test presented below, the text passages and video clip were presented as page tabs. The first page tab contained a text passage titled “Energy Story” (Figure 28). The second page tab contained a video clip titled “Hands-on Science with Squishy Circuits” (Figure 29). The third page tab contained a second text passage titled “Conducting Solutions” (Figure 30). Across all three page tabs, text passages and the video clip were

presented on a scroll section on the left, while a constructed long answer item (i.e., essay) was presented on the right. This display format using page tabs would allow for multiple contents to be displayed, including multimedia contents such as video clips or audio recordings.

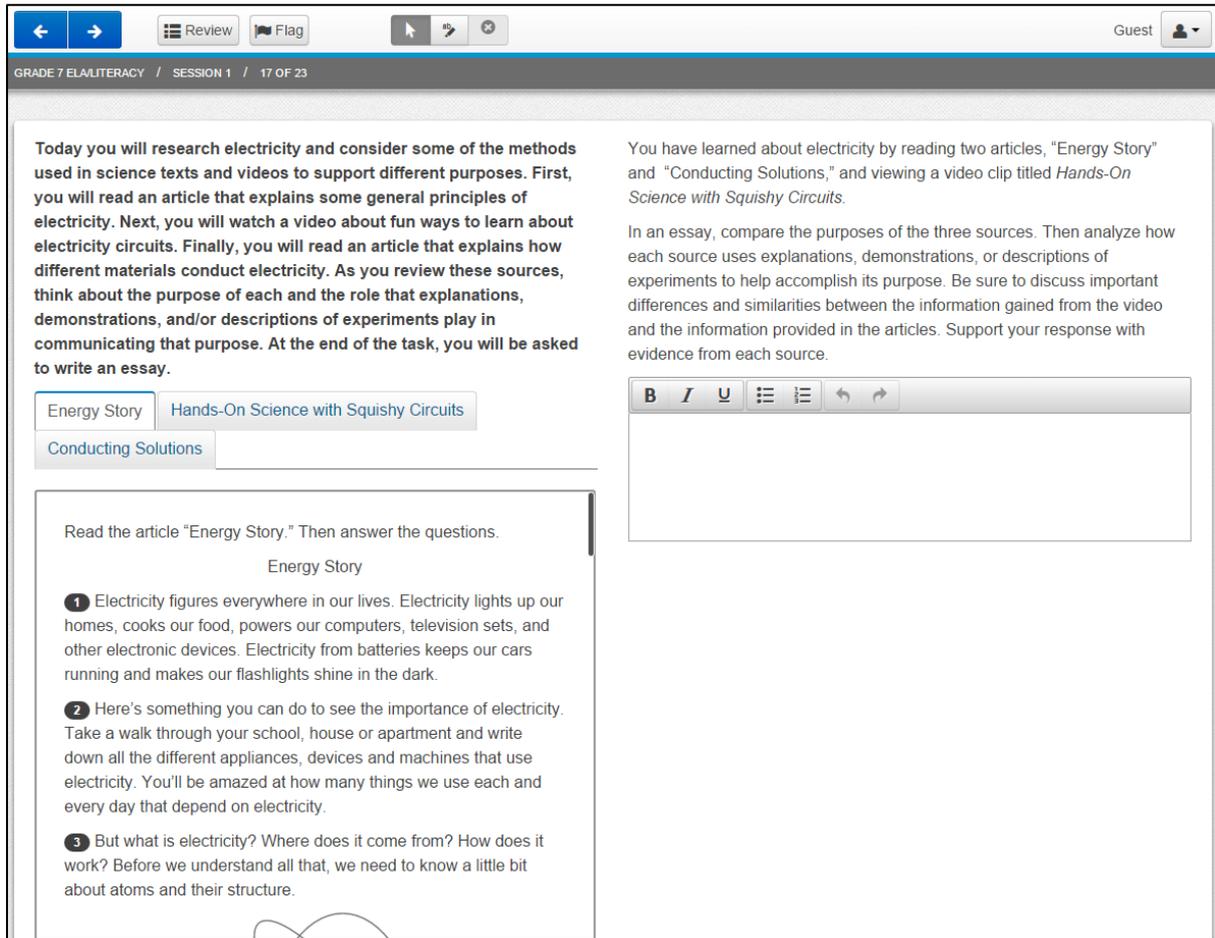


Figure 28. A sample test from the PARCC assessment. The first page tab contained a text passage titled “Energy Story”. Screen shot taken from <http://epa-parcc.testnav.com/client/index.html#getitem/7914>

The screenshot shows a web-based assessment interface. At the top, there is a navigation bar with a 'Review' button, a 'Flag' button, and a 'Guest' user indicator. Below this, a breadcrumb trail reads 'GRADE 7 ELA/LITERACY / SESSION 1 / 17 OF 23'. The main content area is divided into two columns. The left column contains a bolded paragraph of text: 'Today you will research electricity and consider some of the methods used in science texts and videos to support different purposes. First, you will read an article that explains some general principles of electricity. Next, you will watch a video about fun ways to learn about electricity circuits. Finally, you will read an article that explains how different materials conduct electricity. As you review these sources, think about the purpose of each and the role that explanations, demonstrations, and/or descriptions of experiments play in communicating that purpose. At the end of the task, you will be asked to write an essay.' Below this text are three tabs: 'Energy Story', 'Hands-On Science with Squishy Circuits' (which is selected), and 'Conducting Solutions'. The right column contains a paragraph: 'You have learned about electricity by reading two articles, "Energy Story" and "Conducting Solutions," and viewing a video clip titled *Hands-On Science with Squishy Circuits*. In an essay, compare the purposes of the three sources. Then analyze how each source uses explanations, demonstrations, or descriptions of experiments to help accomplish its purpose. Be sure to discuss important differences and similarities between the information gained from the video and the information provided in the articles. Support your response with evidence from each source.' Below the text is a large text input area with a toolbar containing buttons for bold (B), italic (I), underline (U), bulleted list, numbered list, and undo/redo arrows. The video player section on the left shows the title 'Hands-On Science with Squishy Circuits' by AnnMarie Thomas and a video player with a play button.

Figure 29. A sample test from the PARCC assessment. The second page tab contained a video clip titled “Hands-on Science with Squishy Circuits”. Screen shot taken from <http://epat-parcc.testnav.com/client/index.html#getitem/7914>

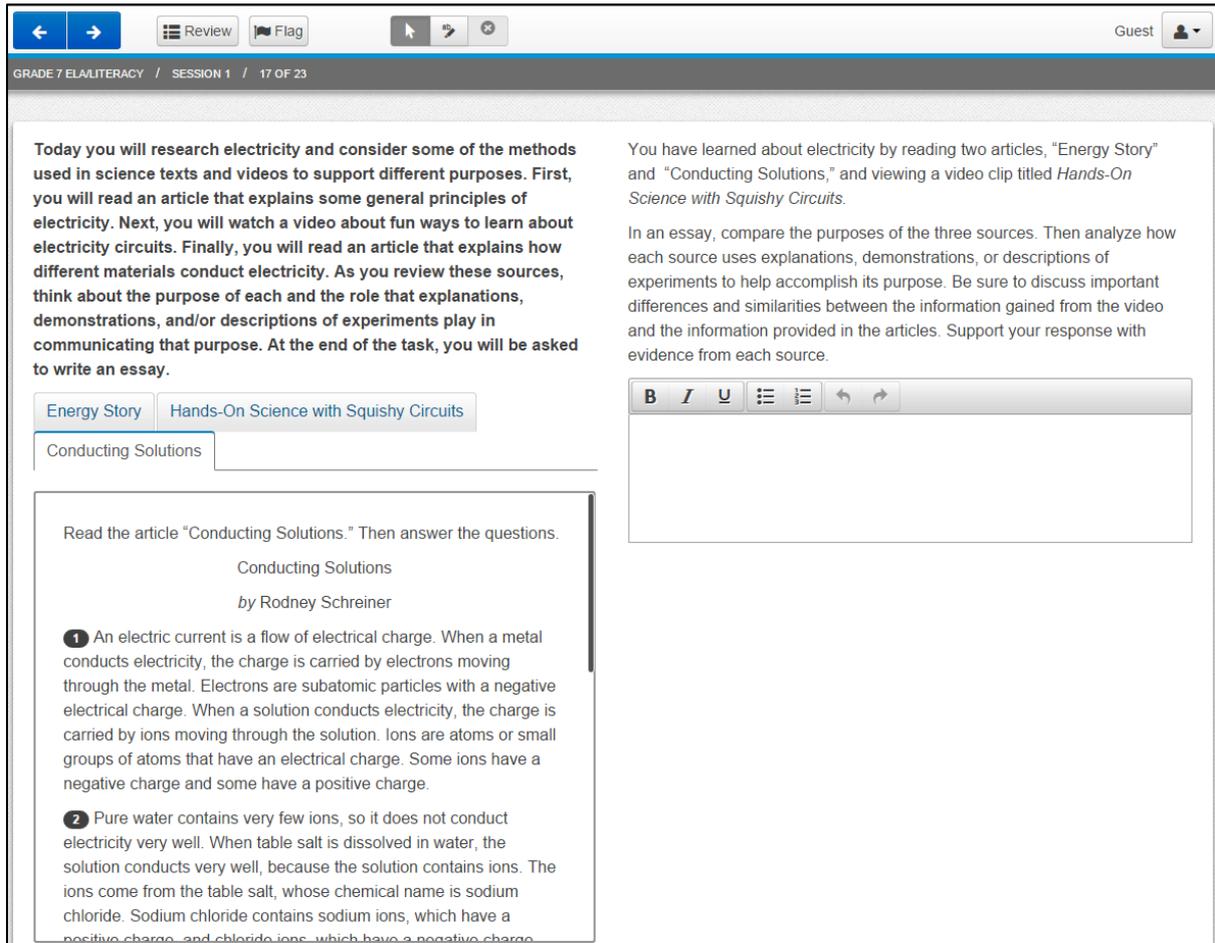


Figure 30. A sample test from the PARCC assessment. The third page tab contained a text passage titled “Conducting Solutions”. <http://epat-parcc.testnav.com/client/index.html#getitem/7914>

Alternatively, multiple media contents can be displayed on a page as hyperlinks instead of page tabs. In one of the Assessment Research Centre’s ICT creative problem solving assessment, different types of task content and information (i.e., video collection, poem text) were presented as hyperlinks on the web page (see Figure 31). Students were able to access the contents by clicking on the hyperlinks. Contents appeared on a separate pop-up window. The window could not be re-sized or minimised, and could only be closed. To our knowledge, this display format has not been used for compare-and-contrast reading task comprising of multiple media contents (e.g., text passages and video/audio files). A possible disadvantage of this method for reading assessment is that questions and response options may get covered or blocked by the pop-up window. If so, students would need to open and

close the content window in order to view the items again. The use of this display format for compare-and-contrast reading task using multiple-media would need to be tested with school children, especially with younger students, to determine the ease of use and to examine potential impact on students' performance.

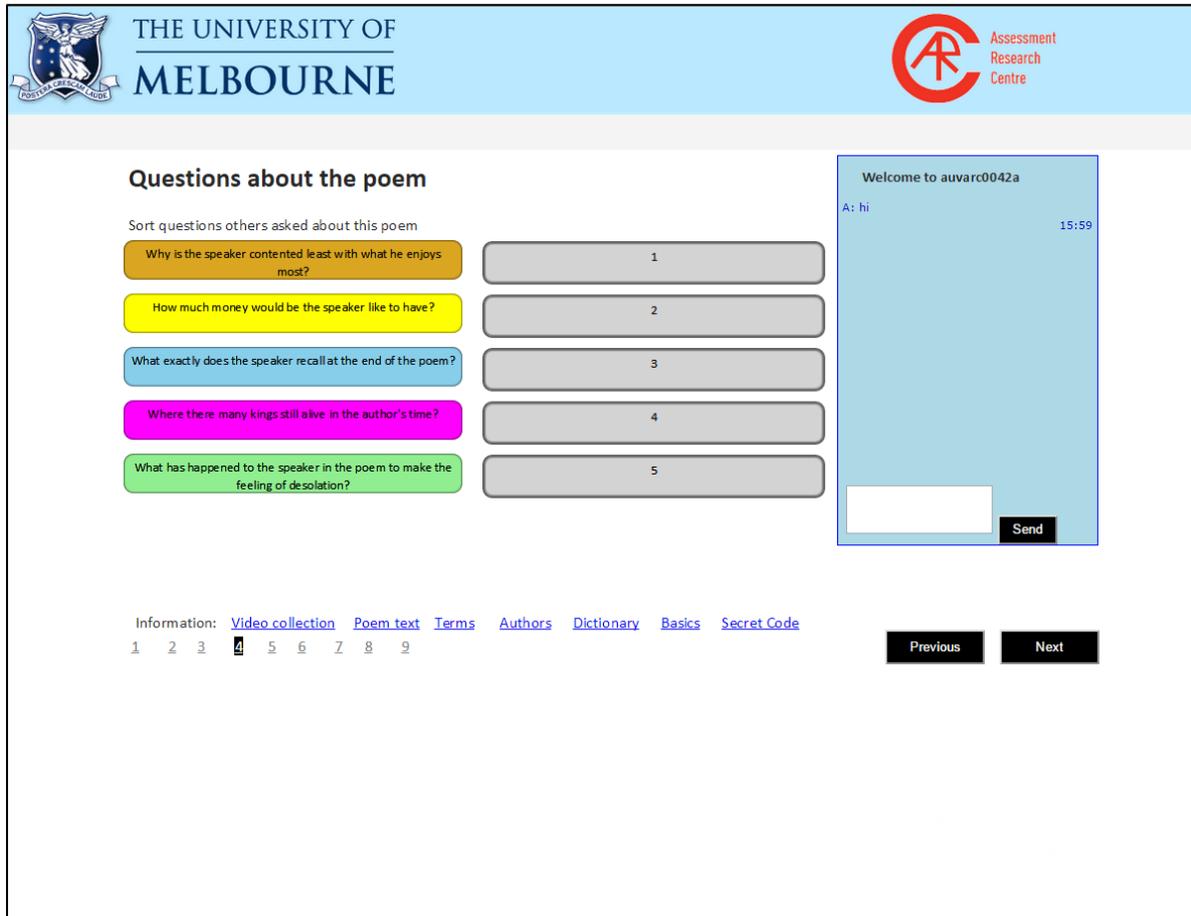


Figure 31. A screen shot from an ICT collaborative problem solving task designed by the Assessment Research Centre. Students had the option of accessing information contents by clicking the hyperlinks (in blue) displayed below the question. Screen shot taken from <http://arc-assessment.com>

## 5.2 Item Types

Traditionally, paper-based assessments have included multiple choice questions (MCQ) or constructed response items. However, in online assessments, technology-enhanced items such as drag and drop were used in addition to MCQ and constructed response items. In the following section, we review and discuss the different options found in currently available

online assessment tools and programs. We also provide further suggestion and recommendation of item types in relation to NAPLAN reading assessments.

### 5.2.1 Drop-down List

Three main types of questions from the NAPLAN paper test were identified: multiple choice questions (MCQ), constructed response items and order or sequence of event. All three types of items can be easily presented on screen using the formats explained and illustrated in section 5.1. However, for order-of-event type items, students would be required to type in the numbers using a keyboard on a desktop computer or tablet (built-in or external keyboard). Research studies have revealed that some students had difficulty bringing up the built-in numeric keypad on tablets (Davis et al., 2015). To resolve this, an alternative method to consider is to use a drop-down menu or list containing number options. A drop-down button could be placed on the side of each answer box (Figure 32). To reveal a list of number options, students would click on the grey downward arrow. Students will then be able to click and select the correct number option using a mouse instead of a keyboard. Nevertheless, the usability of this method should be tested with primary and secondary school children.

### 5.2.2 Drag and Drop Items

Drag and drop items contain “dragers” and “drop bays”. Students respond to a question by moving an answer option (also referred as “dragers”) to a “drop bay” displayed on the screen. This feature can be used on both computers and tablets.

The drag and drop technique can also be used in cloze passages, where students can drag the correct words into the designated blank space to replace missing words. Likewise, sequence or order of events type questions can be answered using the drag and drop technique; an example of this was found in the sample PARCC assessment (Figure 33). This eliminates the need for students to type in their responses.



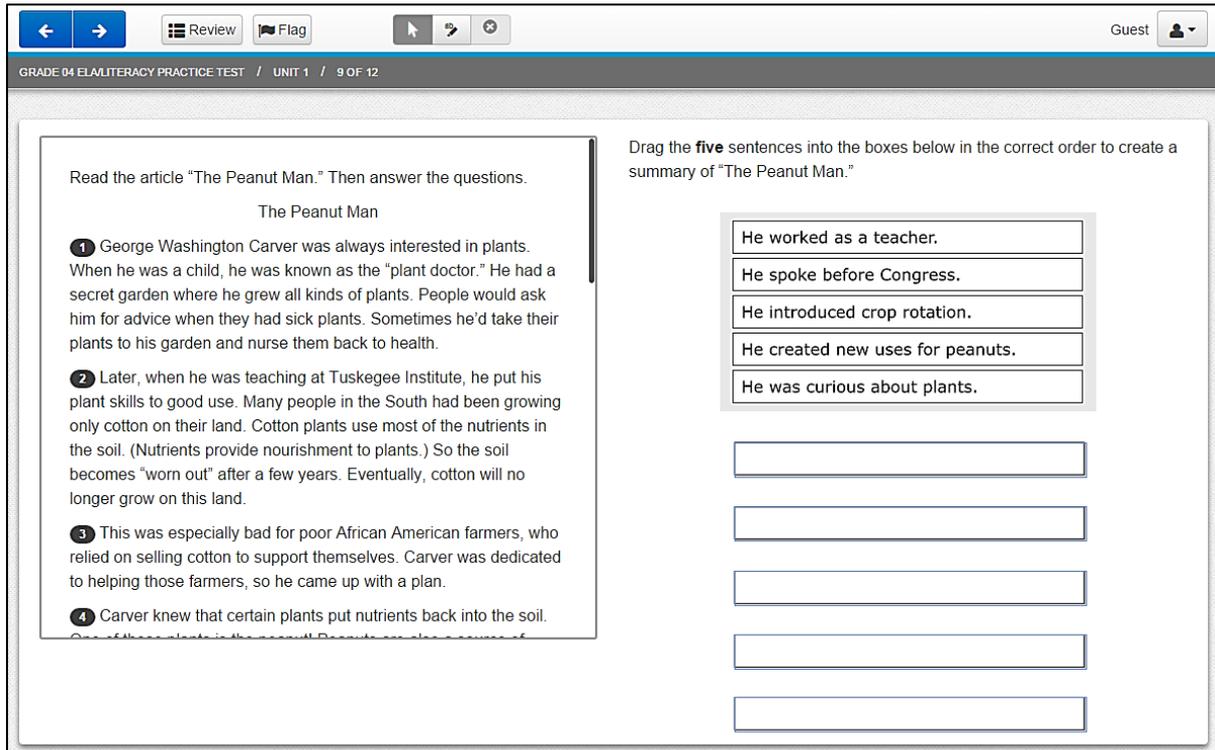


Figure 33. A sample test from PARCC which used drag and drop to list the sequence or order of a story. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

### 5.2.3 Click and Select or Select and Highlight Answers

Instead of typing responses using the keyboard, students could “click and select” or “select and highlight” sections of texts to indicate their answer(s). Apart from texts or phrases, the click and select method can also be applied to objects and graphics, or sections of an online page. The click and select method was used in the PARCC assessment as shown in Figure 34. In this example, the students were asked to select three excerpts from paragraphs 34-42 that best supported the answer to the previous question. In the passage, paragraphs 37-42 were already highlighted in light blue. Students were only required to click on sentences to indicate their answers. Sentences appeared in yellow once it had been clicked. It is important to note that, in the PARCC assessment, many questions made reference to a specific paragraph (e.g., “In paragraph 4, what is the meaning of the phrase ... ?”). Therefore, the paragraphs were numbered to help students identify corresponding paragraphs more quickly. This consideration should be taken into account if NAPLAN questions include reference to a specific paragraph.

GRADE 9 ELA/LITERACY / SECTION 1 / 5 OF 23

Today you will read and analyze a short story and a passage from another short story. As you analyze these texts, you will gather information and answer questions about each text and its relationship to the other so that you can craft a written response.

37 As he waited, he could hear a woman's fretful voice and the impatient jerk and jar of kitchen things, indicative of ill-temper or worry. The longer he stood absorbing this farm-scene, with all its sordidness, dullness, triviality, and its endless drudgeries, the lower his heart sank. All the joy of the home-coming was gone, when the figure arose from the cow and approached the gate, and put the pail of milk down on the platform by the pump.

38 "Good-evening," said Howard, out of the dusk

39 Grant stared a moment. "Good-evening."

40 Howard knew the voice, though it was older and deeper and more sullen. "Don't you know me, Grant? I am Howard."

41 The man approached him, gazing intently at his face. "You are?" after a pause, "Well, I'm glad to see you, but I can't shake hands. That damned cow had laid down in the mud."

42 They stood and looked at each other. Howard's cuffs, collar, and shirt, alien in their elegance, showed through the dusk, and a glint of light shot out from the jewel of his necktie, as the light from the house caught it at the right angle. As they gazed in silence at

**Part A**

How does the author's development of the brothers' relationship help to create tension in the story?

- A. The author provides flashbacks to show why the brothers' relationship is strained.
- B. The author highlights the brothers' differences by emphasizing the difficulty one has in recognizing the other.
- C. The author develops how the brothers have changed by including an unexpected argument.
- D. The author contrasts the brothers' relationships with their mother by revealing past disagreements.

**Part B**

Select **three** excerpts from paragraphs 37-42 that **best** support the answer to Part A.

Figure 34. An example of a “click and select” item response type. Part B of the question required students to select three excerpts from paragraphs 37-42. Paragraphs 37-42 appeared in blue highlight while the three selected texts or excerpts appeared in yellow. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

#### 5.2.4 Hot Spot Items

A hot spot item is a technology-enhanced item type that allows students to click on one or multiple areas on an image to indicate their answer (see Figures 35 and 36). To date, this item type is less commonly found in online assessment compared to “drag and drop” or “click and select”. This may be due to its scoring difficulty especially when more than one correct answer options is available.

### Hotspot Question Samples

Question 1 of 10

**Mark the auditory canal on this diagram of the ear.**

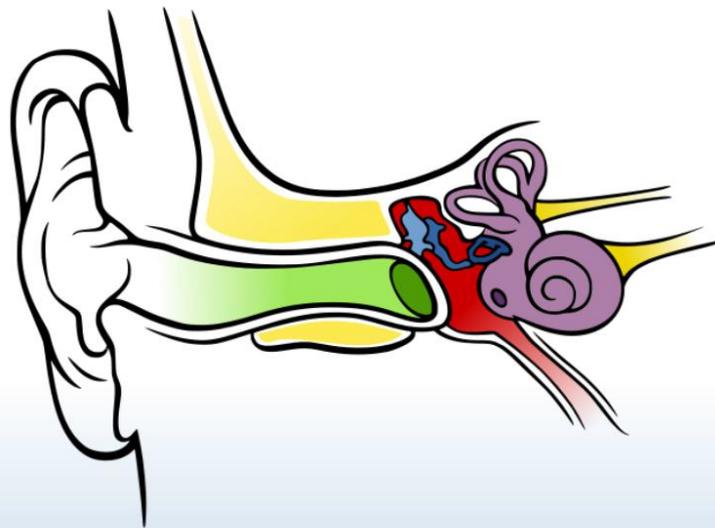


Image by Chittka L, Brockmann derivative work: M•Komorniczak (Anatomy\_of\_the\_Human\_Ear.svg) [CC-BY-2.5 (www.creativecommons.org/licenses/by/2.5)], via [Wikimedia Commons](#)

SUBMIT

Figure 35. In this example, students were asked to indicate the location of the auditory canal on a diagram of the ear. The answer options are displayed in different colours. Sample retrieved from <https://articulate-heroes.s3.amazonaws.com/stories/articulate-community/jeanette/Hotspot/PublishedSample/quiz.html>.

### Hotspot Question Samples

Question 4 of 10



Figure 36. In this example, students were asked to indicate the location of Lake Superior on a map. Sample retrieved from:

<https://articulate-heroes.s3.amazonaws.com/stories/articulate-community/jeanette/Hotspot/PublishedSample/quiz.html>.

## 5.3 Navigation options

Navigation refers to the process of traversing through a network of information resources on the interface layout. The way information is organised on paper and on screen may be quite different, and the way online content is arranged may affect how students navigate and interpret assessment contents. Presentation layout involving navigation methods such as scrolling and paging, and other innovative methods are discussed below.

### 5.3.1 Progress Bar

A progress bar is a helpful element to help readers' visualise their progress on a task, especially when the scope of the task is unknown. In an online environment, such control gives readers an idea of how far along they have come or how much more is left to be completed; this is somewhat similar to page numbers in a paper-based test which gives readers an indication of content location in context of the whole book. Several options of the progress bar are presented below:

#### ***Option 1 – Percentage marker***

A percentage marker progress bar uses some indicator such motion or percentage-completed to show that progress is taking place. Such a progress bar gives readers an idea of how much of the total task has been completed and how much is left. However, it fails to indicate which subtasks or sections in the task are completed and which ones are not. Many online surveys and assessments use this type of progress bar (on the top or bottom of the page) to show test progress (see Figure 37). If this type of control is used for online assessment, students will only know what percentage of items they have completed and what proportion is left, but will not know which items they have attempted or not. This is not very helpful for students as they would not know which item they need to review or which unattempted items were left out intentionally or by mistake. Students would need to go through each item in the assessment to identify items that were unattempted.

The screenshot shows a web-based test interface. At the top, there is a header with the text 'Numerical Aptitude Test - Free Sample Test' and a 'Test Progress' bar that is approximately 50% full. To the right of the progress bar, it says '58'. In the top right corner, there is a 'Time Remaining' indicator showing '7:29 Minutes'. The main content area is titled 'Question #6' and contains a table with the following data:

	Factory 1	Factory 2	Factory 3	Factory 4
Annual sales (in millions of dollars)	14.432	18.48	16.422	22.94
Total products manufactured	943	625	729	1245
Number of faulty products manufactured	41	9	15	98
Manufacturing wages (in millions of dollars)	2.5	2.65	2.4	2.82
Number of manufacturing staff	35	42	47	58
Material cost	10.1	15.7	11.3	18.1

Below the table, the question asks: 'Which factory has the highest percentage of good products manufactured?'. There are five radio button options: 'Factory 1', 'Factory 2', 'Factory 3', 'Factory 4', and 'Factory 2 & 4'. At the bottom of the interface, there are navigation buttons for 'Previous' and 'Next', and a copyright notice for the Institute of Psychometric Coaching.

Figure 37. Example of a progress bar with test progress percentage information (indicated by a green bar presented above the question) found in a Verbal Aptitude Test by the Institute of Psychometric Coaching. Screen shot taken from <http://www.psychometricinstitute.com.au/test-area-free.asp?testID=13>

### **Option 2 – Item panel**

An item panel is a more sophisticated version of a progress bar. Item panels are becoming increasingly popular and frequently utilised in online assessments. This concept utilises the web technology to display a range of information at once about the assessment being undertaken. The item numbers usually appear in sequence and displayed as a progress bar, where each numbered button can be clicked to directly go to that item (works like a review option) and the numbered buttons changes colour to indicate whether items have been attempted (see Figure 38). This type of progress bar design is likely to be more useful to students as it gives them a brief overview of their progress on the assessment. In addition, students can easily identify unattempted items and access those items directly as needed. The panel also gives an indication of how many total items there are in an assessment, allowing students to allocate their time for the task as they see fit. The progress bar with the item panel can appear at the top, bottom or side (e.g. see Figure 39) of a page on screen.

PRACTICE  
APTITUDE TESTS

19

48

Question: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

### Numerical Reasoning Test 1

Fee Income by Consultant (£1,000's)

May days by consultant	Kim	Kris	Sam	Tim	Trevor
Daily fees per consultant (£00's)	6.5	11.2	9.5	10.2	10.9
Hours worked (7 chargeable hours per day)	148	130	145	134	124
Mileage expenses claimed* (miles)	350	528	430	296	388

© Practice Aptitude Tests Ltd

Question 20

What was the difference in the travel expenses claimed by Kim, Sam and Trevor, compared to Kris and Tim (to the nearest £10)? Assume mileage is reimbursed at £0.4 per mile.

- £130
- £140
- £150
- £160

« Previous      Next »

TEST RESULTS »

Figure 38. Example of an item panel progress bar in a Numerical Reasoning test designed by the Practice Aptitude Tests. Screen shot taken from <http://www.practiceaptitudetests.com/numerical-reasoning-tests/>

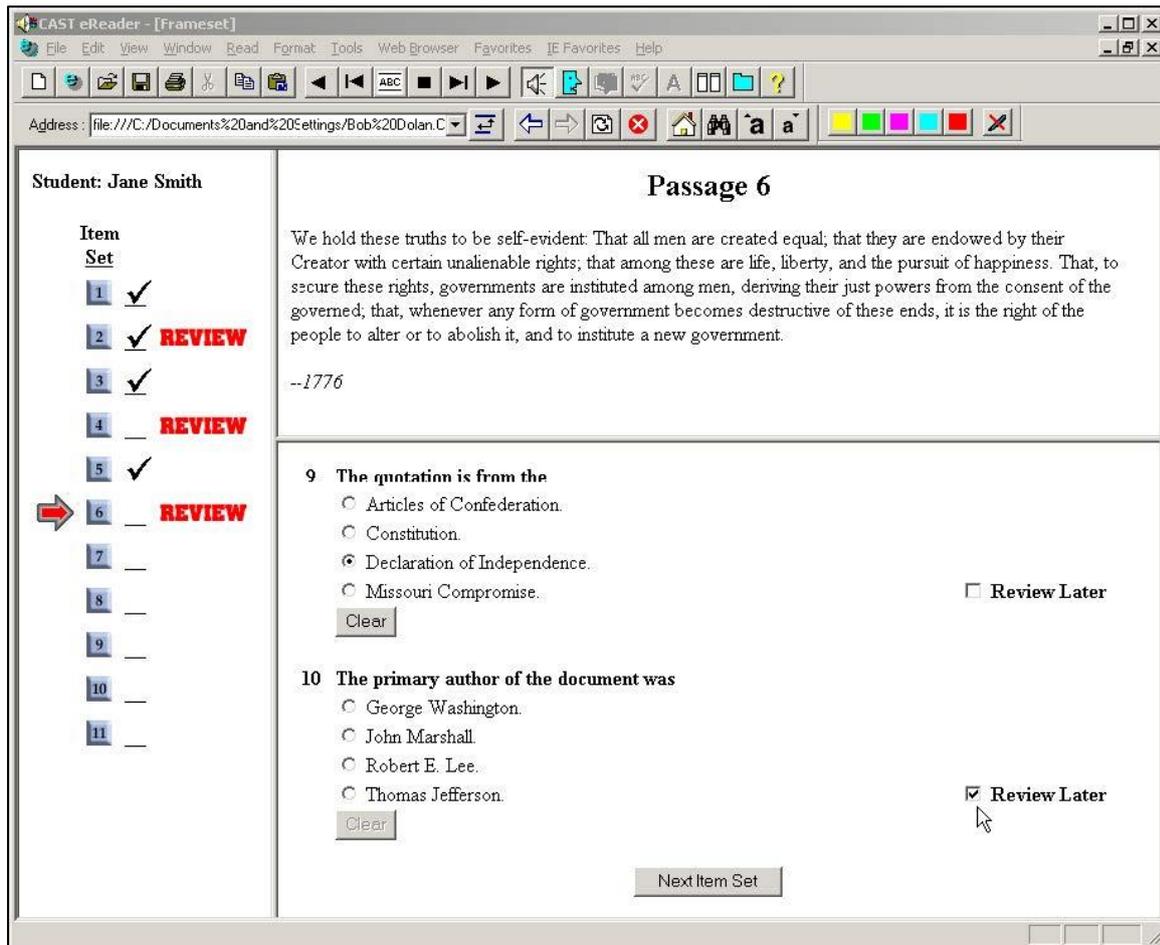


Figure 39. Example of progress bar with item panel and review options as used in a study by Dolan, Hall, Banerjee, Chun & Strangman (2005).

### Option 3 – Popup console

The idea and design of a progress indicator as a popup console is very similar to the item panel option. Unlike the usual progress bar which is permanently placed on the screen, a popup console progress bar can be displayed at the user’s discretion (i.e. the popup console can be opened or closed) (see Figures 40 and 41; screen shots taken from <https://www.scrum.org/Assessments/Open-Assessments/Scrum-Open-Assessment>). Once view capability is expanded, test takers would be able to see items that have been completed and would be able to go directly to any unanswered items.

**Scrum Open**  
Time left: 0:28:22

---

**Question 9 of 30**

What is the main reason for the Scrum Master to be at the Daily Scrum?

- To make sure every team member answers the three questions.
- To write down any changes to the Sprint Backlog, including adding new items, and tracking progress on the burn-down.
- He or she does not have to be there; he or she only has to ensure the Development Team has a Daily Scrum.
- To gather status and progress information to report to management.

[◀ Previous](#) [Next ▶](#)

---

[Display previously viewed questions ▼](#)

Figure 40. Example of a popup console as suggested by the Scrum.org organisation for designing tests. In this screen shot, the pop-up console is minimised, a hyperlink titled “Display previously viewed questions” is presented below the question.

