

Multimodal literacy and large-scale literacy tests: Curriculum relevance and responsibility

Len Unsworth, Jen Cope and Liz Nicholls

Institute for Learning Sciences and Teacher Education, Australian Catholic University

ABSTRACT

An educational chasm divides the Australian National Curriculum, which is permeated with detailed requirements for students to develop multimodal literacy, and the very substantially mono-modal literacy of the reading tests of the Australian National Assessment Program in Literacy and Numeracy (NAPLAN). This paper proposes that mandated centralised large-scale testing approaches to literacy assessment, such as NAPLAN, need to be re-thought and reformed to be consistent with curriculum requirements in relation to multimodal literacy and with the multimodal nature of international tests such as Trends in International Mathematics and Science Study (TIMSS) and the Program for International Student Assessment (PISA) and, most importantly, to influence and support teachers in ensuring that students acquire essential multimodal literacy competencies in the twenty-first century.

This paper firstly demonstrates that the interaction of language and image in multimodal texts are integral to the nature of the multimodal literacy that is increasingly needed to negotiate the vast majority of texts we encounter in all domains of our lives, and are reflected in national curriculum documents. Secondly, analyses of international tests such as TIMSS and PISA show the significant proportion of assessment items that specifically address images and image-language relations for correct responses to test items. Thirdly, we provide analyses of the NAPLAN reading tests from 2008–2016, showing the minimal proportion of test items that deal with images and image-language relations. NAPLAN results contrast markedly with the former NSW Basic Skills Tests (BST), which show a significantly higher proportion of test items that dealt with image-language interaction. Fourthly, we reiterate that the paucity of attention to images in NAPLAN demonstrates that this mandated national literacy assessment does not assess multimodal literacy. This signals a strong need for re-thinking and reforming of NAPLAN, and similar large-scale literacy tests so that assessment approaches will more effectively contribute to enhancing the students' multimodal literacy development. Implications of the paper for supportive action by administrators, teachers and researchers towards such reform are briefly noted.

Introduction

There is an educational chasm that divides on the one hand, the Australian National Curriculum English (ACARA, 2018a), which is permeated with detailed requirements for students to develop multimodal literacy, integrating language and images, and, on the other hand, the very substantially mono-modal literacy assessment of the reading tests of the

Australian National Assessment Program in Literacy and Numeracy (NAPLAN) (<https://www.nap.edu.au/>) that pays minimal attention to assessing students' reading of images and almost exclusively restricts its reading comprehension assessment to print. A similar gulf between curriculum expectations and large-scale literacy testing is also evident in other countries, such as the United States and United Kingdom (Unsworth,

2017). This contrasts with the substantial emphasis on multimodal literacy in international tests such as Trends in International Mathematics and Science Study (TIMSS) (<https://timssandpirls.bc.edu/>) and the Program for International Student Assessment (PISA) (www.oecd.org/pisa/). For both supporters and critics of large-scale state or national literacy testing the issue of the nature and extent to which multimodal literacy is addressed in such tests is an important issue. While it is beyond the scope of this paper to pursue the educational and political contestation surrounding such tests, for those who oppose them, the paucity of attention to multimodal literacy endorses arguments for their irrelevance and lack of social and curriculum responsibility, and for supporters. This is clearly a limitation that needs to be addressed.

This paper extends earlier work to show the progressive decrease to the now extremely minimal attention to multimodality in the NAPLAN tests from 2008–2016, widening the gap between curriculum requirements and high-stakes national testing. It provides a hitherto unavailable independent analysis of the extent to which the role of image-language relations is addressed in reading comprehension in the TIMSS, Progress in International Reading Literacy Study (PIRLS) and PISA tests, enabling for the first time, an accessible comparison of the contrasting treatment of multimodal literacy in these tests with the NAPLAN. Our use of a common detailed coding scheme for analysing the various ways that images interact with language to construct the meanings addressed by test items across different tests may also provide teachers with a practical framework for reviewing the learning experiences they design involving image-language interaction and their classroom-based formative assessments of students' multimodal reading strategy development.

There do not appear to be other Australian or international studies that have challenged the curriculum validity of large-scale literacy tests on the basis of their failure to reflect the multimodal literacy requirements of mandated curricula. Popat, Lenkeit and Hopfenbeck (2017) reviewed studies of how findings of the international large-scale assessment studies (ILSAs) informed teaching practice. They found that 'the majority of texts related to policymaking rather than actual interventions and knowledge transfer' (p. 5) and did not make any reference to image-language relations. In fact, image-language relations in international literacy tests seem to have been almost completely neglected in existing research. One paper (Takayama, 2018) discussed the low achievement of Japanese students in the 2003 PISA test, which was partly explained by the inclusion of questions on 'discontinuous' text that

includes, for example, images, graphs and diagrams, since this kind of literacy was not part of the traditional Japanese reading curriculum. The paper noted that Japanese textbooks now include discontinuous texts with the approach to reading in PISA reflected in the national curriculum.

In this paper we argue from our detailed study of NAPLAN that, where large-scale state or nationally mandated literacy tests are conducted, they need to be designed to reflect the multimodal nature of literacy in the curriculum and in society more generally. We provide evidence that this is possible though our analyses of international tests and state tests that preceded NAPLAN. We indicate that for predominantly monomodal tests such as NAPLAN, renovation along these lines is essential, as these mandated tests significantly influence the emphases given to literacy pedagogy in schools, and furthermore that such a renovation, accompanied by related teacher professional learning opportunities, could substantially enhance students' multimodal literacy development.