**Scrum Open**

Time left: 0:29:24

---

**Question 9 of 30**

What is the main reason for the Scrum Master to be at the Daily Scrum?

---

- To gather status and progress information to report to management.
- He or she does not have to be there; he or she only has to ensure the Development Team has a Daily Scrum.
- To make sure every team member answers the three questions.
- To write down any changes to the Sprint Backlog, including adding new items, and tracking progress on the burn-down.

◀ Previous
Next ▶

---

- 1 ▶ During the Daily Scrum, the Scrum Master's role is to:
- 2 ▶ How much work must a Development Team do to a Product Backlog item it selects for a Sprint?
- 3 ▶ The Product Backlog is ordered by:
- 4 ▶ When does the next Sprint begin?
- 5 ▶ What is the recommended size for a Development Team (within the Scrum Team)?
- 6 ▶ The length of a Sprint should be:
- 7 ▶ The Development Team should have all the skills needed to:
- 8 ▶ Which statement best describes Scrum?

Figure 41. Example of a popup console as suggested by the Scrum.org organisation for designing tests. In this screen shot, the pop-up console is opened by clicking on the hyperlink titled “Display previously viewed questions”. A list of preceding questions is displayed. Test-takers can access specific items directly by clicking the item number.

### 5.3.2 Scrolling

Scrolling is a common navigation method; it allows texts to be moved up and down a screen smoothly by a fixed increment to reveal content that is out of view. Different ideas of scrolling are discussed here.

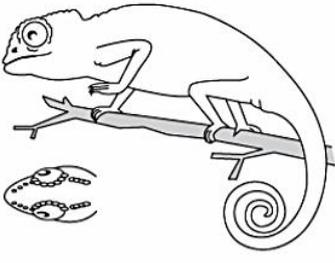
#### ***Whole page scrolling***

With whole page scrolling, online passages are presented as a single block of text. The passage will be published in its entirety and readers will need to scroll up or down to manipulate the view of the text. A passage with all its corresponding items/questions can appear on the page at the same time (i.e. passage at the beginning and all items at the end of the page or passage on the left and all items on the right of the page). Alternatively, the passage can be presented on screen with one corresponding item at a time (i.e. passage at the beginning and one item at the end of the page or passage on the left and one item on the right of the page). Many testing organisations and research studies (e.g. Online PAT-R test) use either of these two options (see Figure 42). This scrolling technique may be easier to manipulate by those who are more familiar with technology, that is, novice users may find it difficult to navigate through a page using whole page scrolling. Moreover, the navigational experience of scrolling on different devices can vary a lot – scrolling on tablets can be perceived easier compared to computers, as tablets allow scrolling with fingers which is faster and a natural process. This scrolling feature with its various positional cues provides readers with limited information about the location of text contents. At the same time, whole page scrolling may disrupt users' mental model of the spatial layout of onscreen content (i.e. structural representation of the text).

« Previous Item | Item 7 | Next Item »

## CHAMELEON

*Our class learned about chameleons. We made a special book. Here is one page from our book.*

Jai	Chameleons eat insects. Chameleons have very long, sticky tongues. When they see an insect, they flick out their tongue and then the insect gets stuck on it
Hilary	Chameleons can change the colour of their skin to match where they are. This helps them to hide from their enemies.
Steven	

*I found this picture of a chameleon on the internet. You see*

### What do Hilary and Karen both tell you?

A: how chameleons grow	0   0%
B: how chameleons move	0   0%
C: how chameleons find food	1   50%
D: how chameleons keep safe	1   50%

**Scale score (patc) difficulty level:** 100.0

**Classification:** Interpreting explicit information

**Description:** Identifies the type of information in a section of a table

Figure 42. Example of the Online PAT-R test item design by ACER with whole page scrolling – passage on the left and one item on the right of the screen. Screen shot taken from <http://www.acer.edu.au/oars/tests-available>

### Section Scrolling

In a whole page scrolling environment (i.e. whole window or screen), readers have to scroll to make text move through a display window. To eliminate issues related to whole page scrolling, assessment content can be presented in sections. Using this method, passages or texts are presented on a scrolled page (either on the left or right, or on the top half of a page), and the questions with response options can be presented on the other side and remain fixed on the page. The items/questions remain static in the same position and students will only scroll the passage section (see examples in Figures 43 to 45). This will enable students to refer to the questions and text at the same time and not lose concentration when searching for information. The question and response options would remain in view for the students to evaluate. The advantages of this method are quite similar to that of dynamic text option mentioned in section 5.3.4, but are less likely to be impacted by students' familiarity with technology. Many research studies and large scale assessments (e.g. PARCC, SBAC) utilise this method for reading comprehension tests. Some researchers have also suggested providing page-up and page-down buttons instead of a vertical scroll

bar to reduce any interference which may affect younger students (Choi & Tinkler, 2002; Piolat et al., 1997).

The screenshot displays a reading assessment interface. On the left, a text box contains the following content:

**Reading Passage: Penny's Scrapbook**

*Penny is a golden retriever puppy from Texas. She is training to be an assistance dog (a dog that helps people who need help in their daily lives). Penny wants to be a guide dog for people who are visually impaired (have trouble seeing). Read about her first year in the story below.*

**Penny's Scrapbook**  
by Rebecca A. Alter

My name is Penny. When I grow up, I want to be a guide dog for someone who is blind or severely visually impaired. You might have heard of junkyard dogs, well, I'm a schoolyard dog. I was born in a second grade classroom in Austin, Texas. The students at Sam Rayburn Middle School raised the money to buy me for the **Southwest Guide Dog Foundation** by collecting pennies. (That's how I got my name!) I will be at Rayburn this school year with my "Puppy Walker," Becky Alter, who teaches art. . . .

**My First Months**  
I joined Becky Alter in April. She will be my "Puppy Walker" until I am about 18 months old. My first weeks of school with Becky were exciting. I was exposed to the sounds in the hallway when classes changed and got

At the bottom of the text box are two buttons: "highlight text" and "remove highlight".

On the right side of the interface, a green header states "There are 5 questions for this passage". Below this is a yellow section titled "Begin Questions".

1. At school, Penny's "Puppy Walker," Becky,  raises money for guide dogs.  teaches biology.  teaches art.  tattoos dogs.

2. Becky takes Penny everywhere she goes because she  does not trust Penny to stay alone.  wants Penny to get used to different places.  does not like to be separated from Penny.  promised she would never leave Penny alone.

Below the questions is a blue button with a downward arrow and the text "more questions".

At the bottom right, there is a row of 19 numbered boxes (1-19) and a red "End Test" button.

Figure 43. Example of a section scrolling layout used in a reading assessment, designed for a research study by Higgins et al. (2005).

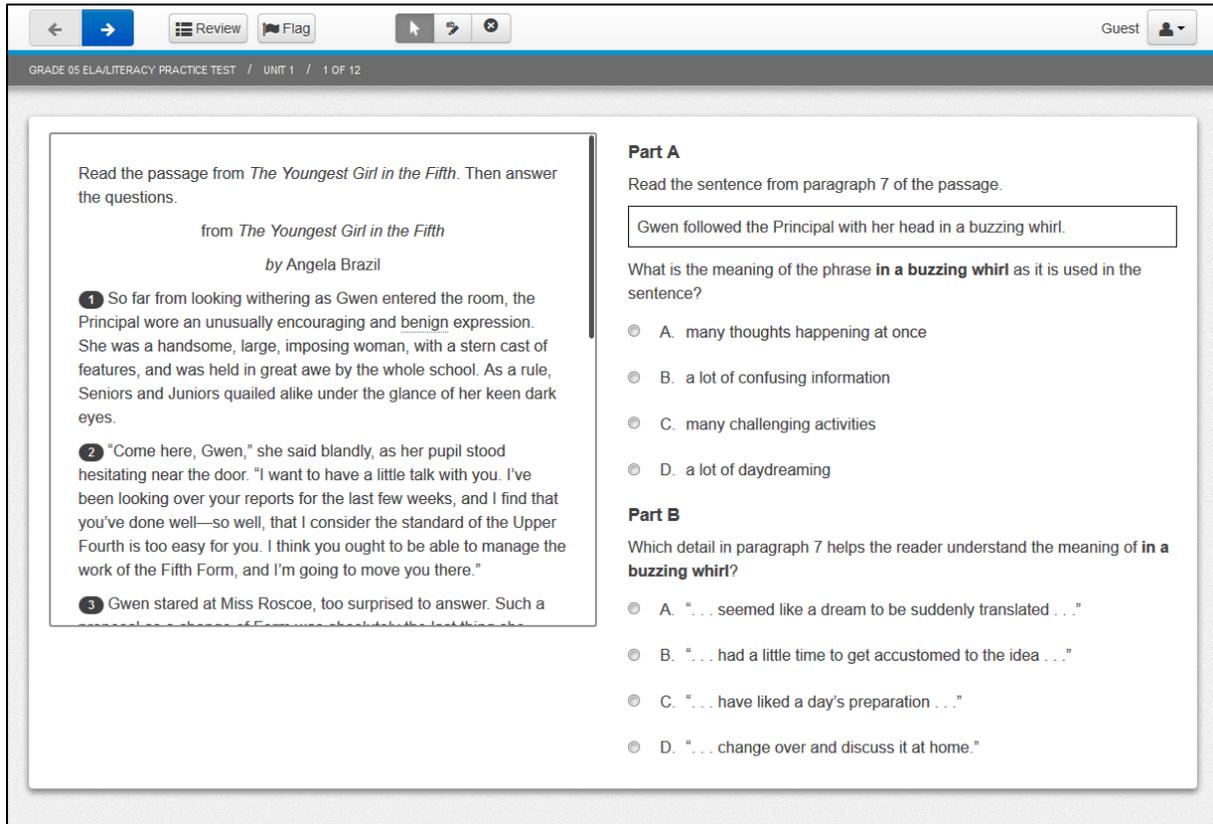


Figure 44. Section scrolling layout design used in a PARCC reading assessments. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

The screenshot displays the Smarter Balanced Assessment Consortium interface. At the top, there are navigation links: "BACK TO SAMPLE ITEMS HOME", "VIEW MORE ENGLISH LANGUAGE ARTS / LITERACY SAMPLE ITEMS", and "ABOUT THIS ITEM". The page title is "English Language Arts / Literacy | Mathematics". On the right, there are buttons for "Feedback and Support" and "Send an Email", along with "BACK", "NEXT", and "ITEM SCORE" buttons. The main content area is titled "Diamonds in the Sky" with an item ID of "43006". The passage text is as follows:

**Diamonds in the Sky**

Stars are not the only objects that glitter in the dark night sky. Scientists have discovered that diamonds are plentiful in outer space. Some of these space diamonds are called "nanodiamonds" because they are incredibly small. A nanodiamond is millions of times smaller than a grain of sugar—more or less the size of a strand of DNA. Nanodiamonds are stardust, created when ancient stars exploded long ago, disgorging their remaining elements into space. Other space diamonds are huge—the size of whole planets—while some may exist in liquid or frozen form. Scientists even suggest that planets in our own solar system may have oceans filled with chunks of frozen diamond "ice."

Diamonds are so common throughout the universe because they are a pure form of one of the universe's most common elements: carbon. Diamonds have a number of amazing properties: they are extremely hard and transparent, and can withstand radioactivity, corrosive acids, and other powerful forces. Diamonds conduct electricity more readily than copper, and are also the best natural conductor of heat that we know of—which is why diamonds feel cool to the touch. Like a prism, diamonds produce rainbows from white light. The melting point of a diamond, 7,362 degrees Fahrenheit, is higher than that of any other known substance.

Graphite and diamonds share the same chemistry—both are carbon. The difference lies in the arrangement of the carbon atoms, known as their "molecular structure." Extreme forces are required to transform dark, soft graphite—the stuff used in pencil lead—into hard, brilliant diamonds. A diamond is formed when carbon is exposed to immense pressure and extreme heat—conditions found hundreds of miles below the surface of the Earth, where most natural diamonds are formed. The heat and pressure squeeze the carbon atoms into a dense, crystalline structure. In the comic books, Superman could create a diamond by simply squeezing carbon in his bare hands, but it normally takes billions of years for carbon to become a diamond.

In space, diamonds are born more quickly. Scientists believe space diamonds often crystallize in no more than a millionth of a millionth of a second, when dust grains containing carbon smash together at extremely high speeds. Another hypothesis for how space diamonds are formed involves the shock waves released by an exploding star, which cook and compress carbon dust until it becomes a diamond.

The question asks: "Read the sentence from the text. Then answer the question. 'Nanodiamonds are stardust, created when ancient stars exploded long ago, disgorging their remaining elements into space.' Based on the context of the sentence, what is the most precise meaning of disgorging?"

The answer choices are:

- (A) scattering randomly
- (B) throwing out quickly
- (C) spreading out widely
- (D) casting forth violently

Figure 45. Section scrolling layout design utilised in the Smarter Balanced reading assessments. Screen shot taken from <http://www.smarterbalanced.org/sample-items-and-performance-tasks/>

### 5.3.3 Pagination

Paging or pagination is the facility to move the text up or down in complete screen sections; this is analogous to turning pages with books or paper documents. For reading assessment, the passage is broken into subsections and each subsection is presented as a single block of text on a page. Passages and items are usually presented in section, where the passage would appear on the left or top using paging and the items would be laid out respectively on the right or bottom (i.e. one item at a time or multiple items together). However, pagination can be used to provide two separate views, one representing detailed content of the area that is in focus and another to show the overall context of that representation.

Researchers have argued that paginated interface is unrelated to the amount of text being presented at once, thus it may appear to provide the structural stability of the page-by-page presentation. Literature suggests that although scrolling mechanism was most frequently used, novices tend to prefer paging perhaps due to its similarity to reading a book in paper form. Such findings are not surprising considering the advantages that pagination can

provide. Pagination allows a passage to be separated into comprehensible pieces while at the same time provides an overview of the interconnections between these pieces. Additionally, in various research experiments, it was found that to some degree pagination helped people to read complex texts faster and also to process them more efficiently (i.e. a positive effect on information processing). This is more likely as it helps readers focus on the gist of the text. Furthermore, pagination is likely to increase the understanding of material that requires sustained attention over a period of time. However, even with the paging type of display, readers have a less accurate idea about the vertical position of any information in the text. The only indication about where in the text the current display is located is usually given by the page number or some other form of numbering mechanism. Nonetheless, it is one of the most preferred methods cited in the literature. Many research studies and large scale assessments (e.g. PISA) utilise this format for reading comprehension test (see Figures 46 to 48).

The screenshot shows a digital reading interface. On the left, under the heading "Reading Passage: Penny's Scrapbook", there are three paragraphs of text. The first paragraph discusses a conference. The second paragraph describes a summer trip with a dog named Penny. The third paragraph, titled "Back to School", describes the start of a new school year. At the bottom of the text area are navigation buttons: "Back" (yellow arrow), "highlight text" (blue button), "remove highlight" (grey button), and "Next" (green arrow). On the right side, a green header states "There are 5 questions for this passage". Below this is a yellow section titled "Begin Questions" containing two multiple-choice questions. Question 1 asks about Penny's "Puppy Walker," Becky, with options: raises money for guide dogs, teaches biology, teaches art, and tattoos dogs. Question 2 asks why Becky takes Penny everywhere, with options: does not trust Penny to stay alone, wants Penny to get used to different places, does not like to be separated from Penny, and promised she would never leave Penny alone. Below the questions is a blue button labeled "more questions" with downward arrows. At the bottom right is a red "End Test" button. A horizontal bar at the very bottom contains a sequence of numbered boxes from 1 to 19, with the 19th box highlighted in red.

Figure 46. Example of pagination layout for reading items found in a research study by Higgins et al. (2005).

The screenshot shows a digital interface for a PISA 2015 reading assessment. At the top, a blue header bar contains the text 'PISA 2015' on the left and navigation icons (a question mark, a left arrow, and a right arrow) on the right. The main content area is divided into two vertical panels. The left panel has a light blue header with the text 'Democracy in Athens' and 'Question 3/5'. Below this, it prompts the user to 'Type your answer to the question below.' and provides a short paragraph about the purpose of a speech in part B. It then asks 'What was ANOTHER purpose of this speech?' and features a large, empty grey rectangular box for the user's response. The right panel is titled 'Democracy in Athens' and is labeled with a '1' in the top left and a '2' in the top right corner of its content area. It contains two paragraphs of text under the heading 'PART A'. The first paragraph describes Thucydides as a historian and military man who lived in the fifth century BC, detailing his role during the Peloponnesian War and his exile. The second paragraph discusses his historical methodology, noting his focus on natural causes and individual behavior rather than fate or divine intervention, and mentions his use of fictitious speeches to explain historical characters.

Figure 47. Pagination layout proposed for PISA online reading assessments. To move through pages, students can click on the folded corner of the page. Image taken the PISA draft reading framework article (OECD, 2013).

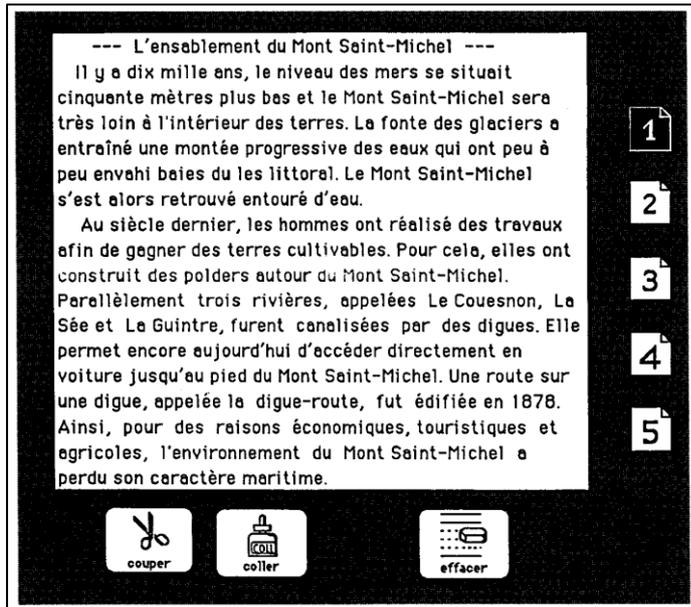


Figure 48. Example of pagination layout for reading items utilised in a research study by Piolat et al. (1997)(1997)(1997). Each page is displayed as a thumbnail on the right side.

### 5.3.4 Innovative options

#### **Option 1 – Dynamic Text**

On a scrolled page, only some part of the page is in view, while other contents are hidden from view. Scrolling a whole page can disrupt a student's ability to locate information asked in the question and to find a reference of it from the passage – this issue is more visible with lengthy texts. Scrolling through the text to find a particular piece of information can be slow, especially if the desired information is far from the current location. For example, if the passage and questions are presented side-by-side on a scrolled page, the student may scroll through to the bottom of the passage but the question is located on the top of the page (interface designed with items on the right and passage on the left). Similarly, students may be reading the beginning of a text passage but the questions appear at the bottom of the page (interface designed with items on the bottom and passage on the top). This layout may impact on their ability to relate questions to information from the passage. To improve reading flow and to eliminate any issues related to spatial mental representation of the text, a new and innovative approach can be used – we term this as *dynamic text*.

The idea of dynamic text in assessment is consistent with some webpage design where some parts of the page contents (such as ads, links, contact us, live chat invitation for service related sites etc.) are presented such that regardless of the reader's scroll position in the page, those elements will appear – as though they are “following” readers during their scroll navigation. Please see <http://www.vodafone.com.au/personal> for the concept and idea of dynamic text. On this web page, a “Hot Offers” tab appears on the right side next to the scroll bar. This tab remains in view and is fixed on screen regardless of the reader's scrolling behaviour.

This concept can be applied to an assessment context, where the dynamic element (e.g., “Hot Offers” tab) can be the item itself (i.e. the questions and options). This approach will allow the entire screen space to be used for displaying the stimulus (i.e., text passage). The dynamic element containing the question and response options could be presented in three different formats: (i) a dynamic text box with questions and response options which is fixed on the screen (Option 1), (ii) a dynamic text box with can be moved around the screen (Option 2), and (iii) a pop-up window/tab which can be minimised and maximised (Option 3; i.e., similar to the Vodafone “Hot Offers” tab, except that clicking the tab/link would not bring readers to a different page).

In Option 1, the dynamic text box remains in the same location regardless of readers' scroll position on a page (see examples in Figure 49). The question and response options (in the case of a MCQ item) are visible and are presented within the text box.

In Option 2, the same presentation format from Option 1 is adopted. However, the text box can be moved to different locations or parts of the screen (similar to example from the Jetstar web page: <http://www.jetstar.com/au/en/home>, shown in Figure 50).

In Option 3, a pop-up window/tab is presented on the right, next to the scroll bar (similar position as the “Hot Offers” tab on the Vodafone web page). When this tab is clicked, the question and response options will be displayed on a pop-up window appearing next to the passage. This window can be minimised or maximised (see example in Figure 51). Using these dynamic text options would ensure that the passage and questions are presented on the same viewing screen – this would allow students to refer to the passage and questions at the same time.

These options are considered enhanced and innovative methods of presenting assessment content. They can be applied across devices, and can be implemented with whole page scrolling interface, or paging/pagination methods. However, it remains to be investigated whether such a novel way of presenting assessment content may disrupt the reading process and impact on students' reading performance.

### ***Option 2 – Interactive Text***

There are various presentation and storytelling tool (e.g., Prezi) available online for presenting ideas on a virtual canvas. Using web technologies, information on screen can be presented in new and innovative ways. For instance, presentation ideas from Prezi-like software could be built into assessment platforms to present complex and lengthy documents – we term this as *interactive text*.

With this technique, all information (i.e. a passage and a question) will be presented on the screen at once. This will give participants a perspective of the entire assessment. The in-built mechanism allows user to zoom in and out of the presented content (including rotate and resize the page content) – allowing users to display, view and navigate through information (i.e. smaller or bigger pieces of text) at their own pace (see <https://prezi.com/>). Such presentation method is likely to eliminate issues related to lengthy text presentation on smaller screens. The concept could be applied to the passage and/or item sections. However, applying this concept to items and questions may be less suitable as younger children may not know how to navigate through the page or access the questions; it is better to leave questions/items in full view. This presentation method is novel and unconventional in online assessments compared to traditional ways of presenting texts and items. Therefore, before implementing this concept to an online assessment platform, this idea will need to be investigated to determine whether this presentation and navigation format would affect students' test performance.

**Jetstar** Log in Travel agents Web check-in Flight status Manage booking Australia (English)

Flights Packages Hotels Cars & extras Special offers Plan & book What we offer Need help? Ask Jess

### Plan your trip

Return One way Multi-city

Origin:  Adults: 1

Destination:  Child 2-11: 0

Depart:  Return:  Infants <2: 0

I need to travel on this date  
 I just want the lowest fare

Children travelling alone

**Find Flights**

Your recent searches: Melb... → Dunedin 25 Dec - 09 Jan

## Get up to 50,000 bonus Qantas Points\*

FIND OUT MORE

Card Services is a division of Macquarie Bank Limited (ABN 46 008 883 542, Australian Credit Licence 237502), which provides and administers credit and is the issuer of the Jetstar MasterCard and Jetstar Platinum MasterCard. \*Credit criteria, fees, charges, spend requirements and conditions apply.

### woohoo BARGAIN BIN

BAG YOURSELF A BARGAIN TODAY!

Departing from Melbourne (Tullamarine) Bag a bargain

Destination	Price from*	Travel Period	Fare Type
Fiji (Nadi) Via Gold Coast	\$217	8 Oct - 9 Dec 2015	SALE FARE
Hobart	\$65	Mid Jul-15 - Late Sep-15	HOT FARE
Perth	\$179	Late Aug-15 - Mid Sep-15	HOT FARE
Launceston	\$55	Early Aug-15 - Late Aug-15	HOT FARE
Bangkok	\$279	Mid Feb-16 - Mid Mar-16	HOT FARE

Travel Alert: Volcanic ash update - Bali

Club Jetstar: Exclusive member sales, partner offers and more! JOIN NOW

Plan & book your next trip anywhere, anytime with the Jetstar app

### woohoo BARGAIN BIN

BAG YOURSELF A BARGAIN TODAY!

Departing from Melbourne (Tullamarine) Bag a bargain

Destination	Price from*	Travel Period	Fare Type
Fiji (Nadi) Via Gold Coast	\$217	8 Oct - 9 Dec 2015	SALE FARE
Hobart	\$65	Mid Jul-15 - Late Sep-15	HOT FARE
Perth	\$179	Late Aug-15 - Mid Sep-15	HOT FARE
Launceston	\$55	Early Aug-15 - Late Aug-15	HOT FARE
Bangkok	\$279	Mid Feb-16 - Mid Mar-16	HOT FARE
Phuket	\$299	Early Nov-15 - Late Dec-15	HOT FARE

\*One-way, checked baggage not included. Hot fares strictly limited, available until sold out. Sale fares available until sale ends, unless sold out prior. Limited availability on peak days and school holidays. Selected travel dates and conditions apply.

Travel Alert: Volcanic ash update - Bali

Club Jetstar: Exclusive member sales, partner offers and more! JOIN NOW

Plan & book your next trip anywhere, anytime with the Jetstar app

FREE Download on the App Store and Google Play

### Packages

View all

**Gold Coast**  
3 nights in a Superior Room at Jupiters Hotel and Casino  
From **\$366**

**Melbourne (Tullamarine)**  
3 nights in a City View at Mercure Melbourne Treasury Gardens  
From **\$338**

**Cairns**  
3 nights in a Standard Room at Novotel Cairns Oasis Resort  
From **\$431**

**Ask Jess**  
Jetstar's virtual assistant

Figure 49. Screen shots from Jetstar web page. A dynamic text box (“Ask Jess”) appears on the right side of the screen, next to the scroll bar. The location of this text box remains the same even when page is scrolled up or down.

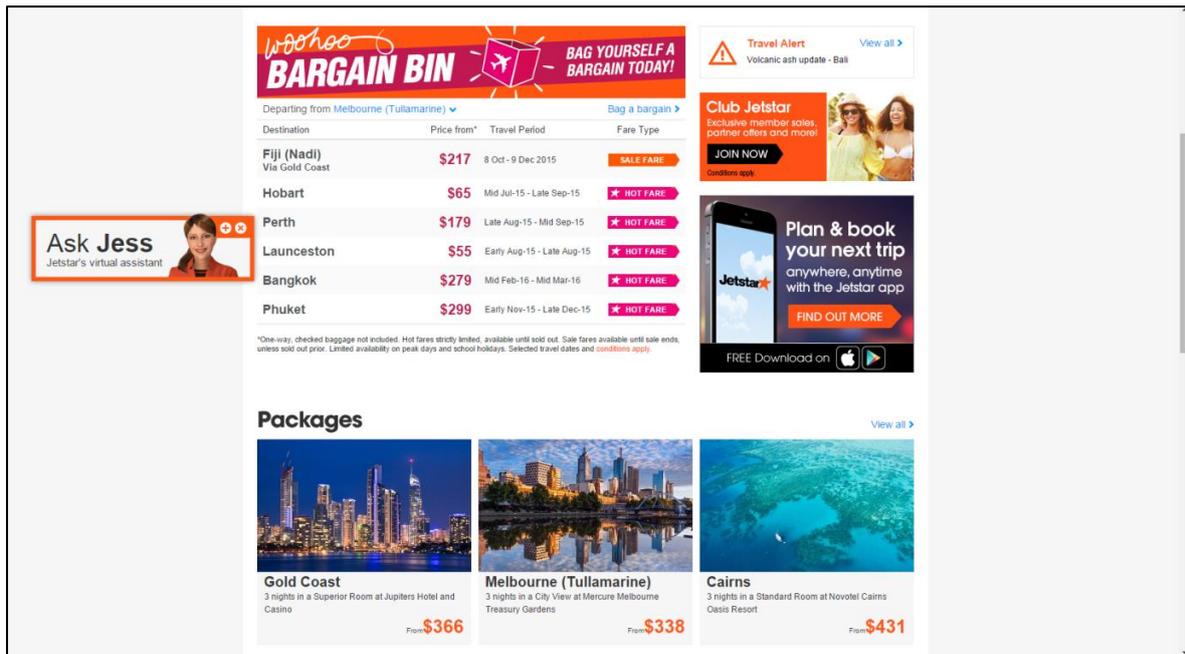


Figure 50. In this screen shot, the dynamic text box (“Ask Jess”) is moved to the left side. This text box can be moved and placed in any parts of the screen.

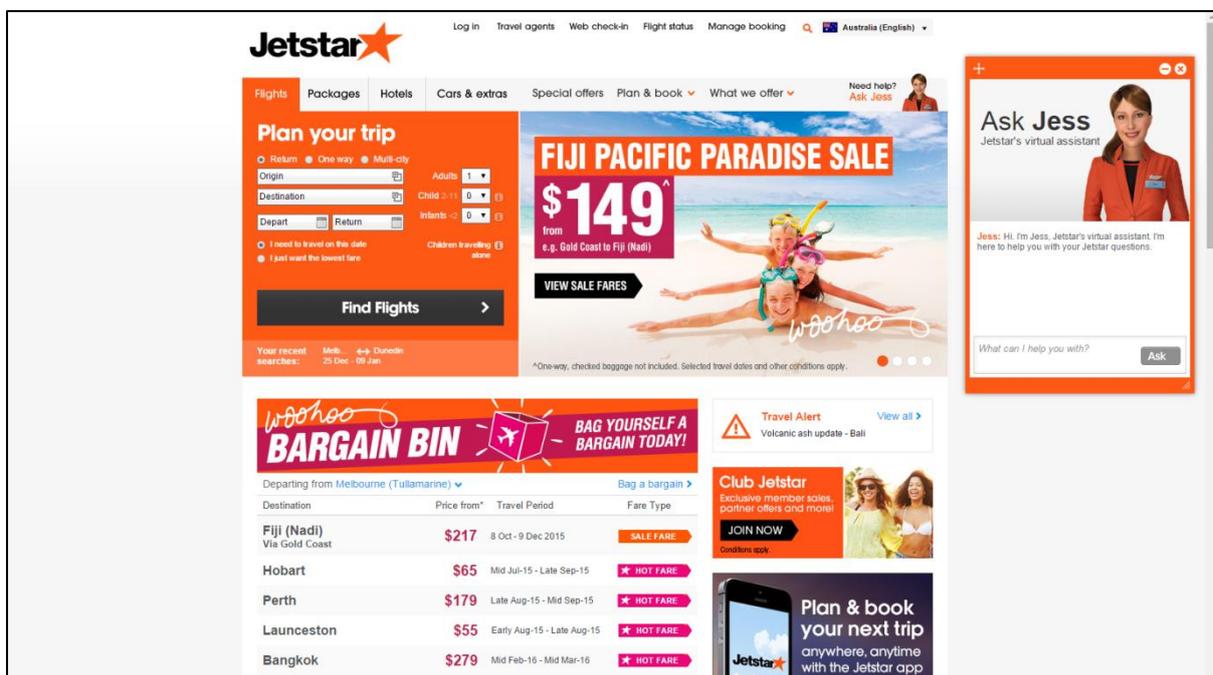


Figure 51. The dynamic text box (“Ask Jess”) is maximised, allowing users to type in their questions to Jess. This concept can be used for online reading assessment, for multiple choice questions and constructed responses. The text box can be minimised and moved to different part of the screen.

## 5.4 Reading Tools

Many software applications and web browsers have included tools to make reading on screen easier. These vary from tools such as highlighter, tools to enlarge the text size or option to change background colour. Some online assessment programs such as the PARCC have also incorporated tools such as a line reader, highlighter and flag tool in their tests to assist students with test-taking. In the following section, we provide details of onscreen reading tools that are currently used in national or international online assessment programs.

### 5.4.1 Highlighter

The use of highlighter on screen is similar to highlighting on paper. On a computer, readers will first select a section of text using the mouse cursor. The text will immediately be highlighted in the predetermined colour. Alternatively, some highlight tools such as the one found in PARCC assessment (see Figure 52) allow readers to select the colour of a highlight.

The screenshot shows a web browser window with a navigation bar at the top containing a back arrow, a forward arrow, a 'Review' button, a 'Flag' button, and a 'Guest' user profile. Below the navigation bar, the page title reads 'GRADE 3 ELA/LITERACY / SECTION 1 / 5 OF 18'. The main content area is divided into two columns. The left column contains the text of a story titled 'A Once-in-a-Lifetime Experience' by Sandra Beswetherick. The text is numbered 1 through 6. A highlighter tool is visible over the text, with a color palette showing yellow, orange, red, and blue. The second sentence is highlighted in blue. The right column contains 'Part A' with a reading instruction and a text box containing a paragraph about a cola bottle. Below the text box is a question about the meaning of 'that bottle erupted like Mount Vesuvius' and four multiple-choice options (A, B, C, D). Below that is 'Part B' with a question about which detail supports the answer to Part A and two multiple-choice options (A, B).

Figure 52. The highlight tool in PARCC reading assessment allows readers to choose the highlight colour. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

## 5.4.2 Line Reader

A line reader can be used to help guide the reader's eyes, especially when reading long documents on screen. Using a line reader would help readers to follow through sentences in between lines and reduce errors such as accidentally skipping or repeating lines when reading. A line reader is usually a horizontal line (similar to a ruler) that is placed below a sentence. The PARCC assessment incorporated a line reader in their tests; however the design of their line reader looked more like a reading window than the conventional horizontal line (Figure 53). Readers could adjust the size of the line reader, "reading window" (i.e., space where text you want to read is displayed) and also the "viewing pane" (i.e., space below the reading window) (Figure 54). The usability of a line reader, including the different designs (e.g., horizontal line vs. reading window) should be tested on different devices and with children of different ages.

Figure 53. A line reader from the PARCC reading assessment. Screen shot taken from

<http://parcc.pearson.com/practice-tests/english/>

The screenshot shows a web browser interface for a reading test. At the top, there are navigation buttons for 'Review' and 'Flag', and a 'Guest' user profile. Below the browser header, the page title is 'GRADE 5 ELA/LITERACY / SESSION 1 / 1 OF 23'. The main content area is divided into two columns. The left column contains instructions: 'Today you will analyze a passage from the novel *Ida B* and a passage from the novel *Moon Over Manifest*. As you read these texts, you will gather information and answer questions about the influence of the narrator's point of view so you can write an essay.' Below this, it says 'Read the passage from the novel titled *Ida B*. Then answer the questions.' The passage text is: 'from *Ida B*  
by Katherine Hannigan  
1 Saturday morning, I was sitting on the front porch, waiting for nothing, with nothing I wanted to do. Rufus sat beside me for a while, hoping I'd be up to something more than misery. But he got tired of waiting and went off on his own, leaving a small sea of spit where he'd been sitting.' The right column is titled 'Part A' and contains the question: 'Read the sentence from paragraph 1. Rufus sat beside me for a while, hoping I'd be up to something more than misery. What does the word **misery** mean as it is used in the sentence?' There are four multiple-choice options: A. confusion, B. "... tired of waiting and went off on his own ...", C. nervousness, and D. unhappiness. A large green rectangular area is overlaid on the passage text, indicating it is expanded.

Figure 54. The line reader can be expanded, including the “reading window” and the “viewing pane”. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

### 5.4.3 Magnifying Tools

There are three ways to magnify or enlarge texts in a passage. The first option is to use a magnifier which is placed over a section of text to enlarge the words. This simulates reading with a magnifier in the real-life context. An example of the magnifying tool from PARCC is presented in Figure 55. The second option is to allow readers to choose different font size options, such as small, medium or large text. This feature is commonly found on websites and e-readers (Figure 56). This option may affect the layout and presentation of passages and graphics, especially on smaller devices such as a tablet. The third option is to include an expandable passage section, as illustrated in a sample test from Smarter Balanced (Figure 57). When the passage section is expanded, the text is displayed on more than half of the screen and thus accommodating most of the text passage. In this view, students will not have to scroll as much when reading. The only disadvantage of using the expanded view is that the question becomes hidden and is no longer in view together with the passage. All

three options should be tested across devices, and the usefulness and usability of these magnifying tools for online assessment purposes should be carefully examined.

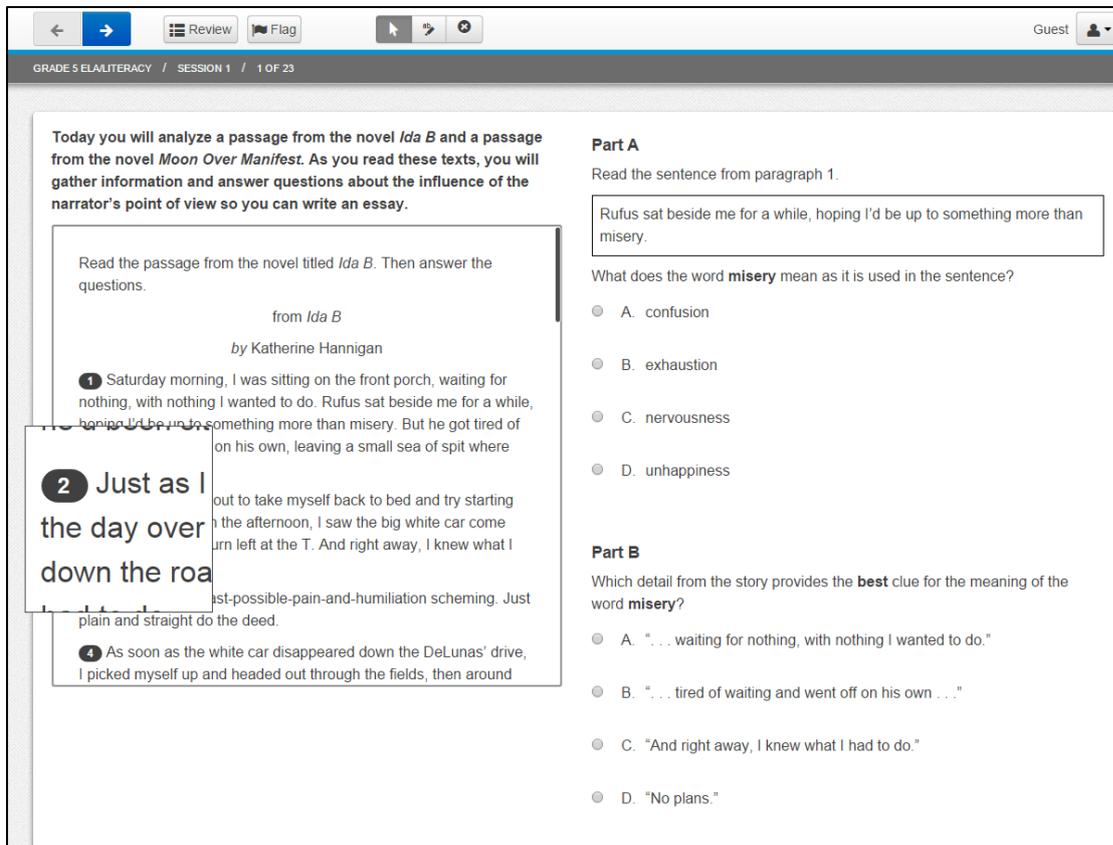


Figure 55. A square-shaped magnifier is applied to a paragraph in the PARCC assessment.

Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

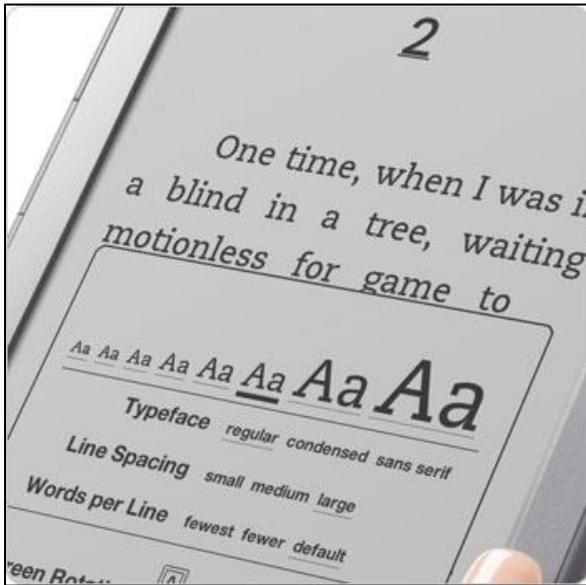


Figure 56. An example of text size selection found on an E-reader. Image taken from [http://g-ecx.images-amazon.com/images/G/01/kindle/tequila/dp/KT-df-03.V166940136 .jpg](http://g-ecx.images-amazon.com/images/G/01/kindle/tequila/dp/KT-df-03.V166940136.jpg)

The figure displays two screenshots of a Smarter Balanced test interface. Both screenshots show the same content: a reading passage titled "Diamonds in the Sky" and a multiple-choice question. The top screenshot shows the non-expanded view, where the text is compact and the question options are visible. The bottom screenshot shows the expanded view, where the text is stretched vertically to fit the screen, and a vertical scrollbar is present on the right side of the text area.

**Top Screenshot (Non-expanded view):**

English Language Arts / Literacy | Mathematics

43006

Read the sentence from the text. Then answer the question.

"Nanodiamonds are stardust, created when ancient stars exploded long ago, **disgorging** their remaining elements into space."

Based on the context of the sentence, what is the most precise meaning of **disgorging**?

(A) scattering randomly

(B)  throwing out quickly

(C) spreading out widely

(D) casting forth violently

**Bottom Screenshot (Expanded view):**

English Language Arts / Literacy | Mathematics

Diamonds in the Sky

Stars are not the only objects that glitter in the dark night sky. Scientists have discovered that diamonds are plentiful in outer space. Some of these space diamonds are called "nanodiamonds" because they are incredibly small. A nanodiamond is millions of times smaller than a grain of sugar—more or less the size of a strand of DNA. Nanodiamonds are stardust, created when ancient stars exploded long ago, disgorging their remaining elements into space. Other space diamonds are huge—the size of whole planets—while some may exist in liquid or frozen form. Scientists even suggest that planets in our own solar system may have oceans filled with chunks of frozen diamond "ice."

Diamonds are so common throughout the universe because they are a pure form of one of the universe's most common elements: carbon. Diamonds have a number of amazing properties: they are extremely hard and transparent, and can withstand radioactivity, corrosive acids, and other powerful forces. Diamonds conduct electricity more readily than copper, and are also the best natural conductor of heat that we know of—which is why diamonds feel cool to the touch. Like a prism, diamonds produce rainbows from white light. The melting point of a diamond, 7,362 degrees Fahrenheit, is higher than that of any other known substance.

Graphite and diamonds share the same chemistry—both are carbon. The difference lies in the arrangement of the carbon atoms, known as their "molecular structure." Extreme forces are required to transform dark, soft graphite—the stuff used in pencil lead—into hard, brilliant diamonds. A diamond is formed when carbon is exposed to immense pressure and extreme heat—conditions found hundreds of miles below the surface of the Earth, where most natural diamonds are formed. The heat and pressure squeeze the carbon atoms into a dense, crystalline structure. In the comic books, Superman could create a diamond by simply squeezing carbon in his bare hands, but it normally takes billions of years for carbon to become a diamond.

In space, diamonds are born more quickly. Scientists believe space diamonds often crystallize in no more than a millionth of a millionth of a second, when dust grains containing carbon smash together at extremely high speeds. Another hypothesis for how space diamonds are formed involves the shock waves released by an exploding star, which cook and compress carbon dust until it becomes a diamond.

In the 1990s, geologists discovered microscopic diamonds embedded in meteorites that had fallen out of the sky, some with the same chemistry as natural diamonds found on earth. Scientists believe these diamonds were created when meteorites collided with asteroids in our solar system. Other diamonds found inside meteorites, however, contain a mixture of xenon gas found only in outer space. These diamonds are useful to scientists because they provide clues about the composition of stars and the history of the universe.

A rare form of diamond found in Brazil and the Central African Republic—called a "carbonado diamond" because of its black color—also appears to have extraterrestrial origins. The hydrogen found in these diamonds indicates that they were probably formed in hydrogen-rich interstellar space. Scientists believe that these black diamonds were once the size of asteroids before they exploded upon impact with the Earth's surface.

Astronomers studying Uranus and Neptune think that diamond icebergs may drift in sparkling diamond oceans on these carbon-rich planets. While this sounds incredible, scientists have discovered that, given the right conditions, it is possible to liquify a diamond. To test this, nuclear scientists used lasers to recreate the extremely high heat and pressure of Uranus and Neptune. Using a normal diamond, they heated it to a temperature of 50,000 degrees and applied pressure equal to 11 million times the pressure on Earth. Under these conditions, the diamond first melted, then froze into icy chunks. In this way, scientists proved that diamonds can melt, freeze, and behave like water.

Scientists have even discovered a diamond planet in our galaxy, 4,000 light years from earth and about five times larger than our own planet. It is heavier than Jupiter, the largest planet in our solar system. Astronomers believe that this diamond planet formed when the carbon core of a massive star collapsed and became a diamond under intense gravitational pressure.

Besides being beautiful to contemplate, space diamonds teach us important lessons about natural processes going on in the universe, and suggest new ways that diamonds can be created

Figure 57. Sample test from Smarter Balanced illustrating the non-expanded (top) and expanded (bottom) reading view. Screen shots taken from

<http://www.smarterbalanced.org/sample-items-and-performance-tasks/>

### 5.4.4 Flag Tool

The flag tool serves as a reminder for students to return to a question or item. This is likened to placing a ‘Post-it’ note on a page when reading on paper. Online assessment programs such as PARCC and Smarter Balanced have included the flag tool in their tests (see Figures 58 and 59).

43016

The following paragraph is an excerpt from a student's report about plant life in the southern United States. After reading the paragraph, you will identify details that are unnecessary and explain why they should be removed.

**The Invasion of Kudzu**

In 1876, Philadelphia, Pennsylvania, hosted the Centennial Exposition in honor of the country's 100th birthday. The Japanese constructed one of the most popular exhibits, primarily due to an amazing vine called *kudzu*. For centuries, the Japanese used the pea vine for many purposes, including medicine, but what attracted the Americans the most were the sweet-smelling purple flowers that covered the vine. The warm, moist climate of the southern states—from Georgia to Florida and north to the Carolinas—was the most hospitable part of the US for the vine. Temperatures in Georgia can reach into the 90s, and the humidity is often above 90%. Residents all over the southeast began planting kudzu. The vine's success caught the attention of many, including Channing Cope who promoted its use for erosion control and animal feed, especially during the Great Depression. Because of the depression, many homes were left abandoned, so there was not anyone to care for them and keep the plants properly groomed. No one predicted, however, that the vine would end up taking over. The vines slowly engulfed pine forests, telephone poles, and even houses, leading to the destruction of native plant life. Pines are not the only trees in the South, however. There are about 250 species of trees in Georgia alone. As kudzu out-competed the local plants, it deprived them of nutrients and, especially, sunlight. Kudzu now covers over 7 million acres of land, and it continues to expand at the rate of 150,000 acres annually. That is almost one foot per day! Attempts to kill it have proven difficult, as it is immune to most herbicides; thus, kudzu continues to smother the southern states. Researchers continue to search for a solution to "the vine that ate the South," but the answer is nowhere in sight.

In the space below, identify the sentences from the paragraph that are unnecessary, and briefly explain why each one should be removed.

Figure 58. An example item from Smarter Balanced; the flag icon is placed on the top right corner above the instructions. Screen shot taken from <http://www.smarterbalanced.org/sample-items-and-performance-tasks/>

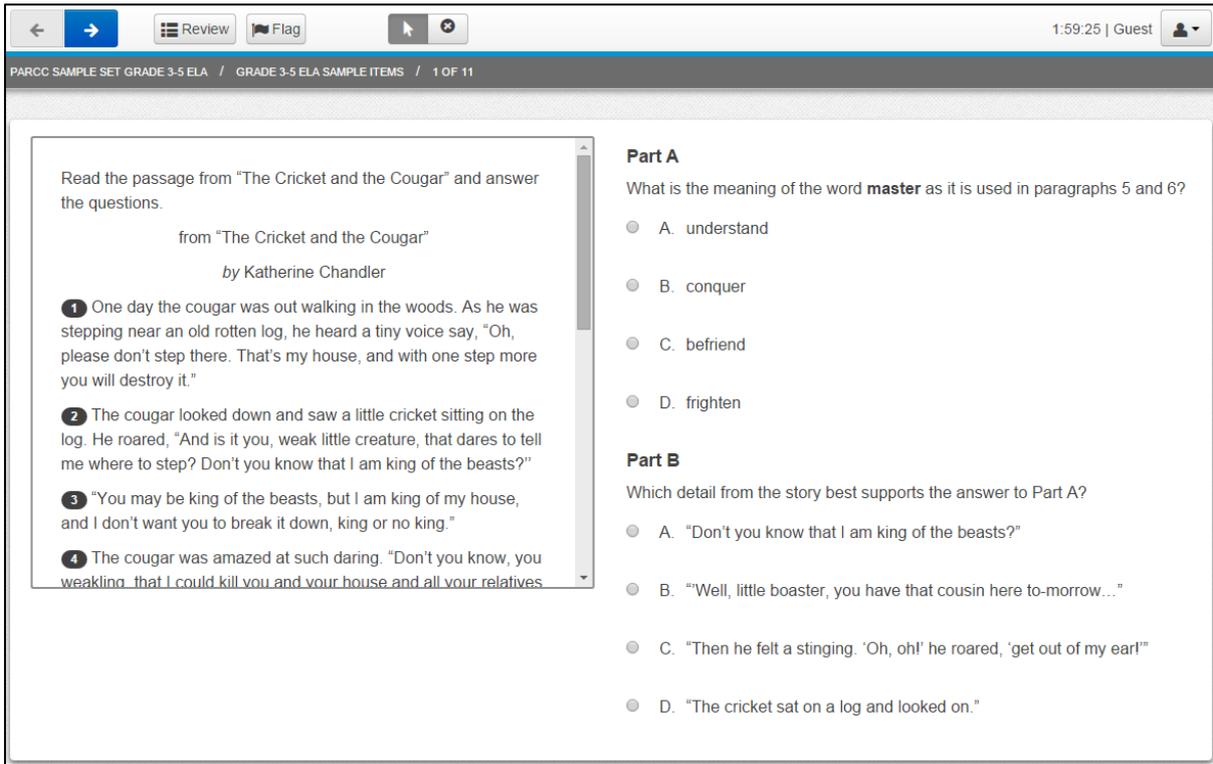


Figure 59. An example item from PARCC; the flag icon is placed on the top row, next to the navigation buttons and review tab. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

### 5.4.5 Review Tool

The review table presents a list of questions with indicators of whether the question has been viewed, answered or flagged. Generally, readers are able to go directly to specific questions by clicking on the item in the review table. The review table is typically presented at the end of the assessment (i.e., after all passages and items are viewed). To ensure that students have answered all questions from a passage before moving to the next passage, the review table could also be placed after each passage instead of at the end of the assessment or test. The ability to amend answers on the review page(s) would also help reduce the time spent on moving back and forth between pages of a test. Alternatively, students can click on the corresponding 'View' button to go directly to a specific item. Figure 60 presents an example of a review table from the PARCC assessment.

The screenshot shows a test review interface for 'Section 1'. At the top, there is a 'Return to Question' button and a 'Guest' user indicator. Below the navigation bar, the section title 'Section 1' is displayed with a right-pointing arrow. To the right of the section title, there are two summary items: 'INCOMPLETE 23 Not Answered' (with a red exclamation mark icon) and 'FLAGGED 1 Question' (with a red flag icon). Below these are three filter buttons: 'All', 'Not Answered', and 'Flagged'. A blue 'End Section' button with a right-pointing arrow is located in the top right corner of the table area.

QUESTION	STATUS	FLAGGED	
Question 1	NOT ANSWERED		View >
Question 2	NOT ANSWERED		View >
Question 3	NOT ANSWERED		View >
Question 4	NOT ANSWERED		View >
Question 5	NOT ANSWERED		View >
Question 6	NOT ANSWERED	FLAGGED	View >
Question 7	NOT ANSWERED		View >
Question 8	NOT VIEWED		View >
Question 9	NOT VIEWED		View >
Question 10	NOT VIEWED		View >
Question 11	NOT VIEWED		View >
Question 12	NOT VIEWED		View >
Question 13	NOT VIEWED		View >

Figure 60. A review table with indicators of whether items have been answered, flagged or viewed. To go directly to a specific question, students can click on the corresponding 'View' button on the far right column. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

#### 5.4.6 Answer Eliminator and Selection Indicator

A tool that may help students with answering multiple choice questions (MCQ) is the answer eliminator. To simulate reading on paper, students will be able to eliminate or cross out answers that they think or know are incorrect. An example of the answer eliminator can be seen in the PARCC reading assessments (Figure 61). In addition, the answer option that a student has selected could appear in a different colour, be highlighted or presented with a border or box (see Figure 62) so that it is clear and visible to the student. This tool, also known as a selection indicator, is only useful for certain types of questions such as multiple choice items, but not constructed response items or drag-and-drop items.

GRADE 3 ELA/LITERACY / SECTION 1 / 8 OF 18 Answer Eliminator

Today you will do some research on animals and their natural environments. First, you will read an article about wolves in Yellowstone National Park. Then you will read an article titled “The Missing Lynx.” As you read these sources, you will gather information and answer questions about animals and their environments so you can write a response.

Read the article “A Howling Success.” Then answer the questions.

A Howling Success  
by Gerry Bishop

**1** In Yellowstone National Park, a gray wolf sends its eerie call into the wild. You might say that it’s howling for joy.

**2** When you look at this photo and think about wolves, what words come to mind? Wild? Scary? Awesome?



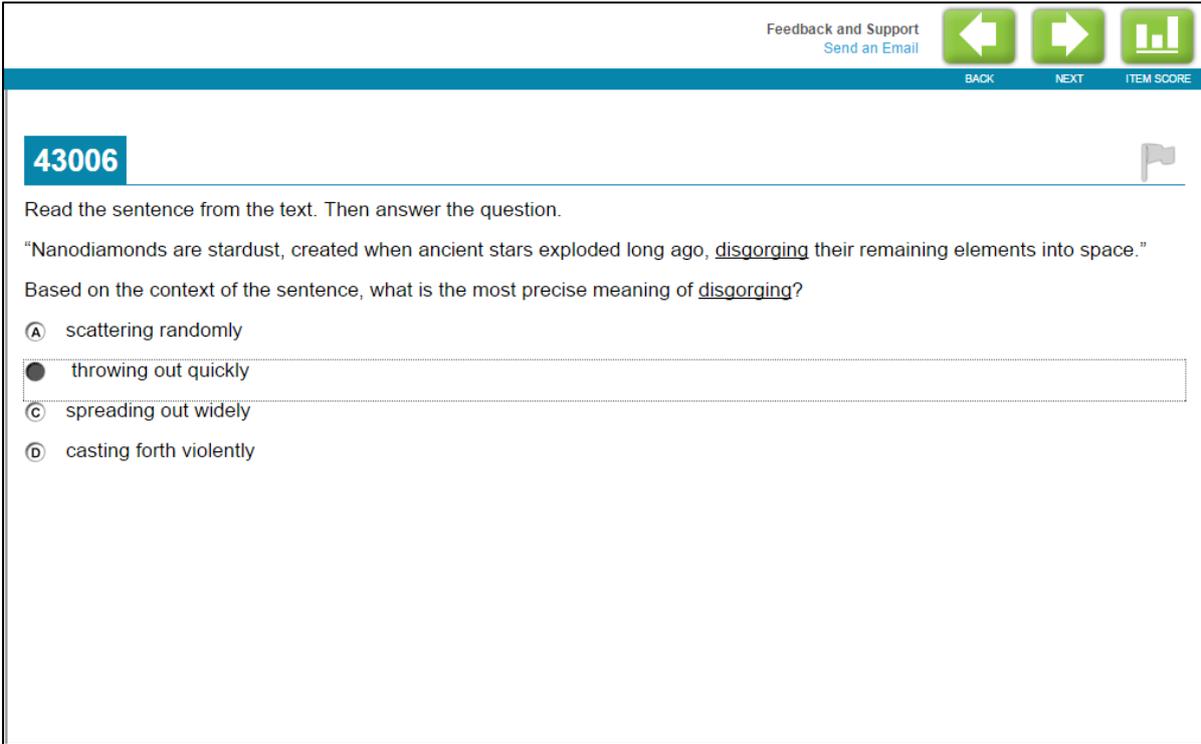
**Part A**  
What is the meaning of the word **thriving** as it is used in paragraph 14 of the article?

- A. ~~thriving~~
- B. hunting
- C. doing well
- D. having fun

**Part B**  
Which phrase from the article helps the reader understand the meaning of the word **thriving**?

- A. “... scattered among 11 packs ...” (paragraph 14)
- B. “Elsewhere in the west ...” (paragraph 14)
- C. ~~... made an amazing comeback ...” (paragraph 14)~~
- D. ~~... are the animals truly safe ...” (paragraph 15)~~

Figure 61. The answer eliminator in PARCC is displayed as a red cross over one or more MCQ response options. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>



The screenshot shows a digital assessment interface. At the top right, there is a navigation bar with the text "Feedback and Support" and "Send an Email" next to a blue link. To the right of this are three green icons: a left-pointing arrow labeled "BACK", a right-pointing arrow labeled "NEXT", and a bar chart icon labeled "ITEM SCORE". Below this bar is a blue header with the question ID "43006" in white text on the left and a flag icon on the right. The main content area contains the following text: "Read the sentence from the text. Then answer the question." followed by a quote: "Nanodiamonds are stardust, created when ancient stars exploded long ago, disgorging their remaining elements into space." Below the quote is the question: "Based on the context of the sentence, what is the most precise meaning of disgorging?" There are four multiple-choice options: (A) scattering randomly, (B) throwing out quickly, (C) spreading out widely, and (D) casting forth violently. Option (B) is selected, indicated by a black radio button and a light blue shaded rectangular box around the text.

Feedback and Support  
Send an Email

BACK NEXT ITEM SCORE

**43006**

Read the sentence from the text. Then answer the question.

"Nanodiamonds are stardust, created when ancient stars exploded long ago, disgorging their remaining elements into space."

Based on the context of the sentence, what is the most precise meaning of disgorging?

(A) scattering randomly

(B) throwing out quickly

(C) spreading out widely

(D) casting forth violently

Figure 62. A screen shot of a question section from Smarter Balanced assessment. The chosen answer (option B) is shaded and presented with a border/box around it. Screen shot taken from <http://www.smarterbalanced.org/sample-items-and-performance-tasks/>

#### 5.4.7 Notepad

To allow students to make notes, a notepad could be included, such as the one found in the PARCC assessment (Figure 63). This notepad appeared as a separate pop-up window, and could be close and re-open at any time. It had an auto-save feature; all written notes were saved automatically without having to click on any 'Save' button. This auto-save notepad will be very useful to include as it allows students to make notes, just as they would when reading on paper.

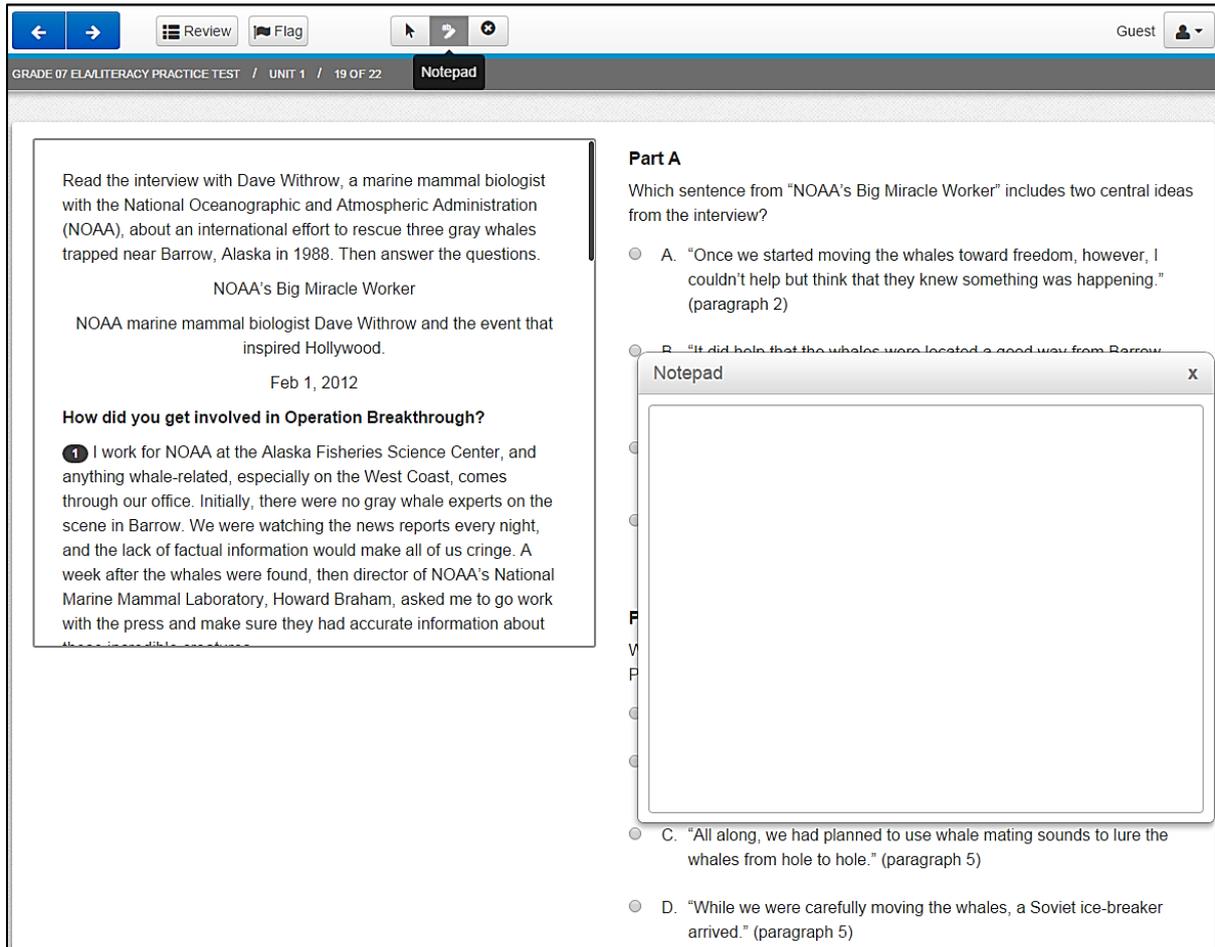


Figure 63. A pop-up notepad featured in a PARCC online assessment. Screen shot taken from <http://parcc.pearson.com/practice-tests/english/>

## 5.5 Summary of recommendations

Recommendations for NAPLAN online are summarised in Table 2. A list of factors and questions to consider is also included in Table 2. Many of these recommendations were adapted from other online assessment tools, while some recommendations were generated using ideas from other online tools or web pages. All recommendations should be tested and trialled with children from Years 3, 5, 7 and 9 using NAPLAN reading materials, and across different devices such as desktop computers and tablets.

Table 2

*Summary of recommendations for NAPLAN online*

<b>Feature</b>	<b>Recommendations</b>	<b>Things to consider</b>
Display format	Text only passages, some with add-on graphics	<ul style="list-style-type: none"> <li>• Length of text</li> <li>• Position of add-on or supporting graphics especially if thumbnails and hyperlinks are used</li> </ul>
	Text with supporting graphic	
	Compare and contrast texts	<ul style="list-style-type: none"> <li>• The ease of reading text displayed side by side or in page tabs               <ul style="list-style-type: none"> <li>○ Are students able to read and navigate between texts and items presented on the screen?</li> </ul> </li> <li>• Placement of graphics within compare-and-contrast texts</li> <li>• The feasibility of bifocal display for presenting compare-and-contrast texts, especially on smaller screen sizes, and with touch-screen interface.               <ul style="list-style-type: none"> <li>○ What are the benefits of using bifocal display to present multiple passages or compare-and-contrast texts?</li> <li>○ Is this method feasible on smaller devices?</li> <li>○ Are there any differences in performance or useability when bifocal display is presented with a touch-screen interface versus using a mouse?</li> </ul> </li> </ul>
	Multiple media content	<ul style="list-style-type: none"> <li>• Placement of multimedia content (e.g., video clip, audio recording).               <ul style="list-style-type: none"> <li>○ Are pop-up hyperlinks useful for displaying multiple media contents? Do they obstruct the view of the items/questions displayed on</li> </ul> </li> </ul>

Feature	Recommendations	Things to consider
		<p>screen?</p> <ul style="list-style-type: none"> <li>• Accessibility of media content                             <ul style="list-style-type: none"> <li>○ Is Internet access required to access the media content?</li> </ul> </li> </ul>
Item types	Drop-down list	<ul style="list-style-type: none"> <li>• Are the drop-down list/menu and drag-and drop items easy to use across all devices?</li> </ul>
	Drag-and-drop items	<ul style="list-style-type: none"> <li>• Are these methods suitable to use when there are many options in the drop-down menu or many items to drag?</li> </ul>
	Click-and-select or Select-and-highlight	<ul style="list-style-type: none"> <li>• Can these methods be applied to different answer options (e.g., words, paragraphs, graphics and etc.)?</li> <li>• Inclusion of selection indicator to make students' selection (i.e., selected answer) more salient.</li> <li>• Colour of highlight</li> </ul>
	Hot spot items	<ul style="list-style-type: none"> <li>• How could hot spot items be used in an online reading assessment context?</li> </ul>
Presentation layout	Progress bar	<ul style="list-style-type: none"> <li>• Position of the progress bar (e.g., top or bottom of page) in relation to text passages and items.                             <ul style="list-style-type: none"> <li>○ The position of the progress bar would be influenced by navigation method (i.e., paging vs. scrolling).</li> </ul> </li> <li>• Type of progress bar (e.g., percentage bar vs. item panel vs. pop-up console)</li> <li>• Features and information included in the progress bar:                             <ul style="list-style-type: none"> <li>○ Can students review and access a particular item directly?</li> <li>○ Would students be able to tell which items have been answered or viewed?</li> </ul> </li> </ul>

Feature	Recommendations	Things to consider
	Scrolling and paging	<ul style="list-style-type: none"> <li>• Navigation methods: whole page scrolling vs. section scrolling vs. paging                             <ul style="list-style-type: none"> <li>○ This would depend on the length and number of texts/passage, number of items (questions) displayed on each page, and device type.</li> </ul> </li> </ul>
	Innovative options: <ul style="list-style-type: none"> <li>• Dynamic text</li> <li>• Interactive text</li> </ul>	<ul style="list-style-type: none"> <li>• To our knowledge, these options have not been used before in an online assessment. Therefore, careful consideration and further testing will need to be conducted to examine the feasibility and usefulness of utilising dynamic text and interactive text concepts in reading assessments.</li> </ul>
Reading tools	Highlighter	<ul style="list-style-type: none"> <li>• Option to select highlight colour</li> <li>• Size of line reader                             <ul style="list-style-type: none"> <li>○ Can the size be adjusted?</li> </ul> </li> <li>• Do students find these tools useful?                             <ul style="list-style-type: none"> <li>○ Is the tool (e.g., a notepad) needed in a reading task?</li> <li>○ Is the tool easy to operate (e.g., option to activate and de-activate the magnifier or answer eliminator)?</li> </ul> </li> <li>• Placement of tool                             <ul style="list-style-type: none"> <li>○ Is the tool icon easy to locate on the screen?</li> </ul> </li> </ul>
	Line reader	
	Magnifying tools: <ul style="list-style-type: none"> <li>• Magnifier (to simulate using a magnifying glass)</li> <li>• Different text size options (e.g., small, medium, large)</li> <li>• Expanded passage section</li> </ul>	
	Flag tool	
	Review tool	
	Answer eliminator and Selection indicator	
	Notepad	

## 6 Cognitive Interviews

---

In the context of generating assessment tasks or items that reflect an identified construct, it is necessary to ensure that the items do in fact target that construct. There are several processes which can be implemented to provide supporting evidence of this, prior to piloting of the items and analysis of test data. In this section, we consider in particular, ‘think-aloud’ methods such as cognitive interviews that are used in the early stages of test development to identify design issues in large scale assessment (Johnstone, Bottsford-Miller, & Thompson, 2006). The methods enable investigation of the mental, or cognitive, processes used by examinees when completing a task and responding to test items (R. P. Dolan, Goodman, Strain-Seymour, Adams, & Sethuraman, 2011). Consequently, they can detect issues related to unclear constructs, instructions and items (e.g. Items that are too difficult, repeatedly answered incorrectly, left incomplete or omitted) as well as misunderstood language, and illegible text or graphics that result in incorrect responses or bias (Zucker, 2004). Various organisations have conducted cognitive interviews (also called cognitive laboratories [labs] and/or field tests) (Griffin & Care, 2015; PARCC, 2014a; SBAC, 2014a; Zucker, Sassman, & Case, 2004) for large scale assessment test development and/or to compare paper based assessments with computer based assessments. As well as cognitive interviews, two other processes - concept checks and logistic checks - are briefly outlined.