Firstly, we will briefly note the well-established international consensus among literacy educators, researchers and curriculum authorities that literacy can no longer be thought of in terms of words alone and that images and the interaction of language and image in multimodal texts are integral to the nature of the multimodal literacy that is increasingly needed to negotiate the vast majority of texts we encounter in our personal, social, civic, academic, professional and vocational lives. We will indicate how this is reflected in national curriculum documents and also in the currently advocated pedagogies in the curriculum areas of science and history. Secondly, we will provide analyses of international tests such as TIMSS and PISA to show the significant proportion of assessment items that specifically address images and image-language relations such that effectively comprehending these is necessary for correct responses to test items. In the third section of the paper we will provide analyses of the NAPLAN reading tests every two years from their inception in 2008 until 2016, showing the minimal proportion of test items that deal with images and image-language relations. We will also provide sample analyses of tests conducted prior to NAPLAN, showing, for example, that the NSW Basic Skills Tests (BST) (New South Wales Department of Education and Training, 2005–7) included a very significantly higher proportion of test items dealing with images and image-language interaction. In the fourth part of the paper we note the clear indications from research, that mandated national literacy assessments strongly influence the nature of the taught curriculum in schools

and that a reform and renovation of NAPLAN would be a productive means of contributing to enhancing the multimodal literacy development of our students. To conclude, we consider the need for collaborative research to optimise the construct validity of the assessment of multimodal literacy in order to ensure that the national literacy assessments would be educationally responsible in terms of national curriculum requirements. We propose that such a plan for NAPLAN, if articulated with sustained professional learning opportunities for teachers at the intersection of the national curriculum requirements and national assessment data on multimodal literacy, could contribute to substantial improvements in students' multimodal literacy development.

The significance of image-language integration in contemporary literacy curricula and international literacy assessment

It has long been recognised globally among literacy educators, researchers and curriculum authorities that making meaning from images and language, and their prevalent interaction in multimodal texts, is a significant and increasing requirement in contemporary literacy (Andrews, 2004; Bezemer & Kress, 2008; Hull & Nelson, 2005; Kamil, Intrator, & Kim, 2000; Kress, 2000a, 2000b, 2000c; Leu, Kinzer, Corio, Castek, & Henry, 2013; Luke, 2003; Mayer, 2008; Richards, 2001; Rowsell, Kress, Pahl, & Street, 2013; Russell, 2000). Examples of the significance of image-language interaction in interpreting multimodal texts from the extensive research in this area can be seen in the studies by Hull and Nelson (2005) with narrative and by Bezemer and Kress (2008) with science texts. Hull and Nelson provide an illustrative analysis of a digital multimodal story, 'Lyfe-N-Rhyme', authored by a young man, Randy. They focus on the couplings of images and language, showing how this orchestration transcends the separate contributions of the images and the language. In a 13-second introductory segment concurrently presenting a narration of four sentences and a series of five images, the author verbally communicates his search for identity visually locating this through a succession of images which symbolise African American struggles and Black masculinity. Hull and Nelson point out that the thematic thread running through the succession of images is mapped onto the meanings of the first-person narration, so that it is the linkage of language and image that constructs the meaning of this orientation to the story of a young African American man's search to reconcile personal identity with culture and history. They argue that the powerfully organic connection of the universal themes,

symbolised in the succession of five images with the young author's life and personal identity, could only have been accomplished through the interaction of image and language in what they call 'the multimodal laminate' (Hull & Nelson, 2005, p. 239). Bezemer and Kress (2008) point out the essential role of image-language interaction in constructing an understanding of the digestive system from a school science textbook. For example, in this book the written text conveys the shape of the oesophagus as 'a narrow, muscular tube' (p. 186). But, as Bezemer and Kress note, this does not indicate its shape relative to the other organs involved in digestion, and this relative shape has to be shown in the image. The image however, cannot depict the texture of the oesophagus, which is rendered verbally as 'muscular'. In communicating understanding in multimodal science texts, then, the writing and the image are not simply copies of each other and nor is the image a simplified version of the language. Image and language both offer distinctive epistemological affordances and commitments and interpretation of these image-language ensembles requires an integrative reading strategy for constructing meaning. (For further examples of meaning-making at the intersection of image and language see [Unsworth, 2001, 2006, 2008, 2014a]).

Images occur very frequently and routinely in texts in all spheres of our personal, social, civic, academic, professional and vocational lives. Of particular importance to this paper is the ubiquitous use of images as a fundamental dimension of texts that students read in school and in their extra-curricular activities. Increasingly, such images are not add-ons, but form an integral part of texts that is crucial to their interpretation (Rowsell et al., 2013).

While internationally, government mandated curriculum documents in countries such as Australia, Canada, the U.S., Singapore and Sweden, for example, require literacy pedagogy to address the integration of images and language in multimodal text comprehension and creation (ACARA, 2018a; British Columbia, 2006; New York, 2012; Singapore, 2008; Sweden, 2009), it is apparent that national reading tests in this second decade of the twenty-first century are still not addressing the reality of the prominence of multimodal texts in the lives of students, and this is particularly the case in the U.S., England and Australia (Unsworth, 2017). The extent to which image-language interaction is considered an essential aspect of literacy in the Australian Curriculum: English (ACARA, 2018a) is demonstrated by many references to, for example, 'multimodal texts' (p. 4) and 'contribution of words and images to meaning' (p. 18), as well as 'texts that

incorporate supporting images’ (p. 46) or the ‘analysis of the ways images and words combine’ (p. 113). The English curriculum is very clear that students should understand how the relationships between language and images make meaning in multimodal texts, and that they should be able to utilise these multimodal resources in composing their own texts (see, for example, Content Description Numbers 1661 and 1704, Australian Curriculum: English – Literacy, ACARA, 2018a, p. 45 and p. 102). In Australia, the curriculum areas of science and history also incorporate images as well as language in the disciplinary literacy requirements of the subject area: the Australian Curriculum: Science (ACARA, 2018b), particularly within the Science Inquiry Skills strand, emphasises the use of a variety of methods and tools to observe, represent and communicate scientific ideas, including ‘multi-modal texts’, while many references stipulate ‘drawings’, ‘diagrams’, ‘models’ and creating ‘graphical representations’ (see, for example, Content Description Numbers 060 and 110, Australian Curriculum: Science, ACARA, 2018b, p. 41 and p. 68); the Australian Curriculum: History (ACARA, 2018c), in both content strands, similarly specifies image use in addition to language, such as, ‘using a cross-sectional drawing’, ‘creating a graphic representation’, ‘responding to questions about photographs, artefacts, stories’, and ‘identify the possible meaning of images and symbols in primary sources’ (see, for example, Content Description Numbers 030, 209, 157, and 169, Australian Curriculum: Humanities and Social Sciences: History, ACARA, 2018c, p. 16, p. 26, p. 46, and p. 59); and science and history education researchers are emphasizing the importance of the distinctive multimodal literacy for learning in these subject areas (Derewianka & Coffin, 2008; Oteiza & Pinuer, 2016; Tang, Ho, & Putra, 2016; Tytler, Murcia, Hsiung, & Ramseger, 2017; Tytler, Prain, & Hubber, 2018; van Leeuwen & Selander, 1995).

The inclusion and, in some instances, emphasis on multimodal literacy within several national curricula can easily lead to an assumption that national assessments would similarly address multimodal literacy. Indeed, it could be argued that national tests should be responding to the curriculum requirements in terms of assessing multimodal literacy development (Unsworth, 2014b, 2017; Unsworth & Chan, 2009). Yet, as our analysis of the Australian NAPLAN reading tests (www.nap.edu.au/naplan/reading), taken by Years 3, 5, 7, and 9, will show, despite images being incorporated into almost every reading passage in the stimulus material, very little assessment is directed towards making meaning from the interplay between language and images. This disconnect between the multimodal nature

of curricula and the essentially monomodal nature of national reading tests is surprising, especially given that PISA incorporates 35% of test items in which the interpretation of images is essential to comprehending the text (OECD, 2017), and also that former state-based reading tests, such as the New South Wales Basic Skills Tests (BST), included significant proportions of such test items (Unsworth, 2014b, 2017; Unsworth & Chan, 2008, 2009).

International Assessments: A significant focus on images in test items

Collection of international test data for analysis

Analyses of a number of international tests were undertaken to investigate the proportion of assessment items that specifically addressed images and image-language relations such that effectively comprehending these is necessary for correct responses to test items. The TIMSS tests and the Progress in International Reading Literacy Study (PIRLS) (<https://timssandpirls.bc.edu/>) are both conducted by the International Association for the Evaluation of Educational Achievement (IEA). The IEA is ‘an independent international cooperative of national research institutions and government agencies’ (Mullis & Martin, 2017, pp. 3–4), with offices in Amsterdam and Hamburg, and is directed by the TIMSS & PIRLS International Study Center at Boston College, USA. TIMSS assesses students in Years 4 and 8, every four years, in about 60 countries. In 2019, 70 countries are expected to participate in TIMSS. PIRLS assesses the reading literacy of students in Grade 4 and is conducted every five years; more than 60 countries were expected to participate in PIRLS in 2016 (Mullis & Martin, 2015). The Australian Council for Educational Research (ACER) supports PIRLS to develop items, together with the National Foundation for Educational Research (NFER) in England.

PISA assessments are conducted every three years among 15-year-old students (near the end of their compulsory education) within 72 participating countries and economies of the Organization for Economic Cooperation and Development (OECD). According to the OECD (2017), PISA assesses how students reproduce and apply knowledge in science, reading, mathematics and collaborative problem solving; in 2015, the main focus was science; 540,000 students participated in PISA 2015, which represented 29 million 15-year-olds.

Data analysed from the TIMSS, PIRLS and PISA tests included:

- The released items from the 2011 and restricted use items from the 2015 version of the TIMSS science

tests taken by Grades 4 and 8 in Australia. Many items from TIMSS 2011 science assessments were released onto their website ‘to provide the public with as much information as possible about the nature and contents of the assessment’ (Mullis & Martin, 2013, p. 94); TIMSS 2011 released science items were accessed as PDFs via the TIMSS and PIRLS website (<https://timssandpirls.bc.edu/timss2011/international-released-items.html>). However, the policy on releasing TIMSS items to the public changed for the TIMSS 2015 assessment; permission was requested and granted from the IEA to allow the researchers access to TIMSS 2015 Restricted Use Items, which we received as PDFs for the purpose of our research.

- Released passages and items in reading from the 2011 version of the international PIRLS taken by Grade 4 in Australia were released to the public, similar to TIMSS 2011, and obtained as a PDF from the TIMSS and PIRLS website (<https://timssandpirls.bc.edu/pirls2011/international-released-items.html>).
- Items from PISA focused on the science domain in 2015; five units of example PISA 2015 test questions from the two-hour assessment were accessible via the OECD website (<http://www.oecd.org/pisa/test/>).

The majority of TIMSS test items are standalone individual questions with a few comprising two or three-part questions. TIMSS 2011 science test items totalled 162, and TIMSS 2015 science test items totalled 171. Just under 50 % of TIMSS test items did not contain any type of image. The 2011 PIRLS papers consist of four reading passages, ranging between two and six pages, with each passage containing at least four images; 12–16 questions correspond to each PIRLS text. The answers are divided between multiple-choice options and open-ended responses. PIRLS 2011 consisted of 54 questions. PISA 2015 sample computer-based test items, or units, contain a short written passage and one or more images; between three and six questions correspond to each PISA test unit, and totalled 18. Similar to PIRLS 2011, PISA 2011 questions are ‘a mixture of multiple-choice questions and questions requiring students to construct their own responses’ (OECD, 2017, p. 13).