Cognitive interviews are used primarily to answer the following questions:

- What are the students’ and teachers’ perceptions of and reactions to new items?
- What are the cognitive processes employed by students while responding to these items, and are these the processes aligned with the intended constructs?
- How do the cognitive processes employed by students when responding to innovative items compare to the processes used while responding to corresponding traditional items?
- Are there usability issues that impact on how students interact with innovative items and to what degree is usability issues a potential source of construct-irrelevant variance?

There are two associated processes which can usefully be undertaken to complement cognitive interviews with students. These comprise engagement with teachers of the targeted student population ('concept check'), prior to presentation of new tasks in their beta versions to students; and engagement with test administrators and IT coordinators in schools post piloting of tasks with students ('logistics check'). These processes are described here within the sequence in which they would normally occur.

## 6.1 Concept check

In this discussion, 'concept check' refers to a process of checking whether assessment tasks or items appear to be accessing the targeted skills. The process is akin to the establishment of face validity. Do the tasks look as though they sample the skills from the viewpoint of those who will be responsible for assessing and teaching the skills? The approach could be beneficial in this current project on reading assessments on an online platform.

### 6.1.1 Task Concept Check Example

The international ATC21s project aimed to establish new forms of assessment for 21<sup>st</sup> century skills (e.g. collaboration and digital literacy) to support teachers in assessment and learning by providing accurate information about students' skill levels. The assessments were developed for an online platform, and due to the innovative nature of the task design and presentation, task concept checks were conducted with teachers to ascertain whether the assessment task concepts appeared reasonable to teachers in schools.

Researchers took alpha versions of tasks out to schools for the purpose of task concept checks. The sessions involved describing the skill set of interest to the teachers, and demonstrating the tasks being designed for student use. The teachers were asked whether the assessment tasks would:

- be appropriate for students in terms of prerequisite knowledge
- be appropriate for students in terms of socio-cultural context
- require different amounts of time for students
- differentiate between students
- engage students.

In addition, the following questions were asked of teachers:

- What skills or capabilities do you think the tasks are targeting?
- Considering the capabilities of your students, are there any questions or activities that appear too far above or too far below the general range of their capabilities?
- What other comments do you have about the task?

Following this exercise, modifications were made to the tasks, as well as consideration being given to the appropriateness of the age range targeted. The learning from this process was not to go too far with development of innovative tasks before engaging in this task concept check process. Face validation of the concepts with teachers is a critical step.

## 6.2 Cognitive Interviews

Typically, a cognitive interview consists of an interviewer administering items to a student in one-on-one sessions. There are two types of cognitive interviews. The first is the concurrent oral report (i.e. participants share thoughts with the interviewer as they complete a task) while the other is the retrospective oral report (i.e. participants share thoughts with the interviewer immediately after completing the task) (Zucker, 2004). Retrospective oral reports are typically used for the lower-grade-level students (K–3). Research has found that both are useful. The primary issue in the concurrent type lies in the cognitive load on the student when required to verbalise thought processes while engaging with the task, that is, the multi-tasking. This can be particularly an issue with younger children. The disadvantage of the retrospective report is the loss of information about processes that can occur as a result of the time delay, and is dependent on the skills and knowledge of the interviewer to an even greater extent than in the concurrent implementation. Notwithstanding, the retrospective oral report can help clarify the processes and responses, and provide additional information about test performance (Dolan et al., 2011). Participants need to be representative of the population that the assessment will be administered to subsequently (Johnstone et al., 2006). There is no point in having ‘expert groups’ simulate the test-taking process, since their familiarity with the concepts underlying the tasks prevents their bringing a naïve eye to the process.

Data collection and analysis processes are time-intensive and therefore hinder the use of large samples (Dolan et al., 2011). At the same time, the nature of the activity negates the need for large samples. The process basically encapsulates an audit of a respondent's cognitive and affective processes and responses in the interaction with the task.

### 6.2.1 Cognitive Interviews with Students – Concurrent Report

For cognitive interviews, a protocol is carefully designed to guide interviewers in eliciting and recording student responses accurately. The purpose of the interviews is to check the assessment tasks and ensure that they had the capacity to elicit evidence of the required skillsets from students (Griffin & Care, 2015). When starting the interview, participants are given a brief introduction to the purposes of the session (i.e. that it is on the test items, how the respondent functioned and reacted, as opposed to the correctness or quality of their answers) and an overview of the test environment and interview process. In addition to the protocol, a cognitive interview booklet is developed to guide the interviewer's record of observed behavioural data according to the protocol and can include checklists of the possible behaviours that the student could demonstrate while responding to an item. Observers are provided with different lists of behaviours to audit depending on which skillset is being observed. An example of some guidelines for completing checklists is presented below, and a checklist component of a cognitive interview is shown in Table 3.

#### Guidelines for completing checklist:

- Record the date, time, location and duration of the cog lab in the 'Notes' section. Also the number of participating students (by first name) alongside their logins (e.g. [specific login]) in the designated section.
- Record the nature and duration of any technical or connectivity issues you encounter, as well as the approach you (or the student) took to fixing them.
- If possible, record the upload and download statistics for each student for the duration of their interaction with the task.
- Where applicable, select Yes or No against each of the student actions specified by colouring the appropriate box green (e.g. Y).

- Ensure that any comments you make about a student’s interaction with a particular page of the task are inserted in the appropriate comment box and identified using the student’s login.
- Comments or observations of a general nature can be appended to the end of the template document.

Table 3

*An example of component checklist for cognitive interview*

<b>Actions Taken</b>	Login ID 1	Login ID 2	Login ID 3	Login ID 4
Click links to definitions of Mood and Meaning	Y <input type="checkbox"/> N <input type="checkbox"/>			
Use of radio buttons	Y <input type="checkbox"/> N <input type="checkbox"/>			
Copy / paste poem	Y <input type="checkbox"/> N <input type="checkbox"/>			
Move poem	Y <input type="checkbox"/> N <input type="checkbox"/>			
Use mapping tools (hand, pen / arrow, text / label, bubble)	Y <input type="checkbox"/> N <input type="checkbox"/>			
Use recycling bin (if required)	Y <input type="checkbox"/> N <input type="checkbox"/>			
Add ideas	Y <input type="checkbox"/> N <input type="checkbox"/>			
Connect ideas	Y <input type="checkbox"/> N <input type="checkbox"/>			
<b>Teach Aid</b>	Y <input type="checkbox"/> N <input type="checkbox"/>			
<b>Usability Issues</b>	Y <input type="checkbox"/> N <input type="checkbox"/>			

During the concurrent oral report, interviewers are expected to prompt students strategically for simultaneous commentary, to minimise distractions during the completion of the task, or inadvertently leading a student to a particular problem solving approach (Ericsson & Simon, 1993). Interviewers are trained to avoid influencing the student’s response through verbal or visual cues, such as offering assistance or indicating whether the student’s response was correct. They are expected to use neutral prompts in cases where the student stop verbalizing. A typical question would be “What are you thinking now?” Different sequences of questions may be asked when students are observed to be taking

both expected and unexpected paths through task solution. In order to obtain the richest information possible from the students in the ATC21s project, interviewers used the following points as a guide during the test administration:

- Sit near the student but not in their personal space
- If the student is silent for more than a few seconds, prompt with “Keep talking” or “What’s happening?”
- Be sure that the student is actually entering her/his responses and/or taking action, not just talking, in which case prompt with “Please enter your response”;
- If the student is having trouble entering responses and talking simultaneously, have the student talk first and then enter her/his responses.
- Be attentive with body language by head-nodding and occasional, non-evaluative comments in response to students.
- Do NOT tell the student if s/he is getting an answer right or wrong.
- Do NOT tell the student if s/he is doing well/poorly on the activity.
- Do NOT show bias for certain tasks, items or item formats (e.g., do not say anything like, “This is not a very good problem” or “Problems like these don’t test many skills”).

For this stage of item/test development, it is important to identify the degree to which students of different abilities interact with the task. Students are asked to complete each item, thinking aloud as they work through the task, while the interviewer records observations on the student’s behaviour in the booklet. Interviewers are expected to take detailed notes of the participant’s interactions with the items. Where possible, they need to record which steps were taken and the order these steps took place. Interviewers also need to be aware of the different user interface elements (e.g., drag-and-drop tool, highlighter tool) listed for each item, and the degree to which a student could interact with the elements as intended, detailing each student-element interaction when appropriate. Observations can include noting if the student:

- Shows engagement with the task
- Shows frustration with the task
- Shows signs of enjoyment in the task
- Shares any reasoning for their actions
- Takes a systematic approach to solving the problem
- Completes the tasks
- Were there Usability Issues (specify)?

Questions that can be asked include:

- Did you know how to answer the question?
- Were you clear what you were being asked?
- What features of the item made it easy to use or difficult to use?
- Have you ever seen an item like this?
- Did you like working on this item? Why or why not?
- What would make this item easier to use?
- How does this item compare to items that you typically see on a test?

### 6.2.2 Cognitive Interviews with Students – Retrospective Report

After the students report that they have completed an item, the interviewer may guide a retrospective oral report from the student through a series of probing questions related to the students' experiences and/or interactions with the specific features of the item (e.g. student is encouraged to elaborate on challenges of aspects of an item in terms of content or usability). Therefore, the interview booklet may include questions to be asked of the student immediately after he or she completes the item ("How did you get your answer to this question? What made you do that?") in order to gather data about the thought process behind the students' response. After the retrospective interview for each item or question, the participants are instructed to proceed to the next question.

At the end of completing all the items or questions, students can be probed on more general questions about the assessment tasks and the skills tested.

Example questions from other assessment projects include:

- Have you taken such online tests before?
- How difficult was the test?
- What skills do you think were being assessed? Were the items “fair” assessments of your skills?
- Did you have enough time to finish the test?
- Did you understand all the directions on the test?
- Were the tasks enjoyable? Did you learn anything from the experience?
- What technological or technical difficulties or challenges did you face?
- Was it easy to move back and forth between passages and/or stories?
- Did you use the pop-up hyperlinks for other media? Were they easy to navigate?
- Was it easy to type the answers?
- Were the drop-down list / drag-and-drop items easy to use?
- Did you use any of the reading tools provided on the online reading assessment task (e.g. highlighter, line reader, magnifying tools, flag or review tool, notepad etc)?
- Did you find these tools easy to locate on the screen?
- Did you find these tools easy to operate during the assessment? E.g. Were you able to adjust the size of the passage using the magnifying tool?

In addition, based on observer judgment, problematic questions or issues for students can be reviewed with the students by going back into the task environment to discuss points at which observers did not understand what the student was doing or at screens where the student/s had seemed to struggle. Prompts can include:

- “What do you think this question is asking you to do?”
- “How could we make any wording of the task clearer?”

### 6.3 Data Analysis of the Cognitive Interviews

To verify the objectivity of written observations, a computer-based video capture or audio recording of each one-on-one session can be made for later review and verification by

assessment specialists. This can control for bias that may exist when the interviewer omits information or includes their interpretation of information shared by the interviewee. Combining the structured interviews with video capture and/or audio recording assures a more objective data collection. When all observations and interviews are completed, the observations are coded by an independent researcher. Using the multiple data sources – interviewer notes and checklists, and video/audio capture, the coder identifies instances of interest in each session. These instances relate to questions such as the degree to which the cognitive processes captured reflect the intention of the test designers, and the degree to which the actual assessment environment impinges on student performance.

## 6.4 Logistics checks

At the point at which assessment tasks can be piloted at class level, additional information concerning the logistics of test administration, student engagement at group level, and facilities and infrastructure is required.

Teacher test administrators and coordinators should provide feedback about the class experiences with the tasks, and the online assessment system, and about student engagement. IT coordinators should provide feedback about the ease of the infrastructure set up, and system load and connectivity processes. The feedback can be provided either through interview or survey. In addition, the observation of sessions by researchers ensures that the teacher and IT coordinator feedback can be interpreted both in the light of the actual occurrence as well as in the light of the assessment intent. Sample protocols for these purposes from the ATC21s project are shown below (from the IT and Classroom Management Checklist for Pilots):

- What were the issues that the IT support staff needed to deal with in order to enable access to the tasks, prior to the session?
- Were there any events/issues, in terms of hardware, software, or internet access that required IT support staff to intervene during the session? If so, what were these?
- How were they were resolved?
- How long did this take?

- What were the repercussions for management of the class?
- What were the repercussions for individual students?
- How long from the time students were at their desks, did it take for all students to be logged into the task site?
- How much time were students able to spend on the actual assessment questions?
- Were there any difficulties for specific students in engaging with the online environment? If so, what were these?
- Were these difficulties associated with individual student characteristics such as physical or cognitive disabilities or dysfunctions?
- Were these difficulties associated with technical issues? For example, slow connection speeds, inappropriate screen sizes, error messages?
- Did the teacher experience any difficulties in terms of communicating to students what would be required of them?
- Did the teacher experience any difficulties in terms of classroom management during the session?
- To what extent did students completing tasks at different times prove problematic in the classroom?
- What strategies were implemented to facilitate the task/student allocation?
- Is it viable to run such a session with one staff member only in the classroom?
- What advice, comments or suggestions, if any, did the teacher have for the administration of the session?
- For future classroom administration of tasks, what are the most important strategies or structures that needed to be put in place to enhance the smooth running?

## 6.5 Summary of the cognitive interview section

The process described of concept check, cognitive interview and logistics check may not be required to its full extent when the construct being measured is familiar to the stakeholders – the task developers, the education system, school, teacher, and student; or where the task environment is familiar to the test takers. The processes are designed to check that the intended skills are being captured, and that this capture is not biased by the assessment

environment. Accordingly, the process is necessary only where the targeted skills are unfamiliar or where the assessment environment is unfamiliar.

By its very nature, the process will require quite different response and checklist templates – customised to the skills, the assessment tasks, and the assessment environment.

## 6.6 Recommendations for the use of cognitive interviews

It is recommended that for the purposes of NAPLAN online readability and presentation layout decisions, the cognitive interview process without the teacher concept check would be sufficient.

In regards to presentation layout, there are three main issues:

1. Location of stimulus text and item sections (i.e., question and response options)
2. Location of stimulus text and item sections when comparisons across texts are required
3. Location of multiple media contents and links (e.g., video clip, audio recording)

The first issue is relatively easy to check through the cognitive interview process. Using the different NAPLAN online presentation layouts recommended in this report, prototypes of NAPLAN tests can be presented in PDF format for students to react to, with linked observation checklist tools for researchers/teachers to complete. Once item placement has been determined, the finer details of navigation methods (i.e., scrolling or paging) and additional reading tools can be considered and determined using feedback from student retrospective reports.

The second issue can similarly be managed. Neither the first nor second issue therefore require a great deal of infrastructure in terms of programming or alpha versions.

The third issue presents a different problem. The access to multiple media contents – either stimulus materials or response options – requires that students navigate around the screen (e.g., click on video, click and select or drag-and-drop item responses). In some cases, students may need to switch or toggle between contents on the page. In order for students to experience this in a way that can be captured through cognitive interview processes, a “live” version of such an environment needs to be provided. Such provision assumes

programming and expense. It is suggested therefore, that different examples of such live environments are used for the purpose of the cognitive interviews; these examples can be sourced rather than creating new ones. There are examples of ICT digital literacy tasks that could be used to facilitate this process. Apart from the provision of the online environment, the actual cognitive interview process remains as with the first two issues, with the templates for observation to be populated with relevant items and questions for the researchers/teachers to use.

## 7 Summary

---

The use of online assessment tools and programs in the field of education has steadily increased over the past decade. More and more paper-based tests are being replaced by computer-based test and online assessments, and online assessments of reading and mathematics are particularly popular. However, there is limited research and information about the transferability of paper-based tests to an online environment. In addition, research findings on factors related to the readability of onscreen text are mixed. Much of the research on readability of onscreen text was conducted on older students (e.g., undergraduate university students) and/or adults, with some studies examining digital reading or web page reading rather than reading comprehension per se. Nonetheless, the existing readability research has focused on factors such as typeface characteristics, text format, stimuli and item presentation layout. Therefore, these factors should be carefully considered when designing an online assessment task or tool.

At present, there are no specific guidelines for designing online reading assessments. Most of our recommendations for NAPLAN online were based on ideas extracted or adapted from existing online assessment tools. We also drew upon the Centre's staff expertise and adapted ideas from some the Centre's ICT projects. We have included multiple recommendations (e.g., layout options and ideas) and some cognitive interview protocols and questions for NAPLAN online reading assessments. However, these recommendations would need to be tested with children from Years 3, 5, 7, and 9 using NAPLAN reading assessment materials. Considerations for children with disabilities would also need to be taken into account when designing and testing NAPLAN online reading assessments. Additionally, given the rapid advancement in technology, features and accessibility of different devices must be taken into account when implementing online assessments on different device type (e.g., computer or tablet).

In conclusion, further research and testing is still needed to examine the optimal way of transferring paper-based test online, and to evaluate the different approaches of online reading assessment, especially when compare and contrast texts and multi-media are utilised. Field tests using cognitive interviews will provide more information and a better understanding of students' and teachers' experiences with online reading assessments.

## 8 References

---

- ACARA. (2014). NAPLAN Online. Retrieved from <http://www.nap.edu.au/online-assessment/online-assessment.html>
- ACER. (2015). Progressive Achievement Tests in Reading. Retrieved from <http://www.acer.edu.au/pat-reading>
- AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*: Washington, DC: American Educational Research Association.
- Almond, R. G., Steinberg, L. S., & Mislevy, R. J. (2002). Enhancing the Design and Delivery of Assessment Systems: A Four Process Architecture. *The Journal of Technology, Learning, and Assessment*, 1(5), 1-64.
- Baker, R. (2005). *The effects of multiple column online text on reading speed, reading comprehension and satisfaction*. (Doctor of Philosophy), Wichita State University, Wichita, Kansas.
- Bernard, M. L., Chaparro, B. S., Mills, M. M., & Halcomb, C. G. (2002). Examining children's reading performance and preference for different computer-displayed text. *Behaviour & Information Technology*, 21(2), 87-96. doi:10.1080/01449290210146737
- Bernard, M. L., Chaparro, B. S., Mills, M. M., & Halcomb, C. G. (2003). Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal of Human - Computer Studies*, 59, 823-835. doi:10.1016/S1071-5819(03)00121-6
- Bernard, M. L., Fernandez, M., & Hull, S. (2002). The effects of line length on children and adults' online reading performance. <http://usabilitynews.org/the-effects-of-line-length-on-children-and-adults-online-reading-performance/>
- Bridgeman, B., Lennon, M. L., & Jackenthal, A. (2003). Effects of Screen Size, Screen Resolution, and Display Rate on Computer-Based Test Performance. *Applied Measurement in Education*, 16(3), 191-205.
- Burnhill, P., Hartley, J., & Young, M. (1976). Tables in text. *Applied Ergonomics*, 7(1), 13-18.
- Choi, S. W., & Tinkler, T. (2002). *Evaluating comparability of paper-and-pencil and computer-based assessment in a K-12 setting*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.
- Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*, 33(5), 593-602.
- Davis, L. L., Kong, X., & McBride, Y. (2015). *Device Comparability of Tablets and Computers for Assessment Purposes*. Retrieved from Chicago, IL: [http://researchnetwork.pearson.com/wp-content/uploads/NCME\\_DeviceComparabilityofTablesComputers\\_040915.pdf](http://researchnetwork.pearson.com/wp-content/uploads/NCME_DeviceComparabilityofTablesComputers_040915.pdf)
- Davis, L. L., Strain-Seymour, E., & Gay, H. (2013). *Testing on tablets: Part II of a series of usability studies on the use of tablets for K-12 assessment programs*. Retrieved from [http://researchnetwork.pearson.com/wp-content/uploads/Testing-on-Tablets-Part-II\\_formatted.pdf](http://researchnetwork.pearson.com/wp-content/uploads/Testing-on-Tablets-Part-II_formatted.pdf)
- Dillon, A. (1992). Reading from paper versus screens: a critical review of the empirical literature. *Ergonomics*, 35(10), 1297-1326.

- Dillon, A., Richardson, J., & McKnight, C. (1990). The effects of display size and text splitting on reading lengthy text from screen. *Behaviour & Information Technology*, 9(3), 215.
- Dolan, R. P., Goodman, J., Strain-Seymour, E., Adams, J., & Sethuraman, S. (2011). *Cognitive lab evaluation of innovative items in Mathematics and English Language Arts assessment of elementary, middle, and high school students*. Retrieved from [http://images.pearsonassessments.com/images/tmrs/Cognitive Lab Evaluation of Innovative Items.pdf](http://images.pearsonassessments.com/images/tmrs/Cognitive_Lab_Evaluation_of_Innovative_Items.pdf)
- Dolan, R. P., Hall, T. E., Banerjee, M., Chun, E., & Strangman, N. (2005). Applying Principles of Universal Design to Test Delivery: The Effect of Computer-Based Read-Aloud on Test Performance of High School Students with Learning Disabilities. *Journal of Technology, Learning, and Assessment*, 3(7).
- Duchnicky, R. L., & Kolars, P. A. (1983). Readability of text scrolled on visual display terminals as a function of window size. *Human Factors*, 25(6), 683-692.
- Dyson, M. C. (2004). How physical text layout affects reading from screen. *Behaviour and Information Technology*, 23, 377-393.
- Dyson, M. C., & Kipping, G. J. (1997). The legibility of screen formats: Are three columns better than one? *Computers & Graphics*, 21(6), 703-712.  
doi:[http://dx.doi.org/10.1016/S0097-8493\(97\)00048-4](http://dx.doi.org/10.1016/S0097-8493(97)00048-4)
- Dyson, M. C., & Kipping, G. J. (1998). The effects of line length and method of movement on patterns of reading from the screen. *Visible Language*, 32(2), 150-181.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. London: The MIT Press.
- Eslami, H. (2014). The effect of syntactic simplicity and complexity on the readability of the text. *Journal of Language Teaching and Research*(5), 1185.
- Farrell, G., & Farrell, V. (2011). *Online assessment: splitting the screen to be seen*. Paper presented at the 23rd Australian Computer-Human Interaction Conference., New York, USA.
- Gil, L., Martinez, T., & Vidal-Abarca, E. (2015). Online assessment of strategic reading literacy skills. *Computers & Education*, 82, 50-59.  
doi:10.1016/j.compedu.2014.10.026
- Gregory, M., & Poulton, E. C. (1970). Even versus Uneven Right-hand Margins and the Rate of Comprehension in Reading. *Ergonomics*, 13(4), 427.
- Griffin, P., & Care, E. (Eds.). (2015). *Assessment and teaching of 21st century skills: Methods and approach*. Dordrecht: Springer.
- Hartley, J., Burnhill, P., & Fraser, S. (1974). Typographical problems of journal design. *Applied Ergonomics*, 5, 15-20. doi:10.1016/0003-6870(74)90253-1
- Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the Effect of Computer-Based Passage Presentation on Reading Test Performance. *Journal of Technology, Learning, and Assessment*, 3(4).
- Hojjati, N., & Muniandy, B. (2014). The Effects of Font Type and Spacing of Text for Online Readability and Performance. *Contemporary Educational Technology*, 5(2), 161-174.
- Huang, H.-c. (2014). Online versus Paper-Based Instruction: Comparing Two Strategy Training Modules for Improving Reading Comprehension. *RELC Journal: A Journal of Language Teaching and Research*, 45(2), 165-180.
- Hung, P. H., Huang, C. J., Yeh, Y. H., & Chang, K. L. (2010). The Development of an Online Reading Literacy Assessment. *Advanced Materials Research*, 108-111, 838-843.

- Johnstone, C. J., Bottsford-Miller, N. A., & Thompson, S. J. (2006). Using the think aloud method (cognitive labs) to evaluate test design for students with disabilities and English language learners (Technical Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Jones, B. (2012). Understanding the F-Layout in Web Design. Retrieved from Web Design website: <http://webdesign.tutsplus.com/articles/understanding-the-f-layout-in-web-design--webdesign-687>
- Josephson, S. (2008). Keeping Your Readers' Eyes on the Screen: An Eye-Tracking Study Comparing Sans Serif and Serif Typefaces. *Visual Communication Quarterly*, 15, 67-79. doi:10.1080/15551390801914595
- Kerr, M., & Symons, S. (2006). Computerized Presentation of Text: Effects on Children's Reading of Informational Material. *Reading and Writing*, 19(1), 1-19. doi:10.1007/s11145-003-8128-y
- Kim, H., & Kim, J. (2013). Reading from an LCD monitor versus paper: Teenagers' reading performance. *International Journal of Research Studies in Educational Technology*, 2(1), 1-10.
- King, L., Kong, X. J., & Bleil, B. (2011). *Does size matter? A study on the use of netbooks in K-12 assessment*. Paper presented at the Annual meeting of the American Educational Research Association, New Orleans, Louisiana.
- Klyszejko, Z., Wieczorek, A. M., Sarzynska, J., Szostek, A., Chmiel, K., Soloch, T., & Brzezicka, A. (2014). Mode of text presentation and its influence on reading efficiency: Scrolling versus pagination. *Studio Psychologica*, 56, 309-321.
- Kolers, P. A., Duchnicky, R. L., & Ferguson, D. C. (1981). Eye movement measurement of readability of CRT displays. *Hum Factors*, 23(5), 517-527.
- Kruk, R. S., & Muter, P. (1984). Reading of Continuous Text on Video Screens. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 26(3), 339-345. doi:10.1177/001872088402600309
- Ling, J., & van Schaik, P. (2007). The influence of line spacing and text alignment on visual search of web pages. *Displays*, 28(2), 60-67. doi:<http://dx.doi.org/10.1016/j.displa.2007.04.003>
- Loh, C. S., Branch, R. M., Shewanown, S., & Ali, R. (2002). *The effects of text spacing after the period on time for on-screen reading tasks*. Retrieved from [http://www.csloh.com/research/pdf/2002\\_IVLA.pdf](http://www.csloh.com/research/pdf/2002_IVLA.pdf)
- Lu, J., & Deng, L. (2013). Examining students' use of online annotation tools in support of argumentative reading. *Australasian Journal of Educational Technology*, 29(2), 161-171.
- Lynch, P. J., & Horton, S. (2008). *Web Style Guide* (3rd ed.). London: Yale University Press.
- Mason, B. J., Patry, M., & Berstein, D. J. (2001). An Examination of the Equivalence between Non-Adaptive Computer-based and Traditional Testing. *Journal of Educational Computing Research*, 24(1), 29-39.
- Mills, C. B., & Weldon, L. J. (1987). Reading text from computer screens. *ACM Computing Surveys*, 19(4), 329-358.
- Ni, X., Branch, R. M., Chen, K.-C., & Clinton, G. (2009). The effects of text spacing on screen reading time and comprehension. *International Journal of Instructional Media*(4), 383.
- Nielsen, J. (1997). How users read on the web. Retrieved from Nielsen Norman Group website: <http://www.nngroup.com/articles/how-users-read-on-the-web/>

- Nielsen, J. (2006). F-shaped pattern for reading web content. Retrieved from Nielsen Norman Group website: <http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/>
- Nor, N. F. M., Azman, H., & Hamat, A. (2013). Investigating Students' Use of Online Annotation Tool in an Online Reading Environment. *3L: Southeast Asian Journal of English Language Studies*, 19(3), 87-101.
- Nor, N. F. M., Hamat, A., Azman, H., Noor, N. M., & Bakar, N. A. (2011). An Investigation on Employment of Interactive Technological Tools to Comprehend Online Reading Texts. *International Journal of Learning*, 18(2), 247-260.
- Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: are they equivalent? *Ergonomics*, 51(9), 1352-1375.
- OECD. (2009). *PISA 2009 Assessment Framework: Key competencies in reading, mathematics and science*. Retrieved from <http://www.oecd.org/pisa/pisaproducts/44455820.pdf>
- OECD. (2013). Draft reading literacy framework. Retrieved from <http://www.oecd.org/pisa/pisaproducts/pisa2015draftframeworks.htm>
- OECD. (n.d). About the survey of adult skills (PIAAC). Retrieved from <http://www.oecd.org/site/piaac/surveyofadultskills.htm>
- Olson, J. B. (2014). *Score comparability for web and iPad delivered adaptive tests*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Philadelphia, PA.
- Paek, P. (2005). *Recent Trends in Comparability Studies*. Retrieved from [http://images.pearsonassessments.com/images/tmrs/tmrs\\_rg/TrendsCompStudies.pdf?WT.mc\\_id=TMRS\\_Recent\\_Trends\\_in\\_Comparability\\_Studies](http://images.pearsonassessments.com/images/tmrs/tmrs_rg/TrendsCompStudies.pdf?WT.mc_id=TMRS_Recent_Trends_in_Comparability_Studies)
- PARCC. (2014a). PARCC Field Test: Lessons Learned.
- PARCC. (2014b). *Passage selection guidelines for the PARCC mid-year and summative assessments, Grades 3-11, in ELA/Literacy (Phase 2 Item Development)*. Retrieved from <http://www.parcconline.org/sites/parcc/files/PARCCCombinedPassageSelectionGuidelinesandWorksheets.pdf>
- PARCC. (2015a). About PARCC. Retrieved from <http://www.parcconline.org/about-parcc>
- PARCC. (2015b). *Technology Guidelines for Partnership for Assessment of Readiness for College and Careers (PARCC) assessments: Version 4.4*. Retrieved from <http://www.parcconline.org/sites/parcc/files/TechnologyGuidelines-PARCCAssessments-January2015.pdf>
- Parshall, C. G., Spray, J. A., Kalohn, J. C., & Davey, T. (2002). *Practical Considerations in Computer-Based Testing. [electronic resource]*: New York, NY : Springer New York : Imprint : Springer, 2002.
- Pedley, M. (2006). Designing for Dyslexics: Part 3 of 3. Retrieved from <http://accessites.org/site/2006/11/designing-for-dyslexics-part-3-of-3>
- Pinelli, T. E., Cordle, V. M., & McCullough, R. (1986). A Survey of Typography, Graphic Design, and Physical Media in Technical Reports. *Technical Communication*, 33(2), 75-80. doi:10.2307/43089602
- Piolat, A., Roussey, J.-Y., & Thunin, O. (1997). Effects of screen presentation on text reading and revising. *International Journal of Human - Computer Studies*, 47(4), 565-589. doi:10.1006/ijhc.1997.0145

- Pommerich, M. (2004). Developing Computerized Versions of Paper-and-Pencil Tests: Mode Effects for Passage-Based Tests. *Journal of Technology, Learning, and Assessment*, 2(6).
- PWS. (2009). Text justification - issues and techniques. Retrieved from [http://www.pws-ltd.com/sections/articles/2009/justified\\_text.html](http://www.pws-ltd.com/sections/articles/2009/justified_text.html)
- Richardson, J., Dillon, A., & McKnight, C. (1988). *The manipulation of screen-presented text: Experimental investigation of an interface incorporating a 'movement grammar'*. HUSAT Memo #431. Loughborough University of Technology.
- Sandene, B., Horkay, N., Bennett, R. E., Allen, N., Braswell, J., Kaplan, B., . . . Westat, I. R. M. D. (2005). *Online Assessment in Mathematics and Writing: Reports from the NAEP Technology-Based Assessment Project, Research and Development Series*. Retrieved from <http://nces.ed.gov/nationsreportcard/pdf/studies/2005457.pdf>
- SBAC. (2014a). *Smarter Balanced "Tests of the Test" Successful: Field Test Provides Clear Path Forward*. Retrieved from [http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/10/SmarterBalanced\\_FieldTest\\_Report.pdf](http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/10/SmarterBalanced_FieldTest_Report.pdf)
- SBAC. (2014b). *The Smarter Balanced Technology Strategy Framework and Testing Device Requirements*. Retrieved from Los Angeles, CA: [http://www.smarterbalanced.org/wordpress/wp-content/uploads/2011/12/Tech\\_Framework\\_Device\\_Requirements\\_11-1-13.pdf](http://www.smarterbalanced.org/wordpress/wp-content/uploads/2011/12/Tech_Framework_Device_Requirements_11-1-13.pdf)
- SBAC. (n.d.). Smarter Balanced Assessments. Retrieved from <http://www.smarterbalanced.org/smarter-balanced-assessments/>
- School Curriculum and Standards Authority. (2015). *Online Literacy and Numeracy Assessment (OLNA)*. Retrieved from [www.scsa.wa.edu.au/internet/Documents/.../OLNA\\_Handbook.pdf](http://www.scsa.wa.edu.au/internet/Documents/.../OLNA_Handbook.pdf)
- Schwartz, E., Beldie, I., & Pastoor, S. (1983). A comparison of paging and scrolling for changing screen contents by inexperienced users. *Human Factors*, 25, 279-282.
- Solak, E. (2014). Computer versus Paper-Based Reading: A Case Study in English Language Teaching Context. *Mevlana International Journal of Education*, 4(1), 202-211.
- Strain-Seymour, E., Craft, J., Davis, L. L., & Elbom, J. (2013). *Testing on tablets: Part I of a series of usability studies on the use of tablets for K-12 assessment programs*. Retrieved from <http://researchnetwork.pearson.com/wp-content/uploads/Testing-on-Tablets-PartI.pdf>
- Tortorici, T. (n.d.). Website layouts designed for how the human eye works. Retrieved from Optimized for Humans website: <http://tortoriciinc.com/website-layouts-designed-human-eye-works/>
- VCAA. (2013). Types of Assessment. Retrieved from <http://www.vcaa.vic.edu.au/Pages/prep10/ondemand/types.aspx>
- VCAA. (2014). On Demand Testing. Retrieved from <http://www.vcaa.vic.edu.au/pages/prep10/ondemand/index.aspx>
- Wang, S., Jiao, H., Young, M. J., Brooks, T., & Olson, J. (2008). Comparability of Computer-Based and Paper-and-Pencil Testing in K-12 Reading Assessments: A Meta-Analysis of Testing Mode Effects. *Educational and Psychological Measurement*, 68(1), 5-24. doi:10.1177/0013164407305592
- WARC. (2012). Page length/Size. Retrieved from [http://warc.calpoly.edu/building/page\\_design/page\\_length.html](http://warc.calpoly.edu/building/page_design/page_length.html)
- Weisenmiller, E. M. (1999). *A Study of the Readability of On-Screen Type*. (Doctor of Philosophy), Virginia Polytechnic Institute, Blacksburg, Virginia. Retrieved from

<http://scholar.lib.vt.edu/theses/available/etd-102999-110544/unrestricted/WeisenmillerDissertation.pdf>

Williamson, H. (1966). *Methods of book design: The practice of an industrial craft, 2nd edition*. London: Oxford University Press.