Data analysis procedures

The assessment items in all the tests were coded as to whether or not obtaining the correct answer entailed reader attention to the images and, if so, the ways in which the image related to obtaining the answer. The following coding categories were developed:

- YES for when the image was essential to answering the question, that is, the answer could only be completed by looking at the image and could not be found in any written language which might be present.
- NO for when the reading text or test item contained an image, but the image was unrelated to the answer to the question, because the answer was only able to be obtained from the written words.
- SUPPORTS was used for when the answer could be found in the written words, but the image helped to support the answer, that is, it provided a visual aid to what was written and might help to infer the answer.
- REFERENCES was created to account for needing to look at an image to answer the question, though only for a written detail such as an object’s name.
- IMAGE IN ANSWER was created because some multiple-choice answers contained visual images. Sometimes the images were only in the multiple-choice answers and sometimes these occurred in addition to an image in the question or reading text.
- NO IMAGE indicates that the question and answer consisted only of written text and contained no image.

In coding the data, we distinguished between pictorial images and words. Thus if a reading text contained, for example, a facsimile of a web page or a book review and written words provided the answer, regardless of whether the written words were ‘continuous, non-continuous, mixed or multiple’ (Thomson, De Bortoli, & Underwood, 2017, p. 99) and the words were set within the context of the web page or review article, it was deemed that the answer relied on words and did not pertain to an image. For TIMSS items, we decided to treat tables and formulae as text rather than images. Graphs and taxonomies on the other hand, which represented, for example, a hierarchical structure, were considered to be an image. We adopted this position for the purpose of being clear, even though some might contest the demarcations.

A coding manual was created, giving category name, definition, and a description of the category (Table 1). This then enabled sample data to be subjected to inter-rater reliability.

Most test items were single-coded, that is, they were only coded with one category: YES, NO, SUPPORTS, REFERENCES, IMAGE IN ANSWER, or NO IMAGE. Some test items, however, were double-coded. Double-coding applied to those items where answers were composed of one or more visual images: the item was first coded as IMAGE IN ANSWER; and, in most cases, the item was also coded as YES, because

Table 1. Coding scheme for analysis of test questions

Category	Definition	Description
YES	Image is essential to answer the question	The answer can <i>only</i> be completed by looking at the image. The answer cannot be found in any of the written text which might be present.
NO	Image is not needed at all to answer the question	The reading text or test item contains an image but the answer can only be found by reading the written words.
SUPPORTS	Image might help to infer the answer	The answer can be found in the written words, although the image content is considered as a visual prompt in conjunction with the text in helping to answer the question. The image might prompt contextual knowledge.
REFERENCES	Image is required for a minor detail	The image content is not needed to answer the question, but the image needs to be referred to in order to find a detail to answer the question, for example, the name of an object.
IMAGE IN ANSWER	The answer contains one or more images	The answer is composed of one or more images.
NO IMAGE	There is no image present in the question or answer	The test item is composed only of written text. No image is present.

the image was essential to answering the question. Occasional items contained an image in the item or question and an image in the answer; these items were also double-coded as YES and IMAGE IN ANSWER, together with a note made so that the number of items for which this occurred could be indicated.

Inter-rater agreement

Two complete tests – NAPLAN 2016 and TIMSS 2011 – were rated by a second coder. Rater agreement was reached at 92.52%, indicating a very high degree of inter-rater coding reliability. Agreement was marginally higher on the TIMSS test overall than the

NAPLAN test. For TIMSS Grade 4 there was 97.23% agreement. Where disagreement occurred, it tended to be split between one coder using the coding category YES or NO and the second coder using the coding category SUPPORTS. In cases where the second coder used the SUPPORTS category, and the first coder the NO category, the support of the visual image was minimal to answering the question in comparison to the image being considered irrelevant to answering the question.

After the second coder's initial ratings, there were two areas of interest raised. Firstly, whereas the coding manual defined the categories of SUPPORTS and REFERENCES to distinguish between the visual image in the former category acting as a prompt to aid in the answer, and the latter requiring the image to be looked at only for a minor detail such as an object name, the second coder initially regarded the REFERENCES code to be the prompt to contextual knowledge, and SUPPORTS in terms of how well the image supported what the written text was saying. After clarification of these differences, and slight revision of the definition and description in the coding manual, the second coder re-rated items coded as REFERENCES, in most instances revising them to the SUPPORTS category.

Assessment of comprehension of image-language interaction in international tests

The results of the analyses for the interaction between image and language in order to answer test questions are shown in the following tables for the 2011 and 2015 TIMSS, 2011 PIRLS, and 2015 PISA international assessment tests.

Table 2 shows that under half the test items in the 2011 TIMSS science test contain no image; in a substantial 51% (37) of the items that do contain images, the image is essential in order to answer the question; in a further 29% (21) of cases, the image acts as a supportive visual aid in answering the question. The combined total of 80% is thus a hugely significant proportion of assessment items which address image-language relations for correct responses to test items.

The results for 2015 TIMSS science test in Table 3 show an overall high proportion (79%) of items which address image-language relations in terms of answering the questions, as for the 2011 TIMSS science test (Table 2). In the 2015 TIMSS science test, the image is essential in correctly answering the question in 42% (37) of items, which is slightly lower than in the 2011 TIMSS science test. The image is supportive in answering the question in 37% (33) of items however, which is slightly higher than in the 2011 test.