Woody, D. (n.d.). *How document readability and legibility are affected by text justification [White Paper]*. Retrieved from <http://www.zaetric.com/pdf/White%20Paper%20-%20Readability%20Legibility%20and%20Text%20Justification.pdf>

Young, J. (2014). A study of print and computer-based reading to measure and compare rates of comprehension and retention. *New Library World, 115*(7/8), 376-393. doi:10.1108/NLW-05-2014-0051

Zucker, S. (2004). *Administration practices for standardized assessments*. Retrieved from [http://images.pearsonassessments.com/images/tmrs/tmrs\\_rg/AdministrationPractices.pdf?WT.mc\\_id=TMRS\\_Administration\\_Practices\\_for\\_Standardized\\_Assessments](http://images.pearsonassessments.com/images/tmrs/tmrs_rg/AdministrationPractices.pdf?WT.mc_id=TMRS_Administration_Practices_for_Standardized_Assessments)

Zucker, S., Sassman, C., & Case, B. J. (2004). *Technical report: Cognitive labs*. Retrieved from [http://images.pearsonassessments.com/images/tmrs/tmrs\\_rg/CognitiveLabs.pdf](http://images.pearsonassessments.com/images/tmrs/tmrs_rg/CognitiveLabs.pdf)

# NAPLAN Online Readability and Layout Study

February 26, 2016

Laurie Laughlin Davis, Ph.D.  
Emily Lai, Ph.D.  
Rusty Brandt



# Contents

<b>1</b>	<b>Executive summary</b>	<b>3</b>
<b>2</b>	<b>Introduction</b>	<b>5</b>
2.1	Background	5
2.2	Summary of literature review	5
2.3	Purpose of this study	7
<b>3</b>	<b>Method</b>	<b>8</b>
3.1	Participants	8
3.2	Materials	9
3.3	Process	11
3.4	Limitations	12
<b>4</b>	<b>Results</b>	<b>14</b>
4.1	Image placement and size	14
4.2	Passage interface options	24
4.3	Item layout and presentation	34
4.4	General interface observations	40
<b>5</b>	<b>Conclusions</b>	<b>42</b>
5.1	Overall findings and recommendations	42
5.2	Study limitations	46
5.3	Further research	46
<b>6</b>	<b>References</b>	<b>48</b>
	<b>Appendix A</b>	<b>49</b>

# I Executive summary

The National Assessment Program—Literacy and Numeracy (NAPLAN) online readability and layout study was conducted in November 2015 as part of a comprehensive research and development program initiated by the Australian Curriculum, Assessment and Reporting Authority (ACARA) in preparation for the 2017 implementation of the NAPLAN online adaptive testing program. A total of 103 students across four year levels from 11 different schools participated in a series of focus group sessions to evaluate different item display options relative to onscreen layouts. Layout presentations included different options for:

- the location of images in relation to text on screen
- navigation functionality within a single stimulus and across multiple stimuli
- presentation of audio items for spelling
- presentation of item content (including horizontal or vertical response option layouts; font type, line spacing, and line length)

Focus group sessions involved between four and six students from a given year level and lasted for one hour. During this time facilitators asked students to discuss the different features of the displays including what they liked or did not like about them, as well as their preferences between display options. Students viewed the layouts separately on their own devices (PCs, tablets, and tablets with external keyboards) and facilitators followed an interview protocol to prompt students around specific questions about the displays and note their responses. For each set of layouts students were asked to:

- observe and remark on differences between sample displays
- describe for the facilitator how they would interact with the different sample displays
- discuss whether one layout option appeared easier or more difficult than another and why
- suggest how they would improve upon the layouts
- indicate a preference between layout options

The table below provides a summary of recommendations based on student feedback. Students generally found images within reading passages as well as within the stimulus of non-passage based items to be less important than the content of the passage text and the information in the item stimulus. They wanted to read those other elements first as they felt this would set the context for interpreting the information presented in the image. With regard to navigating the passage interface, student preferences were mixed, with younger students preferring a paging interface and older students preferring a scrolling interface. Reasons cited by students for preferring one interface over the other were largely related to their beliefs about ease and efficiency for locating information needed to answer the questions.

Student preferences for line spacing suggest that while younger students preferred double spacing older students preferred the 1.5 line spacing. Results for line width suggest that students at all year levels preferred the text that occupied the full width of the column over text that occupied half the width of the column<sup>1</sup> —although preferences were fairly close for Years 3 and 5 students. In both cases, preferences reflect beliefs about the ease and efficiency of reading the text and a desire to keep the page count to a minimum.

With regard to the response option layout, font, and spelling items, students preferred a horizontal layout of response options, the default font (Helvetica 16 point) for all elements of the item, and the inclusion of a written sentence for spelling items. Students' preference for the horizontal layout of response options was largely driven by efficiency in use of the onscreen space as well as ease of comparison to the information in the question stem. Students' preference for Helvetica was largely driven by a desire to have the same font for all elements of the item and by a general dislike of the Palatino font type. Finally, students' preference for the inclusion of a written sentence to support audio dictation for spelling items reflects a desire for reinforcement of the information through multiple channels, as well as concerns over the ability to understand the audio narration.

#### Summary of recommendations

Layout Option	Year 3	Year 5	Year 7	Year 9
Image placement (relative to passage text)	Right or bottom of text			
Image placement (relative to item components)	Right or inline (with introductory text above)			
Image size	Larger	Larger	Larger	Larger
Passage interface (single passage)	Paging	Paging	Scrolling	Scrolling
Passage interface (multiple passages)	Tabbed Passages	Tabbed Passages	Tabbed Passages	Tabbed Passages
Line spacing	Double spacing	Double spacing	1.5 spacing	1.5 spacing
Line width	Full-column width	Full-column width	Full-column width	Full-column width
Response option layout	Horizontal	Horizontal	Horizontal	Horizontal
Font	Helvetica 16 pt	Helvetica 16 pt	Helvetica 16 pt	Helvetica 16 pt
Spelling items	With sentence	With sentence	With sentence	With sentence

In general, recommendations are consistent across all year levels, but there are two key areas (passage interface and line spacing) where different recommendations are made for younger and older students. Given the differences across year levels in preferences for paging versus scrolling interfaces for passage text (as well as the strength of student opinion on this topic), additional research might be considered to evaluate a hybrid scrolling/paging interface that could be used across all year levels.

In summary, the study yielded many valuable observations and should generally support ACARA in making decisions for online layouts of passage text and items.

<sup>1</sup> Full-column width was defined as 80–100 characters. Half-column width was defined as 40–60 characters or 30–40 characters depending upon the specific passage text.

## 2 Introduction

### 2.1 Background

In preparation for the 2017 implementation of the National Assessment Program—Literacy and Numeracy (NAPLAN) online adaptive testing program, the Australian Curriculum, Assessment and Reporting Authority (ACARA) initiated a comprehensive research and development program that will enable the delivery of tests within an online environment.

The NAPLAN Online Readability and Layout study is a key component of ACARA's online assessment research program. The study was conducted in two phases. In the first phase, a literature review was completed to summarise current knowledge about readability and layout in onscreen assessments, especially reading assessments. The literature review also identified interface, design and layout options that might provide optimal readability for NAPLAN online tests in 2017.

In the second phase of the study, a series of focus groups was conducted in which students across four year levels reviewed and provided feedback on different sample layouts. The second phase of the study was to be based on the findings and recommendations from the literature review. This report contains the results of the second phase of the study.

### 2.2 Summary of literature review

The literature review conducted in the first phase of the NAPLAN Online Readability and Layout study reviewed academic research related to onscreen testing and actual onscreen layouts used by a variety of assessment programs within Australia and globally. Across all online reading assessment tools and programs, four common features and characteristics were observed.

- The stimulus and items were presented in either a top-bottom or left-right format, or a combination of both.
- Although placed in different parts of the screen, “back”, “forward”, “previous” or “next” buttons were used as navigational methods.
- A range of reading tools and features (e.g. highlighter, notepad, etc.) were incorporated into the online reading assessments. These tools were included to facilitate the online reading process.
- The majority of the assessments included multiple-choice questions and/or constructed response items (i.e. short answers). Additionally, some assessments included technology-enhanced items (TEIs) such as drag-and-drop questions, point and click, and select or highlight phrases from a passage.

Research related to the comparability of paper and onscreen assessments revealed a mixed set of results, with some studies finding comparability between paper and onscreen tests, and others finding differences

in student performance between the two modes. The authors state that inconsistencies in earlier findings could be related to variations in the visual quality of the presentation across the two modes, and suggest that improvements in display screen and interface technologies could partially mitigate mode effects resulting from the disparity between the presentations across modes.

In addition to reviewing research in the domain of assessment, the authors reviewed research in web development to shed light on differences in how students may interact with test content across modes. Evidence from eye tracking studies of adult web readers revealed that readers scan and read through web pages differently compared to reading on paper: online readers tend not to read every word on a web page (Nielsen, 1997). Other research (Jones, 2012; Nielsen, 2006) suggests that readers follow an “F-pattern” when browsing or reading web content, such that they first read in a horizontal movement across the first few lines on the top of the screen before moving down the page slowly and reading across in a second horizontal movement (thus creating the top and middle bars of the F), and then finally scanning the left side of the screen in a vertical movement (to create the vertical bar of the F). Using the F-Layout, the best or most important content of a web page should be placed across the top, and the first two paragraphs should state the most important information. The following subsections, paragraphs or bullet points should start with information-carrying words, so that users will notice these sections when scanning down the left side of the screen or content area. It is unknown to what extent this research (conducted with adults) will generalise to students.

The literature review also summarised research on several factors which can impact the ability of students to read and comprehend content onscreen. These factors include typeface characteristics, text layout and format, and presentation layout. Font, line spacing, and line length all play a role with regard to typeface characteristics. Research has found differences between serif and sans serif fonts related to onscreen readability. Specifically, sans serif fonts (e.g. Arial and Comic Sans MS) were perceived as being easier to read and more desirable for online texts compared to serif typefaces (e.g. Times New Roman and Courier New), with a significant effect found for reading accuracy (Bernard, Chaparro, Mills, & Halcomb, 2002; Hojjati and Muniandy, 2014).

Some research suggests that spacing between the lines of a text can impact on the speed of on-screen reading (Dyson, 2004). Specifically, Ling & van Schaik (2007) found that double spacing was better for reading speed and accuracy than single spacing. However, other research finds no significant difference in performance due to line spacing (Loh, Branch, Shewanown, & Ali, 2002). Additionally, Hojjati & Muniandy (2014) found that control of spacing is a significant part of page design as loose spacing can reduce legibility.

To date, studies online length have produced mixed results. Dyson and Kipping (1998) found that the number of characters per line affected participants’ reading rate, such that the longest line was read faster than the shortest line. However, Bernard, Fernandez, & Hull (2002) found no difference in reading time or reading efficiency between three line length conditions for children and adults. Mills & Weldon (1987) reported that readers were more likely to lose their place within the text when reading longer line lengths as they require greater lateral eye movement. The authors of the literature review conclude that it appears that shorter line length may be more suitable for children.

With regard to text layout and format, the literature review authors state that the two most common techniques of navigating or progressing through a passage of text are scrolling and pagination. Depending

on length of texts and device types, scrolling and pagination each have useful features that help readers navigate through a text efficiently. While some research suggests that scrolling through the presentation of the text may disrupt readers' structural representation of the text and negatively impact their reading performance (Kim and Kim, 2013), evidence from older studies suggests that there is no significant difference in reading performance between scrolling and paging (Mills & Weldon, 1987). Additionally, Kim and Kim (2013) also found that most of the students in the study were comfortable using scrolling to navigate through passages. Dillon (1992) noted that experienced users preferred scrolling to manipulate text, possibly due to its speed. Therefore, the authors of the literature review concluded that the advantages and disadvantages of each technique are not conclusive.

If an item includes multiple elements such as stimulus, question prompt, response options, response format and graphics, the positioning of these elements ought to be considered. The literature review authors assert that an onscreen presentation of an item might require a different layout than a paper presentation of the same item. They note a lack of research evidence on the effects of item organisation on reading speed, accuracy, comprehension or cognitive processes and further state that there is very limited information or guidelines regarding graphic placement for the online assessment environment.

## 2.3 Purpose of this study

The purpose of this study is to provide information on the effectiveness of different screen layout presentations, identify key strengths and weaknesses for each presentation, and identify the extent to which differences in the presentation and layout of assessment items affect readability by evaluating students' preferences in response to the different presentations. Results of the study will inform decision making relative to interface development and test content formatting for NAPLAN online tests in 2017. The study focuses on the following key features of item readability and layout.

- the location of images in relation to text on screen
- navigation functionality within a single stimulus and across multiple stimuli
- navigation across multiple stimulus texts
- presentation of audio items for spelling
- presentation of item content (including horizontal or vertical response option layouts; font type, line spacing, and line length)

Results of this study will be discussed in relationship to the research summarised in the literature review. Specifically, results are organised by key differentiating features of the various layout options (as indicated in the summary table above). Within each section of the results, a description of the differentiating feature along with any relevant references to the research literature are provided, followed by a summary table of student preferences for each set of sample displays in which the differentiating feature is addressed. Following the table, a summary of student comments is provided which supplies a rationale for the majority preference and notes a counter-rationale for minority preferences, where relevant. Student quotes (as indicated through the use of “”) are used throughout the rationale to add specificity and give direct voice to the student participants. Figures which show the different sample displays are also included.

## 3 Method

### 3.1 Participants

ACARA nominated 11 schools to participate in the 2015 Readability and Layout study. The participating schools were drawn from four of the eight educational jurisdictions in Australia: Western Australia, South Australia, Queensland and Victoria. All schools were located in metropolitan areas and government, Catholic, and independent schools were represented.

Despite the small sample size a large amount of qualitative data was collected through the cognitive interviews with students. The interviews yielded a rich data set the analysis of which is provided in the results section below.

Each school was asked to nominate between four to six students per NAPLAN year level to participate in the study. This resulted in a total of 103 students. Students were selected from the following year levels; Year 3 (36 students), Year 5 (27 students), Year 7 (19 students), and Year 9 (21 students). In addition, co-educational schools were asked to select both boys and girls to participate. The number of interviews totalled 20 across the 11 nominated schools. Table 1 shows the breakdown of student participation.

Schools were asked to nominate students who typically demonstrate a set of attributes, as follows:

- comfortable sharing ideas and opinions in front of a group
- can express an interest in technology
- can listen to ideas of other students and build upon those ideas
- enjoys working in groups and solving problems
- comfortable talking with adults

Table 1 Student participation in study by state, region and year level

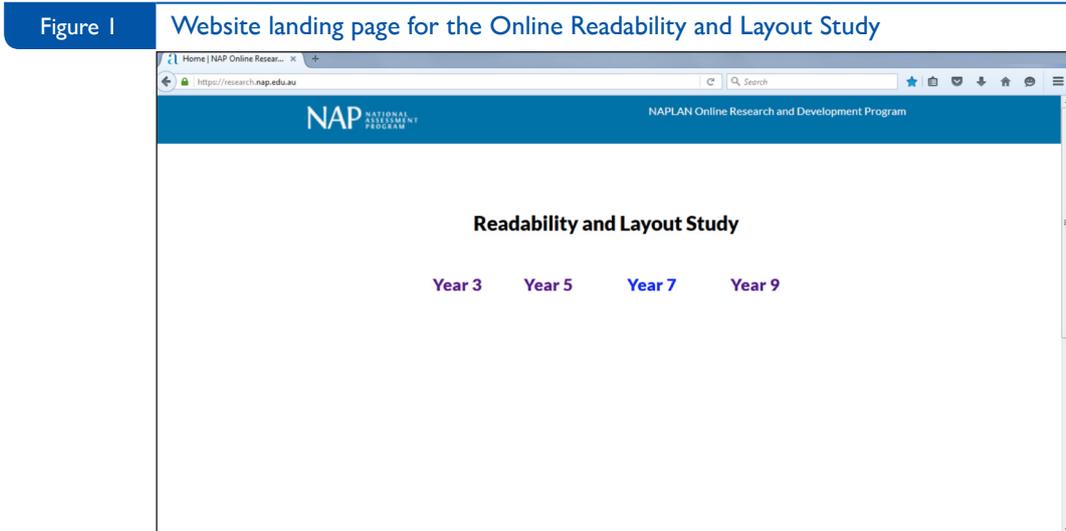
State	Region	Year level	Number of students	Number of interviews	Male students	Female students
QLD	Metro	3	13	2	7	6
VIC	Metro	3	12	2	6	6
WA	Metro	3	11	2	6	5
Totals		3	36	6	19	17
QLD	Metro	5	11	2	4	7
VIC	Metro	5	5	1	3	2
WA	Metro	5	11	2	8	3
Totals		5	27	5	15	12
QLD	Metro	7	9	2	4	5
WA	Metro	7	10	2	5	5
Totals		7	19	4	9	10
QLD	Metro	9	8	2	2	6
SA	Metro	9	6	1	3	3
WA	Metro	9	7	2	5	2
Totals		9	21	5	10	11

## 3.2 Materials

### Item displays

ACARA provided sets of item displays (screenshots) to be viewed by participating students. Each screenshot presented a static view of a test item and associated onscreen functionality. Although students could view the different interface controls, they were not able to interact with them as they would in an operational testing situation. Screenshots were grouped into sample sets of between two to four screenshots that reflected different options for presenting the same item content. For example, a sample set might have shown the same item but with the image to the left of the question in one view and to the right of the question in another view. Different sets of items were made available to each year level. While the items themselves were designed to be appropriate to each year level, each display set generally represented the same variety of layout options.

A website was created specifically for use in this study (see [Figure 1](#)) and students accessed their year-specific displays from this web interface. Students accessed the website using Firefox, Safari or Internet Explorer (version 10 or later) browsers. Google Chrome could not be used to access the last two samples, and for this reason schools were discouraged from using this browser. Item displays were grouped by the domains of Reading, Spelling and Numeracy with students reviewing samples related to Reading passage text displays first, spelling item displays second, and numeracy item displays third. While each sample was presented in the context of a specific domain the presentations and display options were intended to generalise beyond that context.



In addition to the static screenshots, students were also presented with an interactive sample pulled from a publicly available practice set on the Partnership for Assessment of Readiness of College and Careers (PARCC) website. This interactive sample was used to demonstrate a scrolling interface for navigating within passage text as well as a tabbed interface for navigating between two passage texts.

### Devices

The displays were viewed on PCs (including laptops), tablets (including iPad minis) and tablets with external keyboards. Students used their own devices to view the displays. Interviewers did not report any occasions where different devices were used within focus groups. Table 2 provides a breakdown of the number and type of devices used across the different year levels.

**Table 2** Number of devices used by year level and device type

Year Level	PCs	Laptops	Tablets	iPads	iPad mini	Tablets with external keyboards	Total
3	12	7	–	11	6	–	36
5	5	4	6	12	–	–	27
7	10	5	–	–	–	4	19
9	4	11	–	3	–	3	21
Totals	31	27	6	26	6	7	103

## 3.3 Process

### Overall focus group structure

Each school was asked to nominate four to six students in two separate year levels to participate in the study. The study design indicated that research would be conducted through small focus group interview sessions. The interviews were scheduled for one hour in duration. A single facilitator was assigned to conduct the interview with each focus group.

### Study locations and equipment set-up

Schools were advised of the importance of conducting the interview/s in an appropriate room. It was recommended that the room be quiet, free of distraction and allow for internet access. Internet access was required for all students as the expectation was that each student would view the layouts separately on his or her own device. Schools were asked to make the interview room available 30 minutes before the scheduled interview start time, to allow the facilitator to complete the necessary setup beforehand. Schools were able to accommodate this request without concern.

### Interview protocols

A set of interview protocols specific to each year level were developed by Pearson and approved by ACARA. These interview protocols defined the trajectory of the interview from facilitator introductions to guiding and probing questions and finally to ending activities, during which students stated their layout and presentation preferences. Guiding and probing questions were intended to focus students on the salient differences between the sample displays and to solicit their opinions on these differences. Student responses to these questions provided insight into the features of a display that were important to students as well as their general approach to interacting with test content. The ending activities provided an opportunity for facilitators to ask students to identify their preferred layout. The data obtained from the ending activities are summarised in tables in the results section of this report. An example interview protocol is provided in Appendix A.

Facilitators used recording devices in addition to hard copy interview protocol forms to capture the discussion and responses from students. This information was then entered into Pearson's data management software. To obtain the best interview results it was recommended that, where possible, facilitators organise the students in a circle facing one another. This would not only make it easier for the facilitator to record everything, but also encourage discussion among the students, and between the students and the facilitator.

## Facilitator recruitment and training

Pearson recruited four facilitators to complete the interviews in participating schools.

Facilitators were selected on the basis of the following criteria.

- be available from approximately 7:30 am to 3:00 pm for the interview days
- possess a valid working with children clearance (or valid teacher registration)
- be highly organised, flexible, reliable and punctual
- be sensitive to issues relating to the confidentiality and privacy of the interview data and materials at all times
- possess highly developed communication skills
- be familiar with technology and possess the ability to perform basic IT troubleshooting
- possess strong observation and data collection skills
- demonstrate a high level of attention to detail

Facilitators were required to attend an online training session convened by Pearson before conducting interviews in schools. All facilitators received training on a complete set of interview protocols including a group walkthrough. In addition, the training session covered all aspects of the facilitator role including the following:

- pacing – correct pacing through the displays to allow equal time for discussion on each display
- the ideal seating arrangements (students sitting in a circle facing one another)
- the best location for the recording device – non-obtrusive, central location so as to facilitate best recording results
- strategies for making students feel at ease in the interview setting
- the importance of maintaining neutrality, non-evaluative disposition from the facilitator
- probing/clarifying questions – open-ended follow-ups, non-leading follow-up questions
- the debrief and closure of the focus group

## 3.4 Limitations

The study relied on voluntary participation of schools. This was particularly challenging given the time of the year (November–December). The original study design prescribed an ideal sample of six interviews per year level (24 interviews in total). The actual number of interviews conducted was 20. This number was lower than that specified in the study design for two reasons. Firstly, the design had aimed for the participation of 12 schools; however, only 11 schools were available. Secondly, a number of these schools were able to provide one year level only.

As indicated in Table 1, the nominated schools were all located in metropolitan regions. This was a consequence of the voluntary nature of the study coupled with the time of year at which schools were asked to volunteer.

Facilitators reported two instances where students were not sitting face-to-face, or in a circle, but with their backs to one another. Ideal seating arrangements were not possible in these cases due to the architecture of the room coupled with the type of device used (desktop PCs). While this made the task of interviewing the students more challenging, it did not hamper the collection of data as per the interview protocols.

In some cases students encountered errors with the PARCC links (the pages would not load). In such cases facilitators used their own devices (iPads) to present the item layouts to students. Facilitators did not report that this adversely affected the course of the interview or prevented its completion.

# 4 Results

## 4.1 Image placement and size

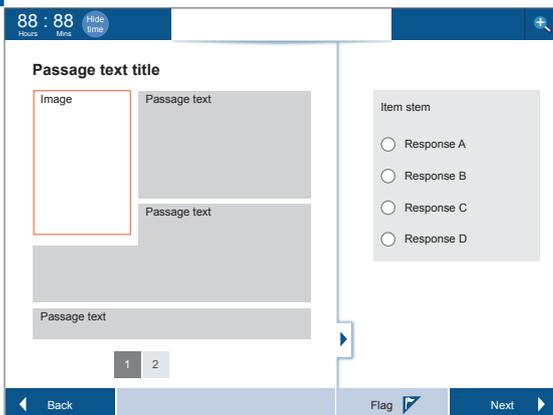
### Image placement relative to passage text

Images are often used in conjunction with passage texts to add illustration or elaboration to the content of the texts and to help students better understand what they are reading. In the current study, passage text was presented in the context of the reading domain, but these findings would generalise to any passage-based assessment domain (e.g. a science or social science exam where students were asked to read informational text and respond to domain specific items). In this study, students were given different options for the placement of these images relative to the passage text. Samples included single-image examples as well as multiple-image examples. Additionally, samples included images interspersed in the passage text as well images outside of the passage text (to the top, bottom, left, or right). With one exception (Years 3 and 5 sample 5) all images were presented in the same left-hand panel as the passage text, with the item and response options presented in the right-hand panel.

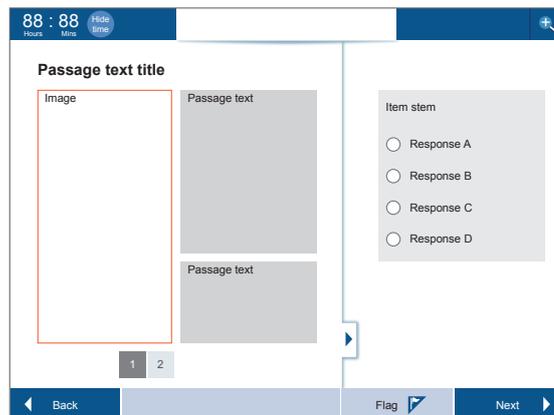
Table 3 presents a summary of student preferences for each set of sample displays which reflected options for image placement relative to the passage text. In evaluating these samples, students' primary consideration was how the information in the image could best support their comprehension of what they were reading in the text.

Two different strategies for how students would use images to support their reading comprehension emerged through analysis of student comments. In the first strategy, students sought to use the images as a pre-organiser for what they were about to read. These students generally preferred to see the images at the top and/or to the left of the text (e.g. Year 7 sample 4 – [Figure 2A](#); Year 5 samples 12 and 13 – [Figure 2B](#)) so that they would naturally be the first piece of information they reviewed when reading from left to right. In the second strategy, students sought to use the images to elaborate on or reinforce what they had just read. These students generally preferred to see the images placed at the bottom and/or to the right of the text (e.g. Years 3 and 5 sample 4 – [Figure 3A](#); Year 5 samples 10 and 11 – [Figure 3B](#)), so that the text itself would be the first piece of information they reviewed when reading from left to right.

Figure 2A-B Image at the top and/or left of the text

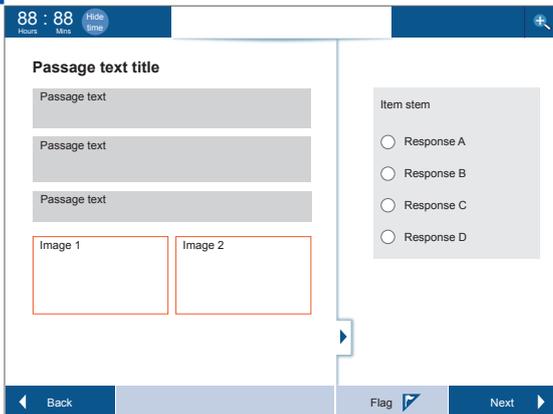


2A-Years 7 and 9 sample 4

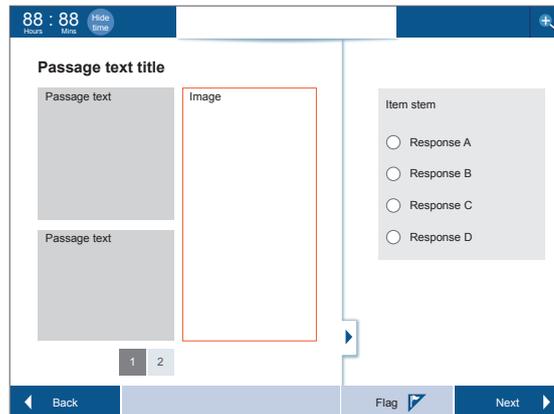


2B-Year 5 sample 12

Figure 3A-B Image at the bottom and/or right of the text



3A-Years 3 and 5 sample 4



3B-Year 5 sample 10

In general, more students appeared to favour the strategy of using the image for elaboration and reinforcement (positioning preferences of right and bottom). Older students frequently referenced a specific reading strategy as part of the rationale to support their preferences. Younger students did not mention specific reading strategies to support their preferences. It is unclear whether they lacked formalised reading strategies or simply could not articulate them to facilitators.

Table 3 Student preference for image placement relative to passage text across year levels

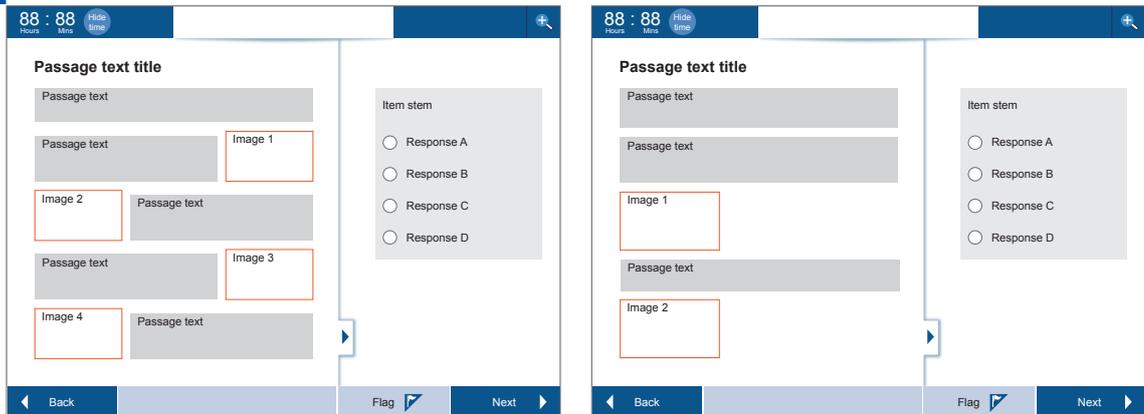
Year level	Item descriptor	Display options	Preference
3	Plants for dinner	Sample 1: Images interspersed left and right Sample 2: Images stacked right only	Sample 1 (11%) Sample 2 (89%)
3	Baby echidnas	Sample 3: Images between paragraphs Sample 4: Images below paragraphs Sample 5: Large images in second column (question hidden)	Sample 3 (28%) Sample 4 (45%) Sample 5 (28%)
3	Banana split	Sample 6: Images left Sample 7: Images right	Sample 6 (50%) Sample 7 (50%)
5	Baby echidnas	Sample 3: Images between paragraphs Sample 4: Images below paragraphs Sample 5: Large images in second column (question hidden)	Sample 3 (15%) Sample 4 (81%) Sample 5 (4%)
5	Injured bat	Samples 10 and 11: Image top right Samples 12 and 13: Image top left	Samples 10 and 11 (81%) Samples 12 and 13 (19%)
7	Ant communication	Sample 3: Image top right Sample 4: Image top left Sample 5: Image bottom left Sample 6: Image above text	Sample 3 (26%) Sample 4 (16%) Sample 5 (37%) Sample 6 (21%)
9	Ant communication	Sample 3: Image top right Sample 4: Image top left Sample 5: Image bottom left Sample 6: Image above text	Sample 3 (14%) Sample 4 (59%) Sample 5 (27%) Sample 6 (0%)

- Year 9 student: “Sample 5 has a picture in the middle and forces students to look at the picture. The picture is not important. Text is important in order to answer the question.”
- Year 9 student: “If the image is on the right, you read the text and glance at the image, you don’t put too much thought into it, but if it’s on the left you kind of have to look at it before you read.”

The overall visual appearance in terms of neatness and organisation of elements within the question was also important to students. They wanted things to be placed in a way that facilitated a quick visual inspection of the information they needed to answer the question. Students tended to dislike options that interspersed multiple images within the passage text (e.g. Year 3 samples 1 and 3 – [Figure 4A](#); Year 5 sample 3 – [Figure 4B](#)) and felt that this looked disorganised. Inherent in this (and mentioned by students across all year levels) was the notion that images could be a “distraction” and that placement was crucial to avoiding this outcome. For example, most Year 3 and 5 students felt that it was better to present the pictures at the end of the passage text, than interspersed in the passage text, noting that this was more organised, less distracting, and prevented students from wasting time.

- Year 5 student: “Sample 3 has stuff everywhere.”
- Year 5 student: “3 is bad ... text then photo ... text then photo ... split up is not good.”

Figure 4A-B Images interspersed within passage text



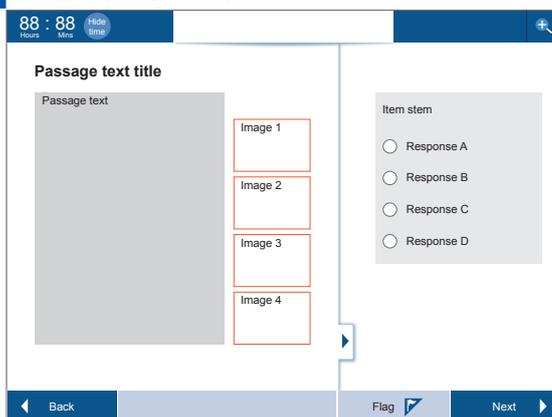
4A-Year 3 sample 1

4B-Years 3 sample 3

Additionally, when multiple images were presented with the text, students generally preferred for the images to be grouped together (either to the right or left side of the text or at the bottom of the text) and ordered in a way that let them more easily connect each image to the text (e.g. Year 3 sample 2 - [Figure 5](#); Years 3 and 5 sample 4 - [Figure 3A](#)). Although images in some of the samples included photo captions, many students did not notice these, or did not feel that this information was sufficient to connect the images to the text.

- Year 3 student: “Sample 2 picture matches with the writing. On sample 1, I'm not sure which picture goes with what writing”
- Year 5 student: “Sample 4 is better because pictures are grouped together.”

Figure 5 Images grouped together



5-Year 3 sample 2

Year 9 students also noted the visual impact of image placement relative to the justification of the passage text. Images placed to the left were sometimes identified as visually more appealing because when left-justified text is used, the image appears to line up cleanly with the text. Images placed to the right were noted by some Year 9 students as being less desirable because of this.

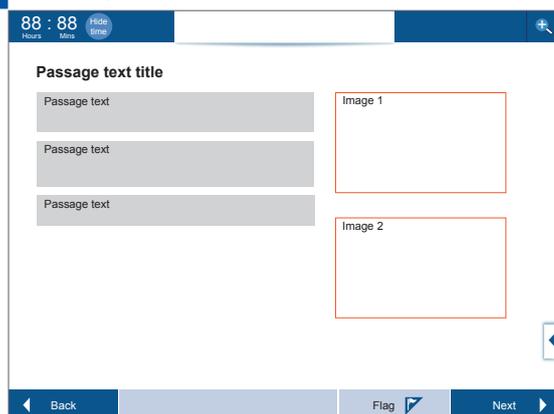
- Year 9 student: “The lines finish at different points. They should finish at the same point.”
- Year 9 student: “Having the image on the left is better because it makes the writing line up and makes it easier to read.”

One of the samples presented to students in Years 3 and 5 (sample 5) – [Figure 6](#), had the text expanded to full screen. This view covers the item. To access the item, a small blue arrow (expansion tab) needed to be clicked in the lower right corner of the screen. Students at both year levels were not always sure where to find the item and while some located the blue expansion arrow and correctly identified its function to minimise the text, some students could not figure out how to locate the item at all. A few students assumed that the item would appear in the blank space directly below the passage text on a real test and that this space was left blank for study purposes. Students who found the expansion arrow suggested that it be made more prominent (e.g. bigger, different colour). Additionally, some students thought that the images in this sample were too large and would be distracting.

- Year 5 student: “Don’t like that you can’t see the question.”
- Year 5 student: “Pictures should be smaller than the text to prevent distraction.”

Figure 6

## Text expanded to full screen



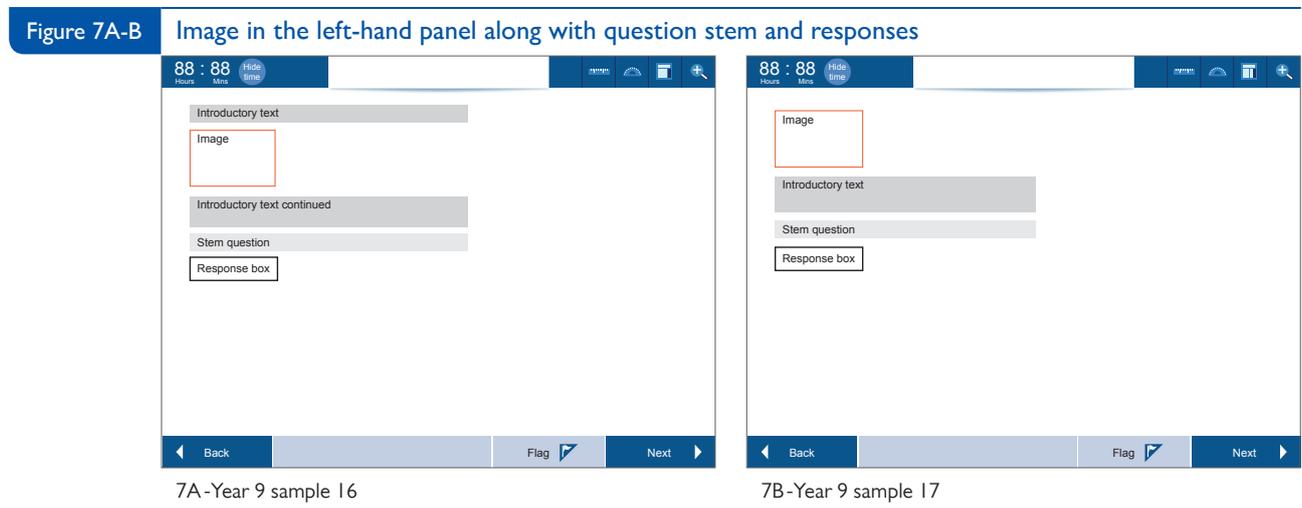
6-Years 3 and 5 sample 5

### Image placement relative to item components

Images such as photos, graphs, data tables, or illustrations are often included in non-passage based items as part of the question stimulus to which students are responding. In the current study, images were presented in the context of the numeracy domain, but these findings would generalise to any domain area which images are used (e.g. a social science exam in which students are asked to view a map). In this study students were shown samples which included images placed in different positions relative to the question stem and responses. Some samples displayed the image in the left-hand panel along with the question stem and responses (Year 3 samples 18 and 19 – [Figures 10A-B](#); Year 5 samples 16, 17, and 20 – [Figures 11A-B](#) and [13B](#); Year 7 sample 17 – [Figure 13C](#); Year 9 samples 16, 17, and 18 – [Figures 7A-B](#) and [12](#) respectively). In these cases image placement varied relative to whether the image appeared first or

introductory text appeared first. Other samples displayed the image in the right or left-hand pane with the question stem and/or response options in the other (Year 5 sample 21 – [Figure 8A](#), and Year 5 sample 22; Year 7 sample 18 – [Figure 8B](#), and Year 7 sample 19 – [Figure 9](#)).

Table 4 presents a summary of student preferences for each set of sample displays which reflected options for image placement, relative to the item stem and responses. Similar to student strategies when viewing images and passage text, student preferences were again related to how they wanted to use the information in the images. However, unlike images within the passage text, students largely saw the images in the item stem as containing important information (such as graphs or data tables) needed to answer the item rather than a potential distraction. As a result, the preference for placement of the image was strongly influenced by how much students felt they needed to know, to engage productively with the image before seeing it.



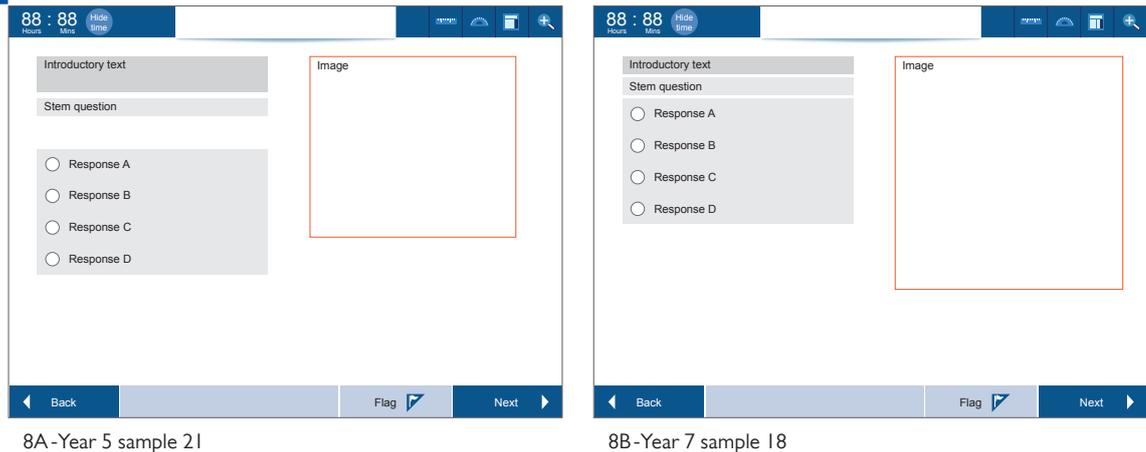
Year level	Item descriptor	Display options	Preference
3	Time line	Sample 18: Image above stem question Sample 19: Image below stem question	Sample 18 (50%) Sample 19 (50%)
5	Paper folding	Sample 16: Image in middle of text Sample 17: Image above text	Sample 16 (85%) Sample 17 (15%)
5	Column graph	Sample 20: Image (small) above stem Sample 21: Image (large) on right Sample 22: Image (large) on left	Sample 20 (4%) Sample 21 (96%) Sample 22 (0%)
7	Triangle graph	Sample 17: Image (small) above stem Sample 18: Image (large) on right Sample 19: Image (large) on left	Sample 17 (6%) Sample 18 (75%) Sample 19 (19%)
9	Temperature	Sample 16: Table inline Sample 17: Table above text Sample 18: Table left of text	Sample 16 (73%) Sample 17 (7%) Sample 18 (20%)

Almost all Year 5 and 7 students preferred to have the graph image to the right of the question for the triangle on grid and column graph items (Year 5 sample 21 and Year 7 sample 18 – [Figure 8A-B](#)) rather

than to the left or above the question stem. These students thought that it was logical to go from left to right and read the question first, then look at the graph. They noted that students could not understand the graph without reading the question first.

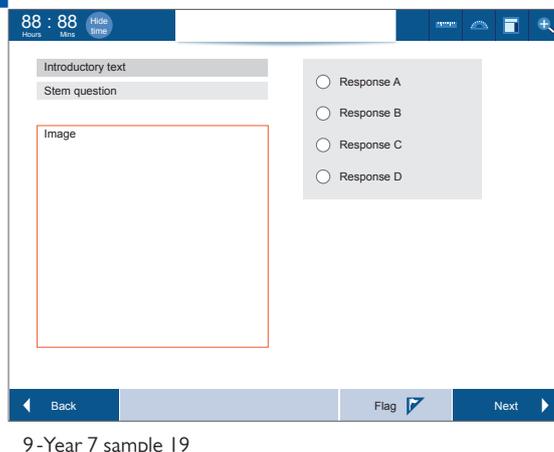
- Year 5 student: “Having pictures below the words gets students to read the text instead of just skipping to the pictures.”

Figure 8A-B Image placed to the right of question



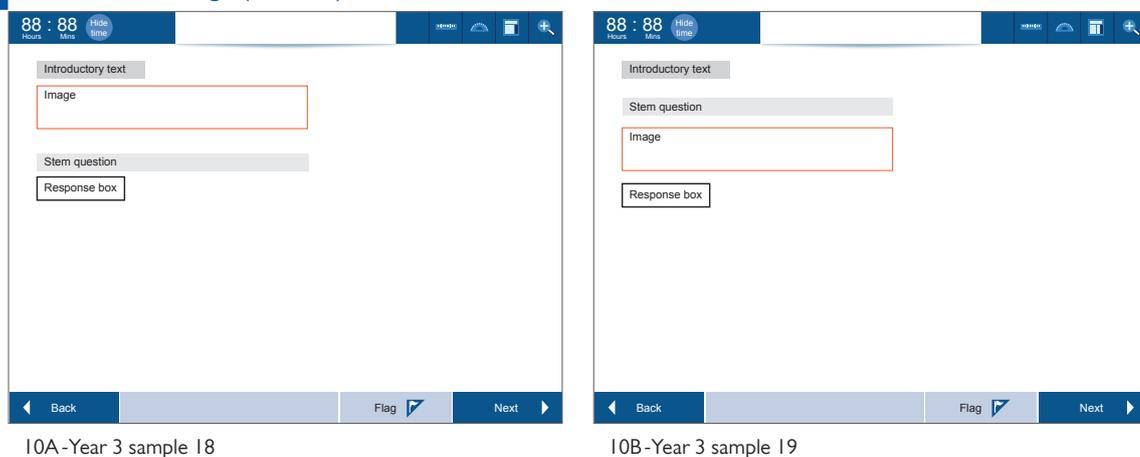
Some students commented that they could see the benefit to having the image to the left of the question, so that they could read/interpret the graph first before reading the question. However, this did not ultimately change their preferences. Additionally, the separation of the question stem from the response options in Year 7 sample 19 (question above the graph image on the left and response options on the right – [Figure 9](#)), made this sample less attractive to students, so may have influenced their preference regardless of graph position. Several students noted that it did not make a difference which side the graph was on and thought that the size of the graph was the more important factor. Image size is also one reason why few students indicated a preference for the samples where the graph image was placed above the question, indicating that the graph image in these options was too small.

Figure 9 Separation of question stem from the response options



Students in Year 3 were mixed in their preferences for where the image of the timeline should be placed relative to the item stem – [Figure 10A-B](#). Half the students thought it was better to have the image above the stem question (Year 3 sample 18 – [Figure 10A](#)) because they wanted to see the timeline first before seeing the question, noting that this was the most logical presentation. These students also observed that this placement resulted in having the answer box directly below the question they were to answer. The other half of students thought it was better to have the image below the question stem (Year 3 sample 19 – [Figure 10B](#)), so that they already knew what to do with the information in the timeline when they saw it. These students liked that both of the stem sentences (the introductory sentence and the stem question) were grouped together so they could read them both before looking at the timeline. It is worth noting that in both samples, the introductory sentence was the first element students encountered and provided some pre-organising context for students.

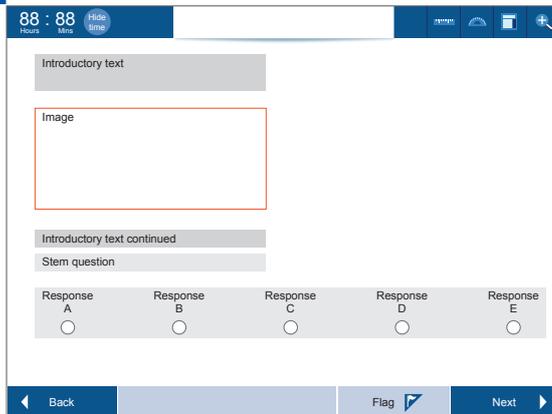
Figure 10A-B Placement of image (timeline) relative to item stem



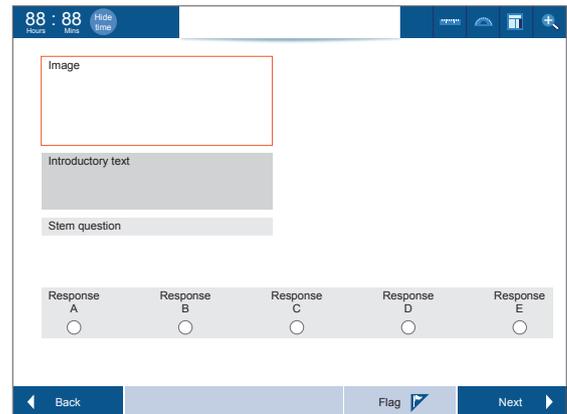
Having pre-organising context prior to the image was also important to Year 5 students in reviewing the samples for the paper folding item. Most students preferred to have the image presented in the middle of the text (Year 5 sample 16 – [Figure 11A](#)), than to have the image as the first element encountered (Year 5 sample 17 – [Figure 11B](#)). These students said that they preferred to read the text first and then see the illustration in the image. Students thought this was a very sequential and logical presentation and that seeing the picture first without having read the introductory text might be confusing. Students described sample 17 as “more overwhelming” and “harder to figure out.” Similarly, most Year 9 students preferred having the data table in line with the text (sample 16) for the temperature item, as it was preceded by introductory text which gave them a clear idea as to what was in the table.

- Year 9 student “like instruction first telling me what the table is about.”
- Year 9 student: “... because if I saw the table first I would try to figure out what it means and interpret without the explanation.”

Figure 11A-B Image in the middle and top of the text



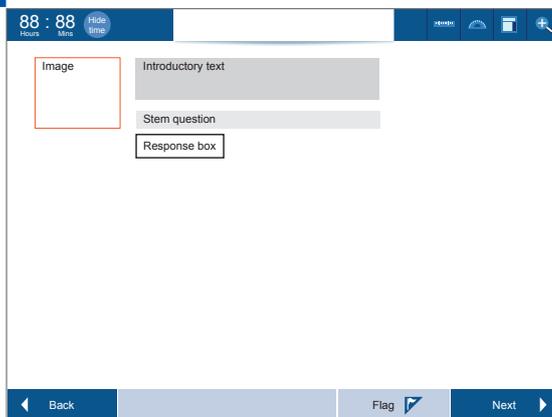
11A-Year 5 sample 16



11B-Year 5 sample 17

Year 9 students who preferred sample 18 (table left of text) – [Figure 12](#) found it more visually appealing than the other samples because they liked reading across rather than up and down, and having the side-by-side comparison to the question. However students also observed that the layout used with sample 18 might not be workable with a larger graphic. Additionally, one student noted that with sample 18, the table seemed “pushed to the side as if not important.”

Figure 12 Table image placed to the left of text



12-Year 9 sample 18

## Image size

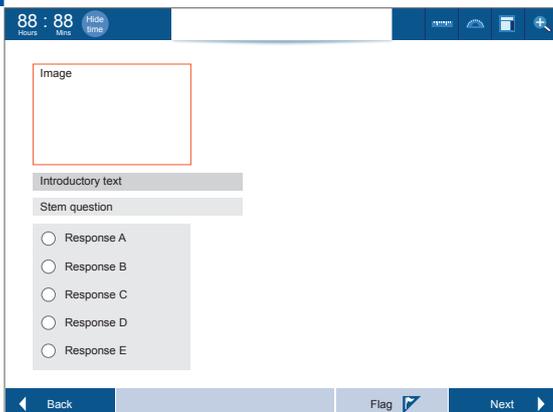
The size of an image can influence the visibility and readability of the information in the image. In this study students were shown images of varying sizes to determine what size would provide the best visibility and readability for students. Table 5 presents a summary of student preferences for each set of sample displays which reflected options for image size. In general, students had a preference for larger images which allowed them to better read and interpret the information in the images. However, there were two instances where students did not endorse the sample with the largest image. In the first case (Baby echidnas Years 3 and 5) – [Figure 6](#), students disliked that the question was not visible because of the larger images and found it confusing that not everything was on one screen. In the second case (Triangle on grid for Year 7 sample 19 – [Figure 9](#)), students disliked the separation of question stem and responses in the sample with the largest graph option. In both situations the size of the image interfered with the preferred presentation of other key elements within the item.

**Table 5** Student preference for image size across year levels

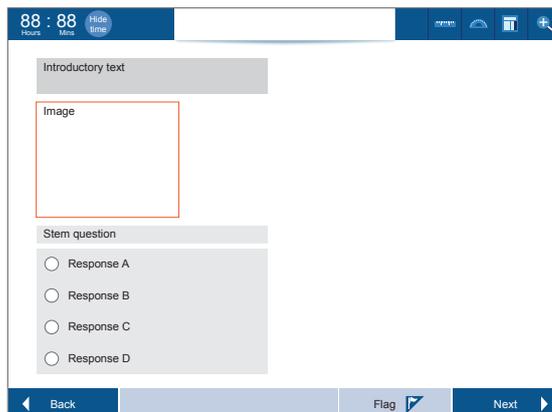
Year level	Item descriptor	Display options	Preference
3	Baby echidnas	Sample 3: Small images between paragraphs Sample 4: Small images below paragraphs Sample 5: Large images in second column (question hidden)	Sample 3 (28%) Sample 4 (45%) Sample 5 (28%)
3	Calculator/Ruler	Samples 20 and 21: Small image size (image above question) Sample 22: Large image size (image to left of question)	Samples 20 and 21 (8%) Sample 22 (92%)
5	Baby echidnas	Sample 3: Small images between paragraphs Sample 4: Small images below paragraphs Sample 5: Large images in second column (question hidden)	Sample 3 (15%) Sample 4 (81%) Sample 5 (4%)
5	Column graph	Sample 20: Small graph (image above question) Sample 21: Medium graph (image to right of question) Sample 22: Large graph (image to left of question)	Sample 20 (4%) Sample 21 (96%) Sample 22 (0%)
7	Triangle graph	Sample 17: Small graph (image above question) Samples 18 and 19: Large graph (right side and left side)	Sample 17 (6%) Samples 18 and 19 (94%)

Students noted that when smaller images were used and placed above the question this could introduce scrolling (Year 3 samples 20 and 21 for Calculator/Ruler) – [Figure 13A](#). The scrolling issue seemed to be exacerbated when students viewed the sample displays on tablets due to the smaller screen sizes (Year 5 sample 20 for Column graph) – [Figure 13B](#). Students also commented that space was not well used when the smaller images were placed above the question (Year 7 sample 17 for Triangle on grid) – [Figure 13C](#).

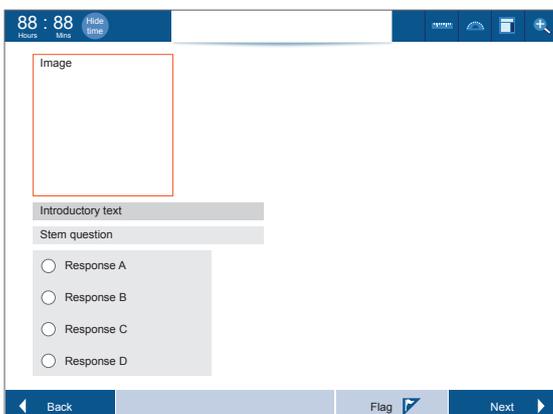
Figure 13A-C Placement of images above question



I3A-Year 3 samples 20 and 21



I3B-Year 5 sample 20



I3C-Year 7 sample 17

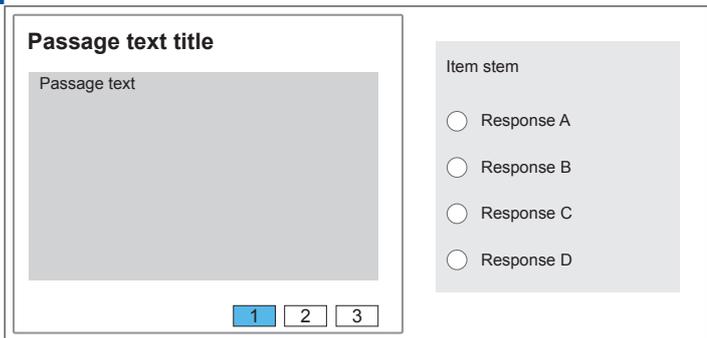
## 4.2 Passage interface options

In the current study, passage text was presented in the context of the reading domain, but these findings would generalise to any passage-based assessment domain (e.g. a science or social science exam where students were asked to read informational text and respond to domain specific items).

### Navigation within a single stimulus text

When the text of a passage is sufficiently lengthy that it does not fit on a single screen, interface decisions must be made about how to support student navigation to the remainder of the text. Two options were presented in this study. In the first option students used a “paging” interface which allowed them to move from one page of text to the next using numbered buttons which appeared at the bottom of each page of text (see [Figure 14](#)). The numbered buttons are highlighted to show the student which page of text is currently showing. Students in the study referred to this display option as a “book system” approach to navigation within the passage text, as it preserves the concept of a page and students use the numbered buttons to turn the page as they would in a physical book.

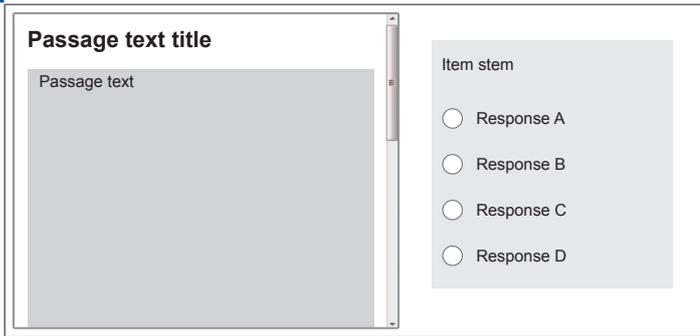
**Figure 14** Paging interface for passage text



14 - PARCC sample; Years 3 and 5 sample 24; Year 7 sample 21; Year 9 sample 22

The second option presented a “scrolling” interface which allowed students to view text as a single continuous piece through the use of an onscreen scroll bar that allowed students to move up and down within the left-hand panel of the screen, while the right-hand panel remained static (see [Figure 15](#)). Students could access this scrolling functionality by using a mouse to drag the onscreen scroll bar, by using the scroll wheel on their mouse, by using arrow keys on the keyboard, or by using their finger on a touch-screen device.

**Figure 15** Scrolling interface for passage text



15 - PARCC sample; Years 3 and 5 sample 23; Year 7 sample 20; Year 9 sample 21

Unlike other sample displays, students were not asked to make a choice between paging and scrolling interfaces, so no summary table is presented. However, student preference was fairly clear based on the comments students made during the focus group discussion. Preference for paging versus scrolling appeared related to both year level and device. Younger students who had less experience with reading onscreen preferred the numbered tabs as they thought that it would be easier for them to locate information using the page references. Older students who had more experience with reading onscreen tended to prefer scrolling and felt that it gave them greater control and that they could more easily locate information related to specific sentence and paragraph references than with the paging interface. Amongst older students, preference for navigation controls also related to the type of device they were using and how easy or difficult it was for them to control the mouse (external mouse versus trackpad) or scroll through content. Students using an external mouse seemed the most positive toward scrolling, with students using a trackpad or a touch-screen tablet less positive.

Many Year 3 students had challenges figuring out the scroll functionality and some had to be shown how to use the scroll bar to view the rest of the story. Many students did not notice the scroll bar at all and suggested making it more visible to indicate that scrolling would be needed. Once they were prompted, students were reasonably successful in using the scroll functionality; however, most students did not like it. Students mentioned that they were more used to flipping pages (like in a book) compared to scrolling. Additionally they did not like having to scroll in two places (within the left panel for the passage and separately in the right panel to view the questions).

By contrast, Year 3 students were largely positive about using the numbered buttons to move between pages in the paging interface. Students noted that it did not take as much time as scrolling and that selecting a page required one button click versus the continuous action of scrolling.

- Year 3 student: “Like it. Can press the boxes easily.”
- Year 3 student: “That’s way easier than scrolling.”
- Year 3 student: “Once you click on the tabs you don't have to do anything else, whereas with scrolling you have to keep going.”

Some students were confused by the numbered buttons or thought that the numbered buttons should be larger or better labelled. Within some of the samples which used the paging interface, students noted confusion between the “next” and “back” buttons and the numbered buttons —indicating that they thought the “next” and “back” buttons might navigate them to the rest of the passage text, whereas the numbered buttons might navigate them to the next item (opposite of the intended functionality). Additionally some students preferred scrolling, or thought that scrolling would have worked just as well as using the numbered buttons to move between pages.

Most Year 5 students recognised the need to scroll to see the rest of the passage text, but did not like the scrolling interaction, indicating that it was “confusing” and that they found it “intimidating and boring.” Year 5 students were also concerned that scrolling might result in students losing their place in what they were reading. Most students at this year level indicated a preference for the numbered buttons to navigate within the reading passage rather than scrolling, indicating that it was much easier and more immediate and direct in terms of navigation.

- Year 5 student. “Prefer buttons because you can immediately click on a page and it will take you there.”
- Year 5 student. “Instead of constant scrolling click on the number to instantly get you there.”

Year 5 students noted that the buttons were easier to notice than the scrolling controls, because they were highlighted, whereas students might not realise that there was a scroll bar. Additionally, students noted that having pages allowed them to use this information to better locate what they had just read. Because the concept of pages is retained, students appeared to have a better idea of where they had read a particular piece of information, and this made it easier for them to retrieve relevant information necessary to respond to the question.

Students in Years 3 and 5 commented that the size of the numbers and buttons should both be enlarged and more clearly labelled. Some students at Year 5 commented that they preferred scrolling, because it gave them greater control in navigating through the passage text.

Preference between paging and scrolling to navigate within the reading passage was mixed at Year 7. Students who preferred the paging interface were concerned that they might lose their place while scrolling, and thought that using numbered tabs would be faster and waste less time. Some of these students had problems using the scrolling functionality because of the browser or the device they used to access the sample. Others thought that the scrolling controls should be made more obvious.

- Year 7 student. “Numbers are faster in getting you from one page to another. If scrolling too fast, students might miss out on some information or lose the spot at which they were at.”
- Year 7 student: “Not clear that you needed to scroll ... have a scroll down sign. Students may not know that you need to hover the mouse over the box in order to scroll.”
- Year 7 student: “Need to be full size thick scroll bar so you can click on it and drag it. The thin one is too difficult to click and drag.”

Year 7 students who preferred scrolling felt that scrolling provided greater control and were concerned about the amount of time that would be wasted waiting for additional pages to load with the paging interface. These students also thought scrolling would be easier in order to find a specific sentence or paragraph references. Students at this year level also knew how to use arrow keys to scroll and some used this method rather than the scrollbar. Additionally, some students did not realise the numbered buttons were there and thought they should be made larger. They also suggested moving the buttons to the top of the passage text to make them more visible. Some students at Year 7 also indicated confusion between the numbered buttons for paging and the “next” and “back” buttons. A few students at this year level also indicated that they thought the blue arrow for expansion tab might take them to the next page of text. Others thought it would expand the text across the full screen and that they would be able to see all of the text at once. One student suggested having the option of either scrolling or using numbered buttons so students could decide what was easiest for them.

- Year 7 student. “Scrolling is easier if you’re looking for something in particular. If you need to compare two things in a text, you can just scroll compared to having to keep clicking between pages. Takes up time.”
- Year 7 student. “I think I like scrolling better. It’s just easier than the buttons. It’s kind of more obvious you are on a computer, if you want to see more you usually scroll.”

A majority of Year 9 students preferred scrolling to paging as a way to navigate within the passage text, noting that it provided greater control of where they are in the text and was quicker. A primary concern of Year 9 students with the paging interface was how long it would take to load the individual pages and whether this would waste time. As with Year 7 students, Year 9 students also mentioned that scrolling made it easier to compare information and to look for a specific sentence or passage number references.

- Year 9 student. “Scrolling is better because you can point to where you are at the top of the screen. Scrolling is quicker to get to the top of the section.”
- Year 9 student. “Scrolling would be easier because it flows better and you don’t have to worry

about clicking each page and whether you remembered the correct page you needed to retrieve information from.”

Although Year 9 students preferred scrolling they suggested that the scroll bar be made bigger and wider. They also observed that it would be difficult to use the scroll bar with a trackpad on a laptop. They suggested that it would be best to have a “proper mouse” but noted that students could use the arrow keys to scroll instead. Additionally, Year 9 students noted some challenges with having separate scrolling controls for the passage text and the item overall.