Table 4 shows that for the 2011 Progress in International Reading Literacy Study (PIRLS), the results are

Table 2. 2011 TIMSS Science Test: Relationship of images to test items

Year Group	YES image is essential to answer	NO image is not needed to answer	SUPPORTS Image helps to infer answer	REFERENCES Image needs to be referenced for a detail e.g. object name	IMAGE IN ANSWER	NO IMAGE IN Q or A	*Total number of Qs
Year 4	15	3	12	3	5	39	72
Year 8	22	8	9	0	8	51	90
Totals	37	11	21	3	13	90	162

Table 3. 2015 TIMSS Science Test: Relationship of images to test items

Year Group	YES image is essential to answer	NO image is not needed to answer	SUPPORTS Image helps to infer answer	REFERENCES Image needs to be referenced for a detail e.g. object name	IMAGE IN ANSWER	NO IMAGE IN Q or A	*Total number of Qs
Year 4	11	14	11	0	7	38	74
Year 8	26	5	22	0	11	44	97
Totals	37	19	33	0	18	82	171

Table 4. 2011 PIRLS: Relationship of images to test items

Year Group	YES image is essential to answer	NO image is not needed to answer	SUPPORTS Image helps to infer answer	REFERENCES Image needs to be referenced for a detail e.g. object name	IMAGE IN ANSWER	NO IMAGE IN Q or A	*Total number of Qs
Year 4	5	46	4	1	0	0	54

Table 5. 2015 PISA: Relationship of images to test items

Age Group	YES image is essential to answer	NO image is not needed to answer	SUPPORTS Image helps to infer answer	REFERENCES Image needs to be referenced for a detail e.g. object name	IMAGE IN ANSWER	NO IMAGE IN Q or A	*Total number of Qs
15 years	9	4	4	0	1	1	18

quite different to those in TIMSS, despite being administered by the same body – the IEA. All test items in 2011 PIRLS contain one or more images, yet the image is essential to answering the question in only 9% (5) of cases, with the image acting as a support in a further 7% (4) of cases. Image-language relations are therefore addressed to some extent in a total of 16% of test items.

The results of the analysis for the OECD's 2015 Program for International Student Assessment (PISA) in Table 5 illustrate that for those test items which contain an image, 53% (9) questions necessitate looking at the image to answer the question correctly. In a further 23.5% (4) questions, the image serves to

support answering the question. The 2015 PISA assessment test thus addresses image-language relations in an overall total of 73.5% cases, considerably higher than the 2011 PIRLS assessment test, and just a slightly smaller proportion than in the 2011 and 2015 TIMSS assessment tests.

Comprehension of image-language interaction in NAPLAN

In this third section of the paper we note earlier analyses of comprehension of image-language interaction in the NAPLAN reading tests up to 2014 (Unsworth, 2017) and we provide analyses of the NAPLAN reading

Table 6. Proportion of reading test questions involving images in NAPLAN

	Year 3 (%)	Year 5 (%)	Year 7 (%)	Year 9 (%)
2008	5	8	2	4
2010	3	3	8	2
2012	3	5	0	2
2014	0	2.5	2	0

tests from 2015 and 2016. Australian primary and secondary school students sit the NAPLAN reading test every two years in Grades 3, 5, 7 and 9. The NAPLAN replaced Australian state-based tests in 2008. These analyses show that a minimal proportion of test items deal with images and image-language relations.

Table 6 shows the proportions of test questions that required readers to attend to images in the tests administered to students in Years 3, 5, 7 and 9 at two yearly intervals from 2008 to 2014.

In the 2012 NAPLAN test, over the four tests for Years 3, 5, 7 and 9, from a total of 171 questions there were only four questions (2%) for which the images were essential to obtain the correct answer. These four questions across all year levels were based on only three images in texts because some stimulus pages, and the questions about them, are repeated over some year levels.

In the 2014 NAPLAN test, only two questions in the entire reading test over Years 3, 5, 7 and 9, required the students to attend to images in order to answer correctly. One of these questions in the Year 5 test shows in the question booklet an image of a person's foot positioned flat on a bicycle pedal. The stimulus booklet lists five steps for checking that the bicycle seat is in the correct position. The multiple-choice answers

were in the form of images only and a selection needed to be made that matched to step two in the stimulus text, which states: 'Sit on the bike and put your feet on the pedals. Your feet should be flat'.

The second question dealing with an image related to a report of a shipping accident which involved many thousands of floating bath toys being lost in the ocean and scientists tracking where these were washed up ashore as a means of studying ocean currents. The text was accompanied by a world map with red lines showing the paths followed by the bath toys across the oceans. The caption indicated that 'A thicker line represents more toys'. The question required the readers to note where the thickest line was to answer the following multiple-choice item:

According to the map, which of these statements is true?

- More bath toys were found in Europe than Australia
- More bath toys were found in South America than Europe
- More bath toys were found in South America than Australia
- More bath toys were found in Australia than South America

The only two questions involving images in the entire 2014 NAPLAN reading tests across four year levels involved very simple literal comprehension processes. The results for the analyses of the 2015 and 2016 NAPLAN are indicated in Tables 7 and 8. They show whether, and to what extent, an image is integrated into test items for the purpose of answering the question. The results are described below each table.

Table 7 shows that in the 2015 NAPLAN reading test, only seven questions (4%) from the whole test over Years 3, 5, 7, and 9 required students to attend to images in order to answer the question correctly, which is a very low number considering that there was a total

Table 7. 2015 NAPLAN Reading Test: Relationship of images to test items

Year Group	YES image is essential to answer	NO image is not needed to answer	SUPPORTS Image helps to infer answer	REFERENCES Image needs to be referenced for a detail e.g. object name	IMAGE IN ANSWER	NO IMAGE IN Q or A	*Total number of Qs
Year 3	1	34	4	1	0	0	39
Year 5	3	35	1	0	0	0	39
Year 7	2	41	0	0	0	6	49
Year 9	1	46	1	1	0	0	50
Totals	7	156	6	2	0	6	177

Table 8. 2016 NAPLAN Reading Test: Relationship of images to test items

Year Group	YES image is essential to answer	NO image is not needed to answer	SUPPORTS Image helps to infer answer	REFERENCES Image needs to be referenced for a detail e.g. object name	IMAGE IN ANSWER	NO IMAGE IN Q or A	*Total number of Qs
Year 3	2	29	7	0	1	0	38
Year 5	2	30	6	0	0	0	38
Year 7	2	47	1	0	0	0	50
Year 9	1	38	4	0	0	7	50
Totals	7	145	18	0	1	7	176

of 171 test items which contained an image. In a further six questions (3%) over the four year groups, the image might have acted as a visual prompt, although the answer could be found in the written language.