A few Year 9 students preferred the paging interface because they felt that it would be “sturdier” and more robust than a scrolling approach. This concern was voiced primarily from students viewing the samples on a tablet. They were concerned about losing control over the scrolling, and scrolling beyond what was intended because of the accelerated scrolling features on touch-screen devices. These students also felt that it would be easier for them to locate information using the numbered buttons. However, students again indicated confusion amongst the “next” and “back” buttons, the blue expansion tab arrow, and the numbered paging buttons. They suggested labelling the “next” button as “next question” to more clearly differentiate its purpose.

- Year 9 student. “Pages are sturdier though. Less can go wrong. Sometimes you can accidentally scroll so this is good.”
- Year 9 student. “A book system is easier for a lot of people because you can’t go wrong. Scrolling is easier to refer to text but a lot could go wrong.”

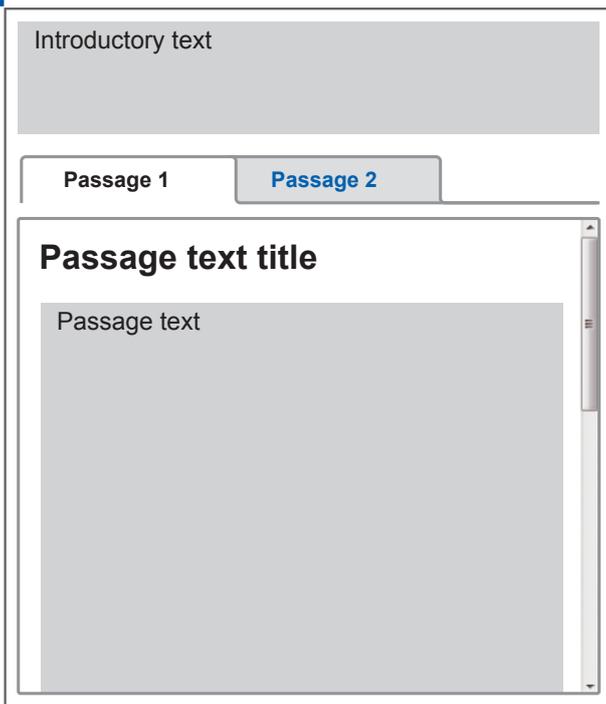
Additionally, students were asked about the number of pages of text they would consider reasonable to read through. Most Year 3 students preferred that text be kept to a single page noting that it was confusing when the text had more than one page and that students would lose interest or forget what they had read on the previous page. Older students were more open to considering multiple pages of text, but there was a common theme that the number of pages should be kept to a minimum. As with the Year 3 students, older students were concerned about being able to stay focused on the text as they read and remember what they had read on previous pages. Students were also concerned about the added effort of having to switch between multiple pages to find text references needed to answer the associated questions. The specific range of responses relative to how many pages would be acceptable was similar across Years 5, 7, and 9, but varied considerably within each year level. Some students indicated that 1 or 2 pages would be ideal, others suggested that 3 or 4 pages should be the maximum, and a few students even said as many as 5 pages would be tolerable. There seem to be a general agreement that 6 or more pages would be too many and that this would lead students to not read the text and just guess at the questions.

### Navigation across multiple stimulus texts

In many passage-based assessment situations, students are asked to read multiple-stimulus texts and make comparisons across those texts when answering the questions. Students in this study were asked to react to a display option where the titles of the passage texts were presented as different tabs within the passage interface (see [Figure 16](#)). Students could navigate between the passage texts by selecting the different titled tabs which would then bring the text of that passage to the foreground. Across all year levels, students were easily able to identify the tabs as the mechanism for switching between passage texts and were able

to do so without issue. Students described this functionality as “clear and straightforward”, “easy”, and “pretty self-explanatory”. Some students suggested that the functionality could be improved by making it more obvious which passage had been selected and wanted more visual differentiation on the tabs (e.g. by adding an outline around the passage title in the tab when it is selected or using a different colour to indicate which passage is selected).

**Figure 16** Tabs for multiple-stimulus texts

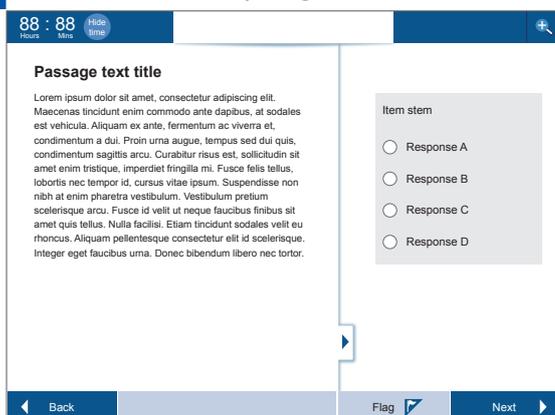


I6 - PARCC sample; Years 3 and 4 sample 23; Year 7 sample 20; Year 9 sample 21

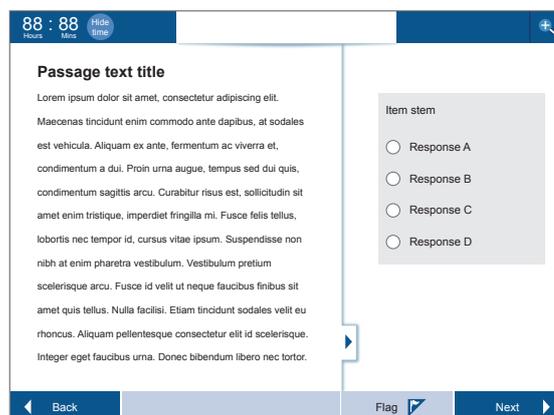
### Line spacing and width

The way in which passage text is displayed can be critical to students’ ability to efficiently read and comprehend the information it contains. In this study, students were shown samples of two different line spacing (1.5 and double spacing) – [Figure 17A-B](#), as well as two different line widths (full-column and half-column width) to determine how these features influenced student impressions of the readability of the passage texts. Implications for the number of pages resulting from each display option were also explored.

Figure 17A-B 1.5 and double line spacing



17A-Years 7 and 9 sample 1



17B-Years 7 and 9 sample 2

Table 6 presents a summary of student preferences for each set of sample displays which reflected options for different line spacing within the passage text. Student preference for line spacing appears to be related to year level and potentially also to their level of reading experience. Specifically, most students in Years 3 and 5 preferred double spacing whereas students in Years 7 and 9 were roughly split between the two options, with a slight majority preferring the 1.5 spacing.

Students at Years 3 and 5 thought that double spacing was easier to read because it was more spaced out, and also thought it would be easier to keep your place while reading and avoid re-reading lines of text. However, many of these students acknowledged that the double spacing might be too much (“too spaced out”; “too much space being used”) and suggested a slight reduction in spacing between the lines. Year 3 and 5 students who preferred 1.5 spacing thought that there appeared to be less writing, since it physically took up less space on the screen. These students also disliked the text wrapping around the image caused by the double spacing. However, they indicated some concern with the size of the text and said they would like the words to be bigger (larger font) but not spaced out more.

Table 6 Student preference for line spacing across year levels

Year level	Item descriptor	Display options	Preference
3	Bath or shower	Sample 8: 1.5 spacing Sample 9: Double spacing	Sample 8 (28%) Sample 9 (72%)
5	Bath or shower	Sample 1: 1.5 spacing Sample 2: Double spacing	Sample 1 (7%) Sample 2 (93%)
7	Choosing a book	Sample 1: 1.5 spacing Sample 2: Double spacing	Sample 1 (53%) Sample 2 (47%)
9	Choosing a book	Sample 1: 1.5 spacing Sample 2: Double spacing	Sample 1 (52%) Sample 2 (48%)

Students at Years 7 and 9 were roughly split in their preference for 1.5 or double spacing. Students who preferred the 1.5 spacing said that having the text spaced together would allow them to read it faster and noted that “if it’s spaced out it takes longer to find the next line and read it all out.” These students also

noted the visual impact of the layout, noting that the 1.5 spacing looked more compact and made a better use of space. Some students said that they were more accustomed to this style of layout and found it less distracting.

Students said that they thought the font size (Helvetica 16 point) was good and that the spacing wasn't too close together. They did not seem to feel that they would have a problem re-reading the same line. Students who viewed the sample displays on a tablet suggested that the text could be enlarged by zooming in if needed.

- Year 9 student: “As a whole you know how much of it there is and it just looks better. Sample 2 [double spacing] looks awkward with that much space. It looks strange.”
- Year 9 student: “sample 2 [double spacing] looks like there's more text to follow as it takes up the whole space.”

Students who preferred double spacing noted that the spacing made it easier to read and would help them keep track of which line they were on and avoid skipping lines. These students mentioned that it was easier to identify the important parts of the passage text (“easier to skim the text”) with the double spacing. However, these students also noted that double spacing takes up a lot of space on the screen (“too much space when you first look at it”). Additionally, students observed that because the text, when double spaced, runs right to the bottom of the screen in sample 2 for Choosing a Book, it's unclear if there is more text to read.

- Year 9 student: “When its bunched up you can lose your line or read the same one twice.”

Table 7 presents a summary of student preferences for each set of sample displays which reflected options for different line widths. Student preference for line width appears to be related to year level. Specifically, students in Years 7 and 9 had a strong preference for text that occupied the full-column (full-column width) whereas students at Years 3 and 5 were more mixed in their preferences, with a slight majority of Year 3 students favouring text that occupied the full-column and a stronger majority of Year 5 students favouring text that occupied half the column (half-column width)<sup>2</sup>.

**Table 7** Student preference for line width across year levels

Year level	Item descriptor	Display options	Preference
3	International Day	Samples 10 and 11: Full-column width Samples 12 and 13: Half-column width	Sample 10 and 11 (53%) Sample 12 and 13 (47%)
5	International Day	Samples 6 and 7: Full-column width Samples 8 and 9: Half-column width	Sample 6 and 7 (37%) Sample 8 and 9 (63%)
7	Can you hear the bells?	Samples 7 and 8: Full-column width Samples 9 and 10: Half-column width	Sample 7 and 8 (89%)* Sample 9 and 10 (11%)
9	Feline: Friend or foe?	Samples 7 and 8: Full-column width (2 pages) Samples 9, 10 and 11: Half-column width (3 pages)	Sample 7 and 8 (100%) Sample 9, 10 and 11 (0%)

\* Note that only 2 of 4 classes reviewed these samples for Year 7.

<sup>2</sup> Full-column width was defined as 80-100 characters. Half-column width was defined as 40-60 characters or 30-40 characters depending upon the specific passage text.

Year 7 and 9 students thought that full-column width was easier to read and made better use of space than half-column width. Students noted that with full-column width there were fewer pages, thought the text was easier to read, perceived it as more "normal" and visually appealing ("like reading a book"), and said it made it seem that there was less to read. Students thought that half-column width was too narrow and noted that people their age like to read across not down.

- Year 7 student: "All the way across; people our age read like that."

Older students were more aware of and concerned by the impact of line width on the number of pages. It should be noted that in the Year 9 sample, *Feline: Friend or foe?*, half-column width created an additional page (went from two to three pages) which made the impact of line width more obvious than in other year levels. The desire for fewer pages also appeared to be related to test taking strategies relative to referencing the question and the text. Students' concerns in this regard seemed to be related to the paging interface. Some students suggested they might have been less concerned about the line width if there was a scrolling interface to navigate through the passage text.

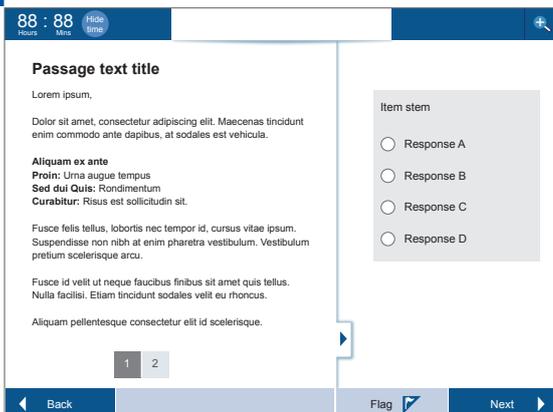
- Year 9 student: "The fewer the pages the better because when you're answering questions you don't have time to go through so many pages of text."
- Year 9 student: "Having it on as few pages as possible makes it easier to compare to the question."
- Year 9 student: "Most people read the question first and then look for it in the text... they won't read if they know thing is too long."
- Year 9 student: "Have a scrolling feature so you don't have to flip through pages and waste time loading."

Year 9 students thought half-column width might be beneficial if they wanted to quickly skim the text, but noted that younger level students would probably not benefit from this, as it is "a higher level technique." Some students mentioned that the full-column width samples could be improved by adding greater vertical spacing between the lines (as they had seen in previous examples).

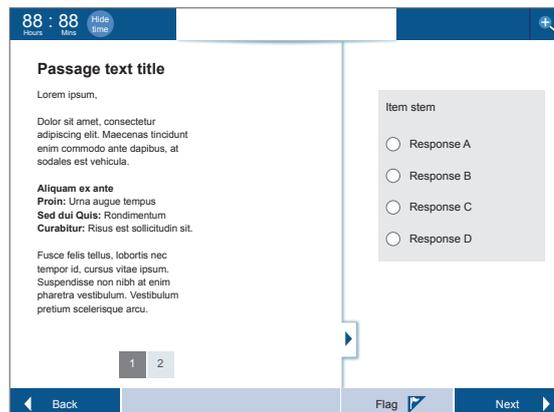
Amongst Year 3 and 5 students, those who preferred the full-column width ([Figure 18A](#)) liked that it was more spread out, and thought this would make it easier to read. These students generally understood that with the half-column width there would be more pages and saw this as a potential drawback.

- Year 3 student: "You don't read down, you read across."
- Year 3 student: "[the half-column width sample] is not good because they only use up half of the page which forces students to scroll between pages more."

Figure 18A-B Full and half-column width text



18A-Year 3 sample 10



18B-Year 3 sample 12

Conversely, students who liked the half-column width (Figure 18B) thought the narrow text would be easier to read. Students expected half-column width to include a second column and thought it might be improved by including something (like a picture) in the second column. Some students thought that the half-column width sample looked more like a poem than a passage text.

- Year 3 student: “Columns are better. It’s easier to read... I like keeping it to half the page instead of going all the way across.”
- Year 3 student: “It seems like you don’t have to read as much, which may appear favourable to some students.”
- Year 3 student: “[the half-column width sample] looks like a poem; neat and easy to read.”

Students who preferred the half-column width also noted several deficiencies with the full-column width sample. Specifically, they noted that the page break between pages 1 and 2 occurred mid-sentence in the full-column width sample which they did not like and thought would be confusing. Additionally, students observed that there was not very much text on page 2 in the full-column width sample and thought something should have been done (e.g. “play with the font”) to keep it all on one page. Both of these issues likely had an influence on student preference for the half-column width. Full-column width may have been preferred by more students if these issues had not been present.

- Year 5 student: “Not wasting time going to the second page since there’s barely anything there.”

## 4.3 Item layout and presentation

In the current study, most individual items (those not associated with a passage text) were presented in the context of the numeracy domain, but findings related to response option layout and font would generalise to any domain area. The layout options for the Spelling items are domain specific with regard to the inclusion or exclusion of a written sentence in addition to the audio sentence. However, findings relative to presentation and labelling of audio controls could generalise to other domains where audio features are included as part of the item stimulus.

### Response option layout

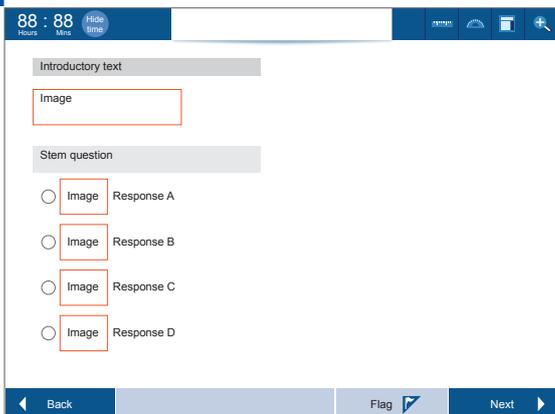
In paper based tests, response option displays typically occur in one of two formats. They may follow a Z- or an N- based pattern where responses are laid out in two columns (A-B-C-D left to right or top to bottom). Alternatively they may be laid out vertically in a single column, with the paper testing booklet in portrait orientation. However, it is unclear whether these layouts are optimal for onscreen testing, because the aspect ratio of most devices assumes a landscape orientation (though tablets can be rotated between portrait and landscape). In this study, students were presented with both vertical and horizontal layouts for response options to evaluate their preferences.

Table 8 presents a summary of student preferences for each set of sample displays, which reflects options for response option layout. Across all year levels, students indicated a strong preference for a horizontal layout of response options, because they are used to reading from left to right. They also thought it made better use of space and might minimise scrolling (especially for students testing on tablets) – [Figure 19C](#).

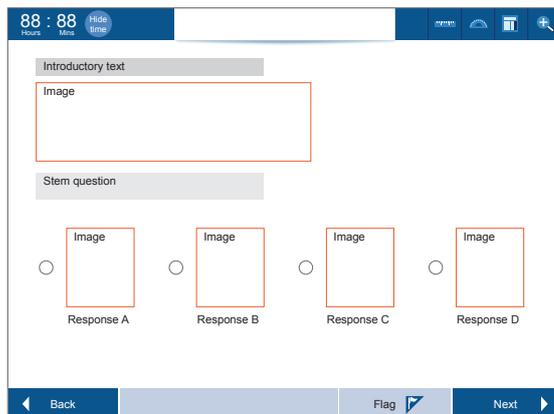
- Year 5 student : “You can see them all on one page so you don’t have to scroll down”
- Year 9 student: “If you scroll down you can’t see the question anymore...then you’d be looking at it without the equation.”
- Year 9 student: “With 14 it would be hard to compare them if you can’t see them at the same time.”

Year 3 students comparing samples 16 and 17 for Fish magnets ([Figure 19A-B](#)) noted that the horizontal layout was easier, since the response options matched the horizontal pattern in the question stem. Students also noted that the pictures in the response options on sample 17 were bigger (less eye strain) and clearer, which made it easier to understand. Students in Year 5 noted that the horizontal orientation of the response options was easier to read and that there was more space to choose between responses. Students who used tablets to view the sample displays noted that the extra space was important, because it was easier to select options with a thumb touch.

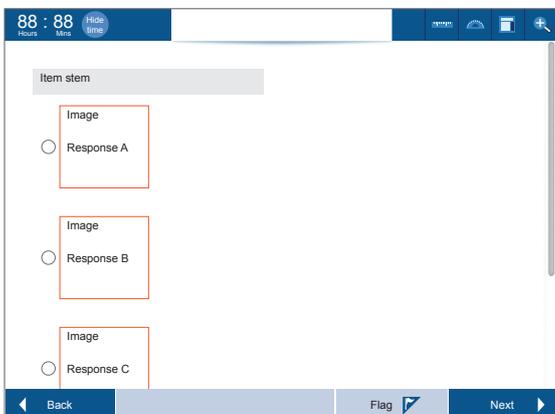
Figure 19A-C Vertical and horizontal layouts of response options



19A-Year 3 sample 16



19B-Year 3 sample 17



19C-Year 9 sample 14 (scrolling required)

Table 8 Student preference for response option layout across year levels

Year level	Item descriptor	Display options	Preference
3	Fish magnets	Sample 16: Vertical layout Sample 17: Horizontal layout	Sample 16: 14% Sample 17: 86%
5	Rope	Sample 18: Vertical layout Sample 19: Horizontal layout	Sample 18: 21% Sample 19: 73%
7	Blocks	Sample 13: Vertical layout Sample 14: Horizontal layout	Sample 13: 5% Sample 14: 95%
9	Graph responses	Sample 15: Vertical layout (scrolling) Sample 16: Horizontal layout	Sample 15: 5% Sample 16: 95%

Students who preferred a vertical layout said it was more familiar and felt that they could more clearly associate responses with radio buttons (i.e. answer bubbles). Year 9 students noted that the placement of the radio buttons relative to the answer choices was confusing in sample 15, since the radio buttons were between the response options and it was unclear which button went with which option. They thought it would be better to place the radio buttons above or below the response options. Additionally, several students at Year 9 suggested the use of a 2 × 2 layout of response options rather than either the vertical

or horizontal layouts. In particular, one student noted that a horizontal presentation might be difficult with a larger number of response options and that stacking response options both vertically and horizontally would be better in this situation.

### Font type

The font used within an item can aid or detract from the readability of the item. Research has suggested that fonts which work well on paper (serif fonts) may not work well onscreen and that sans serif fonts were perceived as being easier to read onscreen. However, the use of different fonts within an item can add emphasis to highlight critical information and may also communicate to students the different types of information they are reading. In this study, students were presented with the same item shown with a single font (Helvetica 16 point) for all parts of the item, and separately with a different font (Palatino) for certain parts of the item stem, image, and response options (Figure 20A-B).

Figure 20A-B Use of Helvetica and Palatino fonts in test items

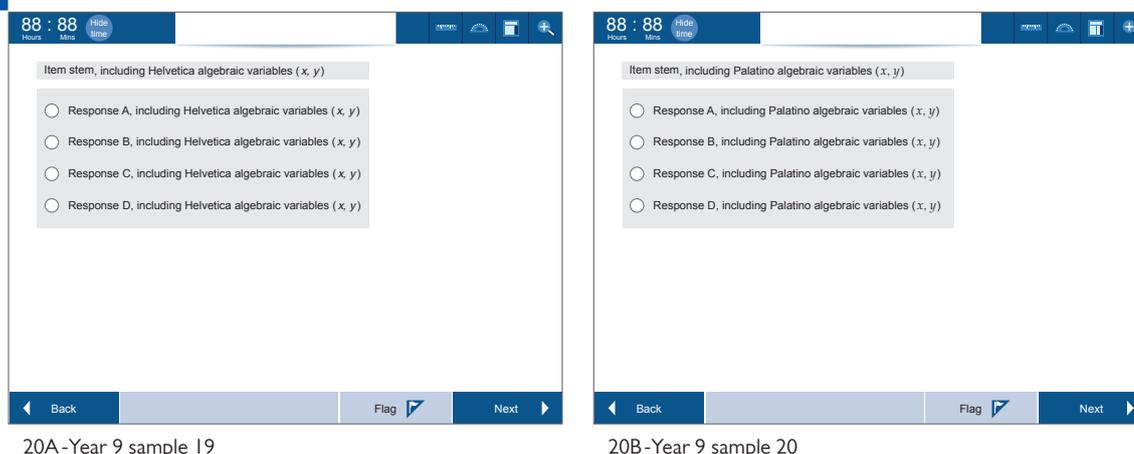


Table 9 presents a summary of student preferences for each set of the sample displays which reflected options for different font types. The majority of students evaluated across year levels, preferred samples with the consistent Helvetica font to samples with the Palatino font used for certain elements of the item. Students noted that the Helvetica font appeared more bold, was easier to read, and that the writing itself appeared larger. While most students preferred the Helvetica font to the Palatino font, when asked to make a choice between the two, many students said that the font was not important to their ability to read or understand question.

**Table 9** Student preference for font type across year levels

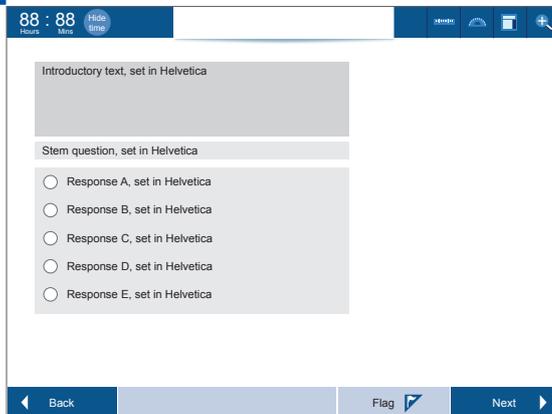
Year level	Item descriptor	Display options	Preference
3	Calculator/Ruler	Sample 20: Helvetica font Sample 21: Palatino font for image labels	Most preferred sample 20 over 21*
5	Alyssa	Sample 15: Helvetica font Sample 16: Palatino typeface	Sample 15: 89% Sample 16: 11%
7	Triangle graph	Sample 17: Helvetica font Samples 18 and 19: Palatino font for point labels	N/A**
9	Expression	Sample 19: Helvetica font Sample 20: Palatino font for equations and variables	Sample 19: 80% Sample 20: 20%

\*Because samples 20 and 21 were also compared against sample 22 (which had an enlarged image), it was not possible to determine specific percentages for preference between 20 and 21, as students made the choice amongst all three options.

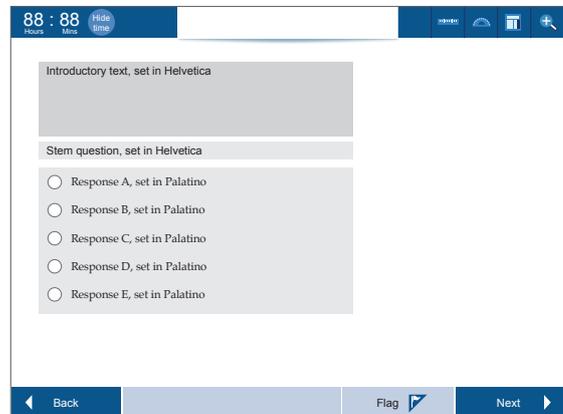
\*\*Although font is a differentiating characteristic between samples 17, 18 and 19, only the letters identifying the points on the triangle reflected a different font and students did not mention the font in their comments. Therefore no percentage information on font preference is presented for these samples.

Students generally liked having the response options in the same font as the question (Year 7 sample 15 – [Figure 21A](#)). However, a few students thought having the answers in a different font made the distinction between the question and the answer more clear (Year 7 sample 16 – [Figure 21B](#)). Some students suggested using a different font for the question stem and options. Students thought the Palatino font was “more formal”, like the font “usually seen in math textbooks and worksheets” and also said it was “nice but a bit thin.”

**Figure 21A-B** Same and different fonts for question stem and response options



21A-Year 7 sample 15



21B-Year 7 sample 16

## Spelling items

It is typical that spelling tests which ask students to spell individual words will include the use of the tested word in a sentence to provide context and avoid confusion for students. Onscreen testing allows audio files embedded within the items to take the place of the test invigilators reading the sentence aloud. These audio files require students to interact with the onscreen controls of the audio player to hear the words in a sentence. In this study students were shown samples of spelling items that contained only the audio controls and were asked to compare these to samples for the same spelling item that contained both the audio controls and a written sentence to support the aural delivery, with the word to be spelled omitted.

Table 10 presents a summary of student preferences for each set of sample displays which reflects options for spelling item presentation. Across all year levels, students indicated a strong preference for displaying the sentence within the text of the item stem, in addition to having it read through the audio player (see [Figure 22](#)). Students thought that without the sentence, the question was confusing and students might not know what to do. Additionally, they mentioned that seeing the sentence written out was important, because students might have difficulty understanding the audio recording. Some students at Years 7 and 9 suggested that it would be helpful to have an option to slow down the speed with which the sentence was read.

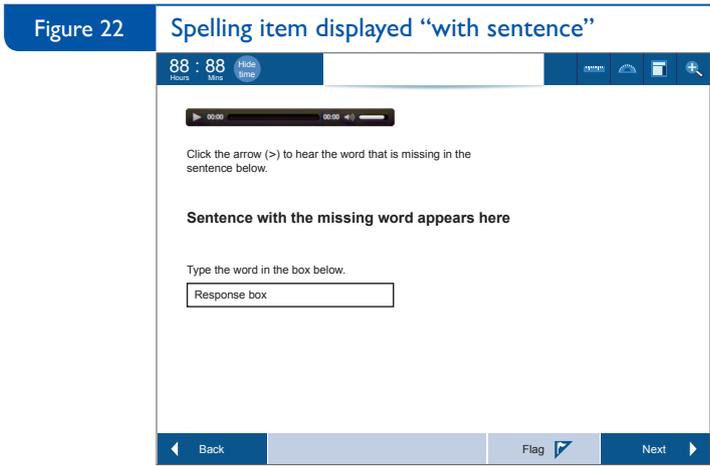
- Year 7 student: “I gather information better when I read it instead of hearing it... and sometimes it’s good to get it twice the information... in case you don’t get it first time... so text and audio combination is best.”

Students who preferred the sample without the sentence said it was visually “less busy”, was quicker because they did not need to read the sentence, and felt that the instructions provided were sufficient. They also noted that having the blank in the sentence added confusion as they did not know where to record your response. Some students who preferred the sample without the sentence also said that it was typical of spelling tests not to have the sentence and so they were used to this type of presentation.

**Table 10** Student preference for spelling item format across year levels

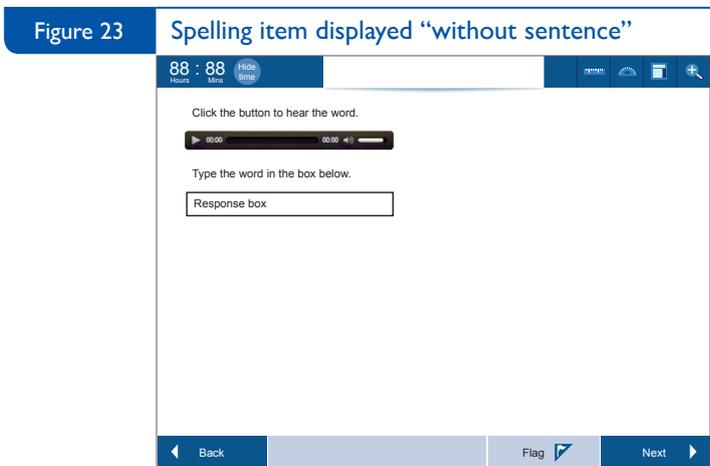
Year level	Item descriptor	Display options	Preference
3	Pushes	Sample 14: With Sentence Sample 15: Without Sentence	Sample 14: 89% Sample 15: 11%
5	Pushes	Sample 14: With Sentence Sample 15: Without Sentence	Sample 14: 96% Sample 15: 4%
7	Disturb	Sample 11: With Sentence Sample 12: Without Sentence	Sample 11: 59% Sample 12: 41%
9	Disturb	Sample 12: With Sentence Sample 13: Without Sentence	Sample 12: 71% Sample 13: 29%

While most students preferred having the sentence written out, many of them thought the sentence should appear above the recording controls (which was not presented as an option in the samples). They wanted to read the sentence first to know for what information they should be listening. Many students at Year 9 said they would click play or listen to the sound as the first action they would take, but mainly because it’s at the top of the page and the first thing they see.



22-Years 3 and 5 sample 14; Year 7 sample 11; Year 9 sample 12

Students at Year 7 were concerned with the placement of the instructions line above the audio player in sample 11 (with the sentence) – [Figure 22](#) and many of them indicated a preference for sample 12 (without the sentence) – [Figure 23](#) for this reason, even though their comments indicated that they would like to see the sentence written out. Some students at Years 7 and 9 noted the difference in wording in the instructions between the samples (“click the arrow” versus “click the button”) and thought “button” was unclear because there were several buttons. Students in Years 7 and 9 also expressed concern about the use of the bracket style arrow (>) as a symbol for the play button in the instructions line and found this confusing. Suggestions were either to write out “play button” or to use an image of the actual button.



23-Years 3 and 5 sample 15; Year 7 sample 12; Year 9 sample 13

## 4.4 General interface observations

### Navigation

Although students generally understood that the “Next” and “Back” buttons would take them to the next question on the test, there was some confusion as to the function of these buttons relative to the numbered buttons associated with the passage text interface. Additionally, some students suggested alternate navigation schemes that included showing thumbnail icons of the questions in a panel on the right-hand side of the screen, or numbered buttons at the top of the screen for each of the questions on the test to allow for direct navigation between questions.

### Expansion tab

The purpose and functionality of the expansion tab within the passage text were not well understood by students. Some students were able to correctly identify that the tab would expand the passage text and hide the question, but did not think this was clearly explained. Some students thought this would allow them to move to the next page of the text. Others thought it would move them to the next question associated with the current passage text. Additionally, students who were able to identify the correct purpose were unclear as to whether the expanded text would include text beyond what they could see on the current page, or whether it would simply expand the current page text wider across the screen. Students suggested that the blue arrow was not clearly visible and suggested making it a different colour, enlarging it, or adding a label to make it more noticeable.

### Tools

Students encountered tool icons in samples for all three domain areas (reading, spelling, and numeracy). Though the tools themselves were not the focus of the study, students were asked their impressions of the tool icons and what they thought they would do. Some students did not immediately notice the tool icons and suggested that they be made larger, or moved to the top-left side of the screen (where the clock is located), so that they would be more likely to see them. Students also suggested providing an explanation of the tools prior to beginning the test (perhaps in some sort of tutorial), or to add instructions which could be accessed through a question mark next to the tool icons within the test. Students were generally confused by the inclusion of numeracy tools within the spelling samples and suggested they be removed.

Within the samples for the reading domain, the only tool icon pictured was the magnifier. Students also saw the magnifier tool with both the spelling and numeracy samples. Some students had some confusion around the icon for this tool and thought it was a search feature. Other students understood that it was a magnifier tool but were uncertain how it would work. Some students speculated that it might bring up a magnifying glass to allow the student to choose what would be magnified. Other students thought it might zoom-in and magnify the entire item. Some students mentioned the existing functionality within their browser or on their tablet to zoom and were not sure how this zoom tool would be different.