Results of the analysis in Table 8 for the 2016 NAPLAN reading test show similar results to the 2015 NAPLAN reading test, with respect to the image being essential to answer the question correctly in only seven questions (4%). The image acting as a visual prompt occurs in 18 questions (10.5%) in the 2016 NAPLAN test, which is somewhat more than in the 2015 NAPLAN test, although the answer can be found in the written language.

Prior to the introduction of NAPLAN in 2008, mandatory group reading comprehension tests were conducted by each of the Australian States and Territories, usually for Year 3, Year 5 and Year 7 children in government schools. In the State of New South Wales, these tests were called the Basic Skills Tests (BST). As part of a larger study the proportion of test items addressing image–language relations in the Year 3 BST for 2005 and the Year 5 BST for 2005 and 2007 were examined (Unsworth & Chan, 2008, 2009). These proportions are shown in Table 9 and indicate the test items that could be answered from the image alone and those that required the reader to obtain the correct answer by attending to both the image and the text.

The format of the BST tests is very similar to that of the current NAPLAN tests. They consist of coloured stimulus magazines with narrative and informational texts replete with images of various kinds and accompanying multiple-choice comprehension test booklets. There does not appear to be any obvious reason why the NAPLAN and other national reading assessments could not include similar proportions of questions addressing image–language relations as those in the BST.

Implications: A new plan for NAPLAN – re-thinking multimodality in curriculum responsible national reading assessment programs

Analyses in this paper have shown that in terms of addressing image–language relations, the international TIMSS science tests for the years 2011 and 2015, and the international PISA 2015 test, contain far higher proportions of test items for which this occurs than in the NAPLAN tests. In TIMSS and PISA tests analysed for this paper, the image is essential to answer the question correctly in over 40% of cases, while the image supports answering the question in a further significant proportion (23.5–37%). The latter tests thus address image–language relations in approximately 80% of all test items. These figures contrast dramatically with the results of the analyses for NAPLAN 2015 and 2016, and highlight the paucity of attention to images in

Table 9. Proportions of test items involving images in the 2005 and 2007 New South Wales Basic Skills Tests

Data on Test Items and their relation to Images	2005 BST Year 3	2005 BST Year 5	2007 BST Year 5
Total number of images in magazine	24	23	34
Total number of test items	36	46	46
Number of test items involving the use of images	12	15	14
Proportion of test items involving images	33%	33%	30%

NAPLAN compared with the international tests. In NAPLAN 2015, only 4% of questions (out of 171) required students to attend to images, while the image might have acted as a visual prompt in a further 3%. Likewise, in NAPLAN 2016, only 4% of questions necessitated attending to the image to answer the question correctly, with a slightly higher 10.5% of questions in which the image might have served as a visual aid.

NAPLAN's very low proportion of attention to images also contrasts starkly with the Basic Skills Tests (BST), the mandatory reading comprehension test conducted in the state of New South Wales prior to NAPLAN's introduction in 2008: analysis of each year group in the BST assessment test showed that the proportion of test items involving images was 30% or more, a much higher proportion which addressed image-language relations than those in the recent NAPLAN tests. The very low proportions for NAPLAN demonstrate that NAPLAN does not assess multimodal literacy and is hence incompatible with the multimodal national curriculum in English as well as the multimodal nature of the literacy required in school science and history curricula. There are clear implications from this study for the reform of literacy assessment policy to address the misalignment between the multimodal nature of national school literacy curriculum requirements and the essentially mono-modal literacy competences addressed in NAPLAN. A new plan is needed that will support the development of students' capacities to interpret the meanings constructed in the increasingly multimodal texts of the twenty-first century.

A number of studies have established the internationally widespread and constant struggle between high-stakes standardised testing/accountability systems and more learning-centred views of classroom assessment (Berry & Adamson, 2011; Klenowski, 2011), and it is clear from the literature that to a very significant extent high-stakes testing narrows curricular content to what is tested (Au, 2007; Stillman & Anderson, 2011). However, while this is predominantly the case, there is also some evidence that the nature of the effects of high-stakes testing on curriculum is highly dependent on the characteristics of the high-stakes tests themselves (Au, 2007). Policy reform to establish a more curriculum responsible national literacy testing regime seems to be an obvious potential pathway to optimise curriculum implementation and achieve the multimodal literacy outcomes intended for students.

To achieve policy alignment across the literacy curriculum and national testing, and for NAPLAN to become a curriculum responsible resource to support teachers in developing the full range of literacy competences students need for effective learning of

curriculum requirements, as well as to engage fully in the multimodal literate world of the twenty-first century, literacy testing agencies need to engage with current research in multimodal literacy. The study of the BST testing in New South Wales schools, which has been outlined here (Chan, 2010; Chan & Unsworth, 2011; Unsworth & Chan, 2008, 2009), is an initial move in this direction, but the outcomes of that study point strongly to the need for further work in theorizing the nature of image-language relations in constructing meanings in the test materials as well as investigating how these relate to readers' strategies in comprehending the texts. Investment in collaborative research along these lines is essential to devising a much-needed new plan for NAPLAN.