Within the samples for the numeracy domain, students saw tool icons for the ruler, protractor and calculator. Students generally understood that clicking on each of these tool icons would bring up the associated tool, but were frequently uncertain how the tools themselves would work. Additionally,

Year 3 students were not always certain about what some of the tool icons were. For the ruler and protractor, students generally understood that they could be used to measure things (e.g. length and angles) and thought that they would click and drag the tool to what they wanted to measure. Some students speculated on how they might rotate the tools, and were concerned that they would be hard to control and might be time consuming to use. Some Year 3 students indicated that they thought the ruler could be used as a straightedge to underline things. For the calculator, there was also concern about whether students would be allowed to use the tool at all or on some items and whether doing so might be considered cheating.

### Timer display

Students generally understood that the timer display would communicate to them how much time they had left on the test. The hide time button was positively received as students were frequently concerned about the timer causing stress. Some students suggested moving the timer to the top-right side of the screen next to the tool icons so that it would not be the first thing that they see. Other students suggested moving the timer to the top-middle of the screen and having tools on either side of it. Some students suggested that in addition to the time it would be good to show how many questions were left.

### Flag for review

Year 3 students were generally uncertain what the “Flag” button would do, with one student speculating that it would show them the flag of the country they were in. Older students had a better understanding of the functionality correctly identifying that it would let them “flag” questions that they could revisit later in the test. However, one student suggested that the button would alert the teacher and send a message to her computer and another suggested it would allow the student to “flag” inappropriate content. Students thought that the label on the button should be clearer and suggested “bookmark” as a possible alternative. Some students did not think they would use the flag button.

# 5 Conclusions

## 5.1 Overall findings and recommendations

This study summarised student preferences for a number of important features of onscreen display for item content. This section will review the findings for each of these features and provide recommendations based on student feedback. Table 11 provides a summary of recommendations for each layout option separately by year level. In general, recommendations are consistent across year levels, but there are two key areas (passage interface and line spacing) where different recommendations are made for younger and older students. Elaboration on, and rationale for each recommendation, as well as suggestions for improving upon the recommended option are discussed below.

**Table 11** Summary of recommendations

Layout Option	Year 3	Year 5	Year 7	Year 9
Image placement (relative to passage text)	Right or bottom of text			
Image placement (relative to item components)	Right or inline (with introductory text above)			
Image size	Larger	Larger	Larger	Larger
Passage interface (single passage)	Paging	Paging	Scrolling	Scrolling
Passage interface (multiple passages)	Tabbed Passages	Tabbed Passages	Tabbed Passages	Tabbed Passages
Line spacing	Double spacing	Double spacing	1.5 spacing	1.5 spacing
Line width	Full-column width	Full-column width	Full-column width	Full-column width
Response option layout	Horizontal	Horizontal	Horizontal	Horizontal
Font	Helvetica 16 pt	Helvetica 16 pt	Helvetica 16 pt	Helvetica 16 pt
Spelling items	With sentence	With sentence	With sentence	With sentence

### Image placement and size

Although students’ eye movement across the screen content was not tracked in this study, the recognition by students of information importance relative to image placement when reading from left to right and the general preferences for image placement suggest that students may be using something similar to the F-layout (Jones, 2012; Nielsen, 2006) to scan test content presented onscreen.

### Image placement (relative to passage text)

Students generally found images within passage texts to be less important than the passage text itself, and wanted to read the passage text first, to set the context for interpreting the information presented in the image. While some students preferred a strategy of using the images as a pre-organiser for the passage text, the majority of students preferred a strategy which used the images to reinforce what they had already read. Additionally, since the information presented in these images was not necessary to answer the question, it should be given a position of lower information priority relative to screen layout. As such, it is recommended that images displayed alongside passage text should be placed to the right or at the bottom of the text, so that students would read the passage text before viewing the images.

### Image placement (relative to item components)

Unlike images associated with passage texts, the content of the images associated with individual items tended to provide necessary information to answer the questions. However, in these situations, students expressed a desire to have pre-organising context prior to the image noting that it was logical to go from left to right and read the question first, then look at the image. Students thought this was a very sequential and logical presentation and that seeing the image first without having read the introductory text or the question itself might be confusing. As such, it is recommended that large images should be placed in the right-hand panel with the question and responses in the left-hand panel. Images which can be made small enough (yet still readable) to be placed with the question in the left-hand panel should be preceded by introductory text, so that the student has some pre-organising context relative to the information contained in the image.

### Image size

The size of an image was also an important factor to students and, as outlined above, interacted with preferences for image placement. In general, students had a preference for larger images especially when the image contained information important to answering the question (e.g. a graph or data table). However, students generally preferred smaller images when the size of the image interfered with the preferred presentation of other key elements within the item (e.g. students did not like having to use the expansion tab to view the item).

It is therefore recommended that images which contain information necessary to answer the question should be large enough for students to easily read the information, but not so large that it displaces other elements of the item.

## Passage interface options

### Single passage

With regard to navigating the passage interface, student preferences were mixed, with younger students preferring a paging interface and older students preferring a scrolling interface. Reasons cited by students for preferring one interface over the other were largely related to their beliefs about ease and efficiency for locating information needed to answer the questions. These results are consistent with those of Dillon (1992) in suggesting that experienced users (in this case older students) were more comfortable using

scrolling than paging to navigate through a text. However, it appears that a paging interface might be more suitable for younger students. This might reflect the heavier reliance of less experienced readers on the structural representation of text (i.e. the concept of pages) to help them locate information they have read. It is therefore recommended that a paging interface should be offered for students in Years 5 and younger and a scrolling interface for older students. For the paging interface, the paging controls should be made larger and more clearly differentiated from other controls in the item (“next”/”back” buttons and the expansion tab arrow). For the scrolling interface, the scroll bar should be made thicker and more apparent to students. Schools should make external mice available to students testing with a scrolling interface on a laptop computer.

### Multiple passages

Students found the tab functionality to navigate between multiple passages easy to recognise and straightforward to use. As such, it is recommended this display be used to allow students to navigate between multiple passages within the same passage text area.

## Item layout and presentation

### Line spacing

Student preferences for line spacing suggest that while younger students preferred double spacing, older students were fairly split in their preferences for double spacing versus 1.5 spacing with a slight majority preferring the 1.5 spacing. Preference for 1.5 spacing amongst the older students reflected both their impression of the text readability, as well as their desire to keep page count to a minimum. It is worth noting that even amongst younger students some thought the double spacing introduced too much space between lines and suggested a compromise line spacing (something between 1.5 and double spacing). Double spacing is recommended for use with Years 5 and younger. The length of passage texts for these year levels should be short enough that the extra space should not result in an unreasonable number of pages for students to navigate. Double spacing may also be considered for older year levels if a consistent standard is desired across the assessment program. However, the impact to the number of pages of text should first be evaluated for passage texts in the respective year level item banks. If the use of double spacing results in an unreasonable page length for some passages, these passages may need to be dropped from the bank. 1.5 spacing is recommended for older year levels and provides an alternative that keeps the number of pages of passage text to a minimum. The impact of 1.5 spacing on passage text readability can be expected to be less of a concern for more experienced readers who are more familiar with this line spacing and have more experience implementing strategies for reading text onscreen.

### Line width

Results for line width suggested that students at all year levels preferred the full-column width over the half-column width—although preferences were fairly close for Year 3 and 5 students. Preference for the full-column width reflected beliefs about the ease and efficiency of reading the text and (as with line spacing) the desire to keep page count to a minimum. It is therefore recommended that full-column width be used rather than half-column width for passage texts. This agrees with the preference of the majority of

students and minimises the number of pages students have to navigate. A review of passage texts should be undertaken once they are formatted in the onscreen interface, to look for issues where sentences wrap across pages and to force pages to end with a complete sentence.

### Response option layout

Students across all year levels preferred a horizontal layout of response options. Students' preference for the horizontal layout of response options was largely driven by efficiency in use of the onscreen space as well as ease of comparison to the information in the question stem. It is therefore recommended that a horizontal layout for displaying response options should be used rather than a vertical layout. Radio buttons should be located either above or below response options, in a horizontal layout to avoid confusion of having radio buttons between response options.

### Font

Across year levels, students had a preference for the consistent use of the Helvetica font. This was largely driven by a desire to have the same font for all elements of the item and by a general dislike of the Palatino font type. It should be noted that although students disliked this font, they did not think it negatively impacted their ability to read and understand the item.

It is recommended that a consistent font type should be used throughout the elements of an item (e.g. item stem, images, and response options) and that Helvetica should be used rather than Palatino.

### Spelling items

Students at all year levels had a preference for the inclusion of a written sentence for spelling items. This reflected a desire for reinforcement of the information through a second channel as well as concerns over the ability to understand the audio narration. It is therefore recommended that Spelling items be presented with the sentence, both read in the audio file and written out onscreen. The written out sentence should be placed above the audio controls. The instructions line should be revised to improved clarity relative to the play button (button versus arrow and bracket arrow versus actual image).

### General interface issues

Students appeared to apply similar logic (importance of information) to the layout of the test interface, in general, as they did when evaluating the placement of images. Students felt that tool information was more important than information about time remaining in the test. While time remaining is certainly an important factor in a testing situation, students felt it served as a distraction and were concerned about the potential for negative impact on their performance. As such, it is recommended that the placement of timer controls and tool icons could be flipped to better reflect the level of importance that students attribute to these features. In addition, instruction should be offered on tools; either through a tutorial prior to the test, help screens within the test, or a combination of both.

Other functionality within the test interface should be clearly identified for students. For example, the label for the “Flag” button could be revised to be clearer to student users. Additionally, the need for and usability of the expansion tab should be evaluated. It is unclear to what extent this functionality is needed. If it is retained, its function should be made clearer to students through improved labelling and button design.

## 5.2 Study limitations

The current study yielded many valuable observations and should generally support strong recommendations to ACARA for online layouts of passage text and items. However, as with all research studies conducted in school settings, some limitations were observed relative to participation, materials (relative to the display samples), and timing. With regard to participation, the number of school and student participants was somewhat less than desired at some year levels, with Year 7 having the fewest participants (4 schools and 19 students). This might limit the degree to which conclusions should be drawn about specific year levels. With regard to the materials, the display samples used in this study were primarily static images that, while displayed onscreen, did not allow students to interact with the functionality represented. Instead students had to speculate about how the interactions would work. Student preferences may have been different if they had been able to click the buttons and experience the items as they would under operational testing conditions. Additionally, some functionality and display options which could have been evaluated with interactive samples could not be included in the study.

The display of static images to students differed somewhat across school locations. In most locations students each had their own device and were able to navigate between the samples and see each one up close. However, in some locations students had to share devices with other students or look at the device used by the study facilitator. This may have limited their ability to evaluate the samples in as much detail as they would have liked. The devices used by students to view the samples in this study also varied across locations and sometimes multiple device types were represented within a class. This variability was in some ways “by design” to allow student comments to be captured relative to a variety of devices, but was not implemented in a controlled enough fashion to allow strong conclusions to be drawn about displays relative to devices.

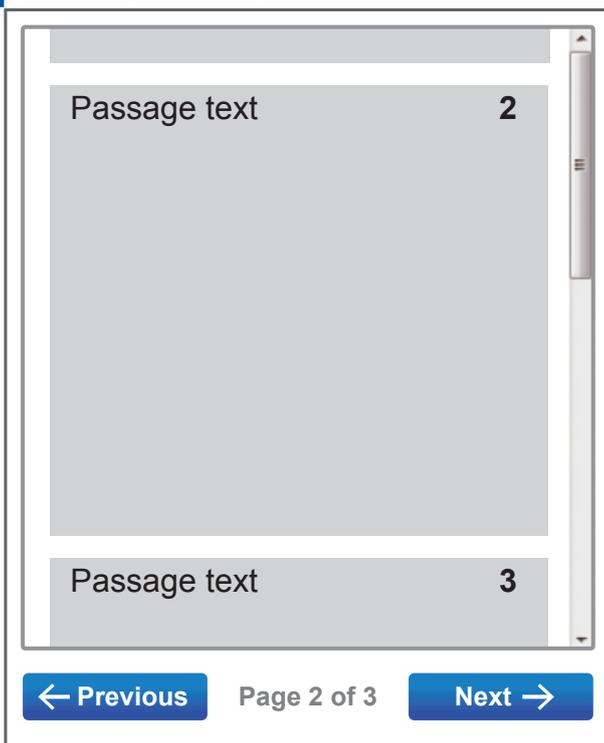
Lastly, with regard to timing, students were available for approximately one hour of participation and were asked to review and provide feedback on as many as 24 different sample layouts during this time. This limited the time that students could spend on any given sample set and, in some cases, meant that some sample sets were skipped by study facilitators.

## 5.3 Further research

Given the differences across year levels in preferences for paging versus scrolling interfaces for passage text (as well as the strength of student opinion on this topic), a hybrid scrolling/paging interface that could be used across all year levels should be evaluated. [Figure 24](#) shows an example of what a hybrid interface might look like. This interface has paging controls similar to the numeric buttons used in this study (though

here labelled “next” and “previous”) and also has a scroll bar control similar to the scrolling interface evaluated in this study. In this example, page breaks are maintained which would help students who find it easier to locate information based on page location. However, students would be able to use either the paging controls or the scroll bar to navigate through the pages as per their preference. Usability research would need to evaluate specific interface elements, but a successful design around this concept would allow for the use of a single interface for all year levels and avoid students having to learn new functionality in later year levels.

**Figure 24** Hybrid paging/scrolling interface



Additional research could also be conducted on the number of pages it is reasonable to expect students to read at each year level. Although students in the current study were asked their preferences for the number of pages, students were almost uniformly in favour of reading as few pages as possible (one page was the preference for many younger students). It is likely that student preferences in this regard may differ from their actual capabilities and research to evaluate how the number of pages may impact student reading comprehension, engagement and fatigue and affect overall test outcome, will help better inform decision making.

## 6 References

- Bernard, M. L., Chaparro, B. S., Mills, M. M., & Halcomb, C. G. (2003). Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal of Human - Computer Studies*, 59, 823-835. doi:10.1016/S1071-5819(03)00121-6
- Bernard, M. L., Fernandez, M., & Hull, S. (2002). The effects of line length on children and adults' online reading performance. <http://usabilitynews.org/the-effects-of-line-length-on-children-and-adults-online-reading-performance/>
- Dillon, A. (1992). Reading from paper versus screens: a critical review of the empirical literature. *Ergonomics*, 35(10), 1297-1326.
- Dyson, M. C. (2004). How physical text layout affects reading from screen. *Behaviour and Information Technology*, 23, 377-393.
- Dyson, M. C., & Kipping, G. J. (1998). The effects of line length and method of movement on patterns of reading from the screen. *Visible Language*, 32(2), 150-181.
- Hojjati, N., & Muniandy, B. (2014). The Effects of Font Type and Spacing of Text for Online Readability and Performance. *Contemporary Educational Technology*, 5(2), 161-174.
- Jones, B. (2012). Understanding the F-Layout in Web Design. Retrieved from Web Design website: <http://webdesign.tutsplus.com/articles/understanding-the-f-layout-in-web-design--webdesign-687>
- Kim, H., & Kim, J. (2013). Reading from an LCD monitor versus paper: Teenagers' reading performance. *International Journal of Research Studies in Educational Technology*, 2(1), 1-10.
- Ling, J., & van Schaik, P. (2007). The influence of line spacing and text alignment on visual search of web pages. *Displays*, 28(2), 60-67. doi: <http://dx.doi.org/10.1016/j.displa.2007.04.003>
- Loh, C. S., Branch, R. M., Shewanown, S., & Ali, R. (2002). *The effects of text spacing after the period on time for on-screen reading tasks*. Retrieved from [http://www.csloh.com/research/pdf/2002\\_IVLA.pdf](http://www.csloh.com/research/pdf/2002_IVLA.pdf)
- Mills, C. B., & Weldon, L. J. (1987). Reading text from computer screens. *ACM Computing Surveys*, 19(4), 329-358.
- Nielsen, J. (1997). How users read on the web. Retrieved from Nielsen Norman Group website: <http://www.nngroup.com/articles/how-users-read-on-the-web/>
- Nielsen, J. (2006). F-shaped pattern for reading web content. Retrieved from Nielsen Norman Group website: <http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/>

# Appendix A

## Example interview protocol

<p><b>Interview protocol</b></p> <p><b>Year 5</b></p> <p><b>Instructions:</b> Use this interview protocol to record your observations of students' responses and reactions to the item layouts.</p>	
School reference	
Year level	
Devices being used (tablets, laptops etc)	
Number of students using each device	
Date	
Invigilator/Researcher	

Purpose Time	Questions	Notes
<b>Introduction</b>		
<b>8 min Total</b>		
Welcome (2 min)	<p><i>Before participants arrive, make sure sticky notes and pens (or some other medium for students to respond with) are placed at each desk. Have name tags/name tents available for students to make.</i></p> <p>Welcome participants Introduce yourself Thank everyone for taking time to participate</p>	
Topic (2 min)	<p>Today we will be talking about test questions and how they look onscreen. We'll be showing you some different ways of presenting the questions and asking for your feedback.</p>	
Guidelines/Ground Rules (2 min)	<p>It is important that we follow a few rules:</p> <p>Only one person speaks at a time. There may be times when you want to jump in while someone else is talking, but it is important that you wait until they have finished.</p> <p>There are no right or wrong answers. We just want to hear what you are thinking.</p> <p>You do not have to speak in any particular order.</p> <p>When you do have something to say, please do so. There are many of you in the group and we want to hear from all of you. Please tell us what you think, and it's okay if what you think is different from what someone else thinks. We do not all have to agree.</p> <p>We want you to share all comments -- whether they are positive or negative. Sometimes, negative comments can be the most helpful!</p> <p>You've probably noticed the microphone. We're tape recording the session because we don't want to miss any of your comments. People often say very helpful things in these discussions and we can't write fast enough to get them all down.</p> <p>We will be on a first name basis, and we won't use any names in our reports. We will keep your answers private.</p> <p>Does anyone have any questions?</p>	
Focus 2 min	<p>So we're going to look at some items and I'm going to ask you for feedback.</p> <p>It is important that you don't focus too much on the questions themselves, but rather how the questions look onscreen. What you will be looking at today is not a final product; in many cases they are just screenshots of what you might see in a final product. So things like buttons or scroll bars will not work in a lot of the examples.</p> <p>We want you to be thinking about things like: is it easy to read the text? Do you see any issues with scrolling? Does the placement of things like images or questions make sense?</p>	
	<p>Let's get started.</p>	

<b>Layouts: Bath or Shower (2 layouts; 9-10; samples 1-2)</b> (5 min)	
<b>Focus:</b> Spacing between lines (single spaced/double spaced) <b>Guiding Research Question(s):</b> Does the spacing between lines (single spaced/double spaced) make a difference to students?	
<b>Let's look at sample 1.</b> <i>(Have participants look at the first layout of Bath or Shower). Pause to give them time to view both).</i>	
Guiding Question(s):	What do you think? What do you see? Just for fun, what do you think the answer to the question is? What do you think the different buttons do? <i>If they don't mention the blue "Back" and "Next" arrow buttons,</i> What do you think would happen if you clicked the blue arrow buttons? <i>Note: We do not want to focus on the interface. After the participants have a chance to share their opinions of the interface, thank them for their input and explain that the rest of the time will be devoted to the questions themselves.</i> Thank you for sharing your thoughts about the different buttons and how the screen looks. The rest of the time, we are going to focus on how the questions look on the screen. <i>Note: If students return to talking about the interface or buttons, gently refocus them on the questions.</i>
<b>Now let's look at sample 2.</b> <i>(Have participants look at the second layout of Bath or Shower).</i>	
Probe:	What do you think? What do you see? <i>If participants have not already mentioned the differences between the two (single versus double spaced text), ask:</i> What are the differences between samples 1 and 2? Feel free to go back to look at both samples again. <i>If participants do not realise the difference, explain that the difference between the two is that the first is single spaced and the second is double spaced.</i> Does the extra spacing between the lines in the second one make it better or worse? Explain. <i>Probe for the specific ways that it is better or worse.</i>
Ending Question/Activity:	Let's pretend you have to answer this question on a test at school. <b>On your sticky note</b> , write which sample you prefer (sample 1 or 2). <b>It's ok to go back and look at both versions if you would like.</b> Be ready to explain why. Which do you like better and why? <i>Have participants share.</i>

<p><b>Layouts: Baby Echidnas (3 layouts; 19-21; samples 3-5)</b> (5 min)</p>		
<p><b>Focus:</b> Images within text block; beneath text block; full screen  <b>Guiding Research Question(s):</b> Does it matter where images are placed in relationship to the text? Do students understand where the questions sit if the text/images are full screen?</p>		
<p><b>Let's look at sample 3.</b>  <i>(Have participants only look at the first layout of Baby Echidnas to start).</i></p>		
Guiding Question(s):	<p>What do you think?                  Does anything bother you with the way the information and pictures are displayed?</p>	
Probe:	<p>How would you change it to make it better?  <i>If necessary, follow-up with:</i>                  What could you change to make it easier to read? Easier to understand?</p>	
<p><b>Now let's look at the sample 4.</b>  <i>(Have participants look at the second layout of Baby Echidnas).</i></p>		
Probe:	<p>What do you think of this version?  <i>If necessary,</i>                  Is it easier to read or harder to read?  <i>If participants have not already mentioned the differences between the two (image placement), ask:</i>                  What's the difference between the two versions?  <i>Note: If participants do not realise the difference, explain that the difference is the placement of the pictures.</i></p>	
<p><b>Now there's one more version for us to look at. Go to sample 5.</b>  <i>(Have participants look at the third layout of Baby Echidnas).</i></p>		
Probe:	<p>Tell me what's going on here.                  Where would you find the questions for this text?                  What do you like, don't you like?</p>	
Ending Question/Activity:	<p><b>On your sticky note</b>, write which sample you prefer (sample 3, 4 or 5). <b>It's ok to go back and look at the other two versions if you would like.</b> Be ready to explain why.  <i>Go around and ask participants to share.</i></p>	

<b>Layouts: International Day (3 layouts; 37-38 and 39-40; samples 6-9)</b> (5 min)	
<b>Focus:</b> Default line width; pagination <b>Guiding Research Question(s):</b> How does line length or multiple pages affect the ability to read the text passages?	
<b>Let's move on. Look at samples 6 through 9.</b> <i>(Have participants look at both layouts of International Day). Pause to give them time to view both. Note: there are 2 samples, but 4 screens for International Day).</i>	
Guiding Question(s):	What do you think?
Probe:	What did you like? What didn't you like? <i>Note: If participants have not already mentioned the differences between the two (line length), ask:</i> What's the difference between the two versions? <i>Note: If participants do not realise the difference, explain that the difference between the two is that the first has longer line length than the second.</i> Let's assume that there's more text/words than you can see here. How would you get to it? Does that make sense? <i>Note: If participants do not understand, explain that the buttons labelled 1 and 2 move you to the next page.</i> Let's pretend you had to read multiple pages of text. How many pages of text would you be okay with having to read? (2-3? 5-6?) How would line length affect the number of pages? Does that matter?
Ending Question/Activity:	Let's pretend you have to answer this question on a test at school. <b>On your sticky note</b> , write which sample you prefer (samples 6-7 or 8-9). <b>It's ok to go back and look at both versions if you would like.</b> Be ready to explain why. Which do you like better and why? <i>Have participants share.</i>

<b>Layouts: How to Help an Injured Bat (33-34, 35-36; samples 10-11, 12-13)</b> (5 min)	
<b>Focus:</b> Image to the right of text; image to the left of text; Pagination required <b>Guiding Research Question(s):</b> Does it matter where images are placed in relationship to the text? What do students think of having to tab to a different page?	
<b>Let's look at samples 10-13.</b> (Have participants look at both layouts of <i>How to Help an Injured Bat</i> ). Pause to give them time to view both. Note: there are 2 samples, but 4 screens for <i>How to Help an Injured Bat</i> .	
Guiding Question(s):	What do you think? What do you see?
Probe:	What did you like? What didn't you like? <i>Note: If participants have not already mentioned the differences between the two (line length), ask:</i> What's the difference between the two versions? <i>Note: If participants do not realise the difference, explain that the difference is that the image is on the right in the first sample and on the left in the second sample.</i> Let's assume that there's more text/words than you can see here. How would you get to it? Does that make sense? <i>Note: If participants do not understand, explain that the buttons labelled 1 and 2 move you to the next page.</i> Does it matter whether the image is on the left or right? Why?
Ending Question/Activity:	Let's pretend you have to answer this question on a test at school. <b>On your sticky note</b> , write which sample you prefer (samples 10-11 or 12-13). <b>It's ok to go back and look at both versions if you would like.</b> Be ready to explain why. Any other thoughts? <i>Ask for responses.</i>

<b>Layouts: X00069295 (Spelling) (2 layouts; 46-47; samples 14-15)</b> (5 min)	
<b>Focus:</b> Inclusion or exclusion of sentence in spelling questions <b>Guiding Research Question(s):</b> Do participants prefer to have a sentence included for spelling questions?	
<b>Now let's look at sample 14.</b>	
Guiding Question(s):	What do you think you would do first? <i>If students say they'd click the play button say, "if you click the play button, this is what you would hear:</i> "Spell <b>"pushes"</b> . The boy gets tired when he <b>pushes</b> his bike up the hill. <b>Pushes"</b> <i>If students do not recognise the play button, point it out and say, "if you click the play button, this is what you would hear:</i> "Spell <b>"pushes"</b> . The boy gets tired when he <b>pushes</b> his bike up the hill. <b>Pushes"</b> What do you think? Does the way this question looks/works make sense to you? If not, how would you change it?
<b>Now let's look at sample 15.</b>	
Probe:	What do you think you would do first? <i>If students say they'd click the play button, say, "if you click the play button, this is what you would hear:</i> "Spell <b>"pushes"</b> . The boy gets tired when he <b>pushes</b> his bike up the hill. <b>Pushes"</b> <i>If students do not recognise the play button, point it out and say, "if you click the play button, this is what you would hear:</i> "Spell <b>"pushes"</b> . The boy gets tired when he <b>pushes</b> his bike up the hill. <b>Pushes"</b> <i>Note: If participants have not already mentioned the differences between the two (inclusion or exclusion of sentence), ask:</i> What's the difference between the two versions? <i>Note: If participants do not realise the difference, explain that the difference between the two is that the first includes a sentence while the second does not.</i>
Ending Question/Activity:	Let's pretend you have to answer this question on a test at school. <b>On your sticky note</b> , write which sample you prefer (sample 14 or 15). <b>It's ok to go back and look at both versions if you would like.</b> Be ready to explain why. Which do you like better and why? <i>Have participants share.</i>

<p><b>Layouts: 5E_05 (Paper Folding) (2 layouts; 67-68; samples 16-17)</b> (5 min)</p>		
<p><b>Focus:</b> Images within text block; above text block  <b>Guiding Research Question(s):</b> Does it matter where images are placed in relationship to the text?</p>		
<p><b>Now go look at samples 16 and 17.</b>  <i>(Have participants look at both layouts of Paper Folding). Pause to give them time to view both).</i></p>		
Guiding Question(s):	<p>Take a minute to look at both versions.                  Does anyone notice anything different on this page?  <i>Note: If necessary, point out the math toolbar at the top-right.</i>                  What do you think of this tool bar? What do you think the buttons do?</p>	
Probe:	<p>What are your thoughts?  <i>Note: If participants have not already mentioned the differences between the two (images within text block or above text block), ask:</i>                  What's the difference between the two versions?  <i>Note: If participants do not realise the difference, explain that the difference between the two is that the first has the image within the text while the second has the image above the text.</i>                  Does the question placement make a difference or not?                  Why or why not?</p>	
Ending Question/Activity:	<p>Let's pretend you have to answer this question on a test at school. <b>On your sticky note</b>, write which sample you prefer (sample 16 or 17). <b>It's ok to go back and look at both versions if you would like.</b> Be ready to explain why.                  Which do you like better and why?  <i>Have participants share.</i></p>	

<b>Layouts: 5B_03 (Rope) (2 layouts; 64-65; samples 18-19)</b> (5 min)	
<b>Focus:</b> Vertical or horizontal layout for answer choices <b>Guiding Research Question(s):</b> How does placement of answer selections in relationship to questions affect ability to answer questions? Does it matter if answer choices are displayed horizontally or vertically?	
<b>Now look at samples 18 and 19.</b> <i>(Have participants look at both layouts of Rope). Pause to give them time to view both).</i>	
Guiding Question(s):	What do you think?  Does anything bother you with the way the information and pictures are displayed?
Probe:	<i>If necessary,</i> Is it easier to read or harder to read?  <i>Note: If participants have not already mentioned the differences between the two (vertical/horizontal placement of answer choices), ask:</i> What's the difference between the two versions?  <i>Note: If participants do not realise the difference, explain that the difference between the two is that the first lists the answer choices vertically while the second lists the answer choices horizontally.</i>  What about the placement of the question? Does it matter where the question is placed in relationship to the clock and the answers?
Ending Question/Activity:	Let's pretend you have to answer this question on a test at school. <b>On your sticky note</b> , write which sample you prefer (sample 18 or 19). <b>It's ok to go back and look at both versions if you would like.</b> Be ready to explain why.  Which do you like better and why?  <i>Have participants share.</i>

<b>Layouts: 5D_13 (Column graph) (3 layouts; 72-74; samples 20-22)</b> (5 min)		
<b>Focus:</b> Layout 72: graphic inline and size reduced to fit page; Layout 73: graphic on right side; Layout 74: graphic on left side <b>Guiding Research Question(s):</b> Does it matter where images are placed in relationship to the text?		
<b>Let's look at samples 20-22.</b> <i>(Have participants look at all layouts of Column graphs). Pause to give them time to view both).</i>		
Guiding Question(s):	What do you think? What do you see?	
Probe:	<i>Note: If participants have not already mentioned the differences between the two (placement of graph), ask:</i> What's the difference between the two versions? <i>Note: If participants do not realise the difference, explain that the difference between the three is that the first has the graph in the middle of the text, the second has the graph on the right, and the third has the graph on the left.</i> Does the placement or the graph make a difference? Why/why not?	
Ending Question/Activity:	Let's pretend you have to answer this question on a test at school. <b>On your sticky note</b> , write which sample you prefer (sample 20, 21 or 22). <b>It's ok to go back and look at the other version if you would like.</b> Be ready to explain why. Which do you like better and why? <i>Have participants share.</i>	

<b>Layouts: Sample 23, Questions 7 and 8</b> (5 min)	
<b>Focus:</b> Scrolling and tabs <b>Guiding Research Question(s):</b> Do participants recognise tabs as a way to move to a different passage? Do participants recognise scrolling as a way to view an entire passage? What do they think of scrolling?	
<p><b>Now let's go to Additional samples.</b></p> <p>Click "Sample 23." You will see a new box pop up. Does everyone see this? <i>Help students navigate to this, if necessary.</i></p> <p>Click the Blue "Start Test Now" button. Then click "Start Section." <i>Pause, make sure students are in the right spot.</i></p> <p>At the top-left, you will see a button that says "Review." Click on that, then click "Question 7."</p>	
Question(s):	<p>Look at Question 7. What do you see? What do you think?</p> <p>Now let's say you're finished reading "Cricket and Cougar" and you want to read the story "Kira-Kira." <b>Think to yourself a minute: How would you get to the story "Kira-Kira"?</b> , raise your hand when you think you know.</p> <p><i>Ask participants to share.</i></p> <p><i>If participants do not recognise that they click the Kira-Kira tab, explain to them that this is how they would navigate to the story.</i></p> <p>What do you think of clicking the tab to get to another story? Is it easy or hard? Why?</p>
<p><b>Now we are going to look at a different question.</b></p> <p>At the top-left, you will see a button that says "Review," click it, then click Question 8.</p>	
Probe:	<p>Look at Question 8. What do you see? What do you think?</p> <p>Can you figure out how to see the whole story (get to the end)?</p> <p><i>Ask participants to share.</i></p> <p><i>If participants do not recognise that they scroll, explain to them that this is how they would navigate through the story.</i></p> <p>What do you think of scrolling to read the story? Is it easy or hard? Why?</p> <p><b>Now, please click out of the Pop-Up window or close it. We are going to go back to the original screen we have been working from.</b></p>

<b>Layouts: Sample 24</b> (5 min)	
<b>Focus:</b> Tabs <b>Guiding Research Question(s):</b> Do participants recognise tabs as a way of scrolling through a passage? Do they prefer scrolling or tabs?	
<p><b>Now you should be back at additional samples. <i>Pause, make sure students are in the right spot.</i></b></p> <p><b>Click "Sample 24."</b> You will see a new box pop up. Does everyone see this? <i>Help students navigate to this, if necessary.</i></p> <p><i>Explain to students that this is just a demo, and that they should ignore all of the PowerPoint stuff at the top, and only look at the slides</i></p>	
Question(s):	<p>What do you see? What do you think?</p> <p>Let's pretend you need to see more of the story. <b>On your sticky note</b>, write what you would do to be able to see the whole story. Be ready to share your answer.</p> <p><i>Ask participants to share.</i></p> <p><i>If participants do not recognise that they click the numbered buttons at the bottom to navigate, explain to them that this is how they would navigate through the story.</i></p> <p>What do you think of clicking the numbered buttons to read the story? Is it easy or hard? Why?</p> <p>Do you like scrolling or numbered buttons better? Why?</p> <p><b>Now, click the right arrow at the bottom to see what it would look like to use the tabs.</b></p> <p>What do you think?</p>
<b>Closing</b>	
(2 min)	
<p>Thank you so much for all of the great ideas and feedback you've given us. We will be taking all of your feedback and putting it together for the product team so that they can implement changes based on what you've shared with us today.</p> <p>(dismiss and provide students with incentive)</p>	