Innovative curricula such as the national Australian Curriculum English (ACARA, 2018a) support the crucial role of schools in mediating the development of twenty-first century literacies to current and new generations of school students. Such curricula recognise the inadequacy of any general presumption that students are developing these literacies informally outside of school, especially as such views sidestep issues of power, ideology and privilege (Bennett, Maton, & Kervin, 2008; Thomas, 2011). If governments mandate centralised large-scale literacy testing, it should, at the very least, support government curriculum initiatives that address the well-established multimodal nature of contemporary literacies. The kind of extreme disjunction between curriculum and assessment demonstrated in this paper, clearly warrants renovation, reform and re-thinking of the bases and approaches to current large-scale literacy assessment. We have sought to crystallise evidence on this particular issue of literacy curriculum and assessment that administrators and teachers can bring to policy debates. We have also sought to provide the kind of practical, accessible analysis of image-language interaction that may be useful to teachers in reflecting on their practices in developing students' multimodal reading strategies. The mismatch in relation to multimodal literacy between international tests like PISA and TIMSS and NAPLAN in Australia, appear to be evident in large-scale literacy tests in countries like the US and the United Kingdom, and between those tests and the respective literacy curricula. As teachers continue to struggle with this persistent, pervasive incongruity, it is imperative that further national and international research seeks to document its impact on teachers' practices and on students' learning and to generate a sound basis for a viable resolution.

References

- ACARA. (2018a). *The Australian Curriculum: English*. Retrieved from <https://australiancurriculum.edu.au/download/DownloadF10>
- ACARA. (2018b). *The Australian Curriculum: Science*. Retrieved from <https://australiancurriculum.edu.au/download?view=f10>
- ACARA. (2018c). *The Australian Curriculum: History*. Retrieved from <https://australiancurriculum.edu.au/download?view=f10>
- Andrews, R. (2004). Where next in research on ICT and literacies. *Literacy Learning: The Middle Years*, 12(1), 58–67.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258–267.
- Bennett, S., Maton, K., & Kervin, L. (2008). The ‘digital natives’ debate: A critical review of the evidence. *British Journal of Educational Technology*, 39(5), 775–786.
- Berry, R., & Adamson, R. (Eds.). (2011). *Assessment reform in education: Policy and practice*. Dordrecht: Springer.
- Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication*, 25(2), 165–195.
- Chan, E. (2010). Integrating visual and verbal meaning in multimodal text comprehension: Towards a model of intermodal relations. In S. Dreyfus, M. Stenglin & S. Hood (Eds.), *Semiotic margins: Meaning in multimodalities* (pp. 144–167). London: Continuum.
- Chan, E., & Unsworth, L. (2011). Image-language interaction in online reading environments: Challenges for students’ reading comprehension. *Australian Educational Researcher*, 38(2), 181–202.
- Derewianka, B., & Coffin, C. (2008). Time visuals in history textbooks: Some pedagogic issues. In L. Unsworth (Ed.), *Multimodal semiotics: Functional analysis in contexts of education* (pp. 187–200). London: Continuum.
- Hull, G., & Nelson, M. (2005). Locating the semiotic power of multimodality. *Written Communication*, 22(2), 1–38.
- Kamil, M., Intrator, S., & Kim, H. (2000). The effects of other technologies on literacy and learning. In M. Kamil, P. Mosenthal, P. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 771–788). Mahwah, New Jersey: Erlbaum.
- Klenowski, V. (2011). Assessment for learning in the accountability era: Queensland, Australia. *Studies in Educational Evaluation*, 37(1), 78–83.
- Kress, G. (2000a). Design and transformation: New theories of meaning. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: Learning literacy and the design of social futures* (pp. 153–161). Melbourne: Macmillan.
- Kress, G. (2000b). Multimodality. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: Literacy learning and the design of social futures* (pp. 182–202). Melbourne: Macmillan.
- Kress, G. (2000c). Multimodality: Challenges to thinking about language. *TESOL Quarterly*, 34(3), 337–340.
- Leu, D., Kinzer, C., Corio, J., Castek, J., & Henry, L. (2013). New literacies: A dual-level theory of the changing nature of literacy, instruction and assessment. In D. Alvermann, N. Unrau & R. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 31765–32703). Newark, Delaware: International Reading Association.
- Luke, C. (2003). Pedagogy, connectivity, multimodality and interdisciplinarity. *Reading Research Quarterly*, 38(10), 356–385.
- Mayer, R. (2008). Multimedia literacy. In J. Corio, M. Knobel, C. Lankshear & D. Leu (Eds.), *Handbook of research on new literacies* (pp. 235–376). New York/London: Erlbaum.
- Mullis, I.V.S., & Martin, M.O. (2013). *TIMSS 2015 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center and International Association for the Evaluation of Educational Achievement (IEA).
- Mullis, I.V.S., & Martin, M.O. (2015). *PIRLS 2016 assessment framework* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center and International Association for the Evaluation of Educational Achievement (IEA).
- Mullis, I.V.S., & Martin, M.O. (2017). *TIMSS 2019 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center and International Association for the Evaluation of Educational Achievement (IEA).
- New South Wales Department of Education and Training. (2005–2007). *Basic Skills Tests*. Sydney: New South Wales Department of Education and Training.
- OECD. (2017). *PISA 2015 Assessment and analytical framework: Science, reading, mathematics, financial literacy and collaborative problem solving* (Revised ed.). Paris: OECD Publishing.
- Oteiza, T., & Pinuer, C. (2016). Appraisal framework and critical discourse studies: A joint approach to the study of historical memories from an intermodal perspective. *International Journal of Language Studies*, 10(2), 5–32.
- Popat, S., Lenkeit, J., & Hopfenbeck, T. (2017). *PIRLS for teachers: A review of practitioner engagement with international large-scale assessment results*. Oxford University Centre for Educational Assessment Report OUCEA/17/1. DOI:10.13140/RG.2.2.10760.01281
- Richards, C. (2001). Hypermedia, internet communication, and the challenge of redefining literacy in the electronic age. *Language Learning and Technology*, 4(2), 59–77.
- Rowell, J., Kress, G., Pahl, K., & Street, B. (2013). The social practice of multimodal reading: A new literacy studies-multimodal perspective on reading. In D. Alvermann, N. Unrau & R. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 32723–33330). Newark, Delaware: International Reading Association.
- Russell, G. (2000). Print-based and visual discourses in schools: Implications for pedagogy. *Discourse: Studies in the Cultural Politics of Education*, 21(2), 205–217.
- Stillman, J., & Anderson, L. (2011). To follow, reject, or flip the script: Managing instructional tension in an era of high-stakes accountability. *Language Arts*, 89(1), 22–37.
- Takayama, K. (2018). How to mess with PISA: Learning from Japanese kokugo curriculum experts. *Curriculum Inquiry*, 48(2), 220–237.
- Tang, K.-S.K., Ho, C., & Putra, G.B.S. (2016). Developing multimodal communication competencies: A case of disciplinary literacy focus in Singapore. In B. Hand, M. McDermott & V. Prain (Eds.), *Using multimodal representations to support learning in the science classroom* (pp. 135–158). Dordrecht: Springer.

- Thomas, M. (2011). *Deconstructing digital natives: Young people, technology, and the new literacies*. New York: Taylor & Francis.
- Thomson, S., De Bortoli, L., & Underwood, C. (2017). *PISA 2015: Reporting Australia's results*. Camberwell, Victoria: Australian Council for Educational Research (ACER).
- Tytler, R., Murcia, K., Hsiung, C.-T., & Ramseger, J. (2017). Reasoning through representations. In M. Hackling, J. Ramseger & H.L. Chen (Eds.), *Quality teaching in primary science education* (pp. 149–179). Cham: Springer.
- Tytler, R., Prain, V., & Hubber, P. (2018). Representation construction as a core science disciplinary literacy. In K.-S. Tang & K. Danielsson (Eds.), *Global developments in literacy research for science education*. Cham: Springer.
- Unsworth, L. (2001). *Teaching multiliteracies across the curriculum: Changing contexts of text and image in classroom practice*. Buckingham, United Kingdom: Open University Press.
- Unsworth, L. (2006). Towards a metalanguage for multiliteracies education: Describing the meaning-making resources of language-image interaction. *English Teaching: Practice and Critique*, 5(1), 55–76. Retrieved from <http://education.waikato.ac.nz/research/files/etpc/2006v5n1art4.pdf>
- Unsworth, L. (2008). Explicating inter-modal meaning-making in media and literary texts: Towards a metalanguage of image/language relations. In A. Burn & C. Durrant (Eds.), *Media teaching: Language, audience, production* (pp. 48–80). Adelaide, South Australia: Wakefield Press.
- Unsworth, L. (2014a). The image/language interface in picture books as animated films: A focus for new narrative interpretation and composition pedagogies. In L. Unsworth & A. Thomas (Eds.), *English teaching and new literacies pedagogy: Interpreting and authoring digital multimedia narratives* (pp. 105–122). New York: Peter Lang Publishing.
- Unsworth, L. (2014b). Multimodal reading comprehension: Curriculum expectations and large-scale literacy testing practices. *Pedagogies: An International Journal*, 9, 26–44.
- Unsworth, L. (2017). Image-language interaction in text comprehension: Reading reality and national reading tests. In C. Ng & B. Bartlett (Eds.), *Improving reading in the 21st century: International research and innovations* (pp. 99–118). Dordrecht: Springer.
- Unsworth, L., & Chan, E. (2008). Assessing integrative reading of images and text in group reading comprehension tests. *Curriculum Perspectives*, 28(3), 71–76.
- Unsworth, L., & Chan, E. (2009). Bridging multimodal literacies and national assessment programs in literacy. *Australian Journal of Language and Literacy*, 32(3), 245–257.
- van Leeuwen, T., & Selander, S. (1995). Picturing 'our' heritage in the pedagogic text: Layout and illustrations in an Australian and a Swedish history textbook. *Journal of Curriculum Studies*, 27(5), 501–522.

Len Unsworth is Professor in English and Literacies Education in the Institute for Learning Sciences and Teacher Education (ILSTE) at the Australian Catholic University. His recent co-authored books include *Functional Grammatics: Reconceptualising Knowledge about Language and Image for School English* (Routledge, 2017) and *Reading Visual Narratives* (Equinox, 2013). *English Teaching and New Literacies Pedagogy: Interpreting and Authoring Digital Multimedia Narratives* (Peter Lang Publishing, 2014) was co-edited with Angela Thomas.
Email: len.unsworth@acu.edu.au

Jen Cope (PhD) is Research Assistant in the ILSTE at the Australian Catholic University in Sydney. Her doctoral thesis (2016) incorporated a pedagogical approach to develop critical literacy skills. Recent publications include book chapters on critical literacy in English for specific purposes (Garnet Education, 2015) and cross-cultural English expressions of blame (John Benjamins, 2018). Jen's research interests include critical and multimodal literacies, image-language relations in assessment tests, and cross-cultural English language variations.
Email: jen.cope@acu.edu.au

Liz Nicholls is Literacy Teaching Educator with the Catholic Education Diocese of Parramatta. She has been working in primary school education for more than 25 years. Liz is a current PhD student researching image language interaction in primary school science discourse. Her PhD supervisor is Professor Len Unsworth in the ILSTE at the Australian Catholic University in Sydney.
Email: lnicholls@parra.catholic.edu.au