

Appendix I

Consultation

Consultation with the following organisations and individuals contributed to the production of this report:

- ACT Department of Education and Training
- Budget and Co-ordination Branch, Schools Division, (DEETYA)
- Department of Education and Arts (TAS)
- Department of Premier and Cabinet (SA)
- Department of Training and Education Coordination (NSW)
- Evaluation Branch, WA Department of Education
- Inter-governmental Relations, Department of Premier and Cabinet (QLD)
- NT Department of Education
- Office of the Queensland School Curriculum Council
- Productivity Commission
- Professor John Keeves, Flinders University, SA
- Queensland Department of Education
- Senior Secondary Assessment Board of SA
- Social Development and Justice Unit (SA)
- Social Policy, Department of Premier and Cabinet (VIC)
- Strategic Planning and School Support, Directorate of School Education (VIC)
- Strategic Planning Unit, Department of Children's Services (SA)
- The Cabinet Office (NSW)

APPENDIX II

MINIMUM AGE ENTRY - STATE/TERRITORY COMPARISON

		4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Age																
SA	BST	Reception	Year 1	Year 2	Year 3	Year 4	Yr 5 (10.1)	Year 6	Year 7	Year 8	Year 9	Year 10	Year 11	Year 12		
WA	MSE	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Yr 7 (11.1)	Year 8	Year 9	Year 10	Year 11	Year 12			
Qld	NET	Year 1	Year 2	Year 3	Year 4	Year 5	Yr 6 (10.1)	Year 7	Year 8	Year 9	Year 10	Year 11	Year 12			
NT	MAP	Transition	Year 1	Year 2	Year 3	Yr 4 (9.1)	Year 5	Yr 6 (11.1)	Year 7	Year 8	Year 9	Year 10	Year 11	Year 12		
Tas.	DART	Preparatory	Year 1	Year 2	Year 3	Year 4	Yr 5 (10.1)	Year 6	Year 7	Year 8	Year 9	Year 10	Year 11	Year 12		
ACT		K	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10	Year 11	Year 12		
Vic.	LAP	Preparatory	Year 1	Year 2	Year 3	Year 4	Yr 5 (9.1)	Year 6	Year 7	Year 8	Year 9	Year 10	Year 11	Year 12		
NSW	BST	K	Year 1	Year 2	Year 3	Year 4	Yr 5 (10)	Year 6	Year 7	Year 8	Year 9	Year 10	Year 11	Year 12		

APPENDIX III

COMPARISON BETWEEN STATE/TERRITORY TESTS

Options

Options to achieve comparison of student literacy across the nation (qualified by the different Year levels that States/Territories insist be used) include those shown in the following tables:

Proposal	Advantages	Disadvantages	Recommendation
1. Single new test	1. Uniform test across Australia. 2. Comparison simplified. 3. Test quality assured by project Consortium. 4. Could design the test to readily relate to the National Profiles.	1. States/Territories reject notion of a national test. 2. Cost of developing, pre-testing and publishing a new test. 3. States/Territories differ in their (assumed) notion of literacy and how to assess it.	Most acceptable from a strictly testing point of view but politically impractical - reject.
2. Each State's/Territory's own test + new common test	1. Allows use of each State's/Territory's own test plus a high quality common instrument for equating and comparing purposes. 2. Politically more acceptable to allow use of own test. 3. Facilitates equating and comparison.	1. Cost of developing, pre-testing and publishing new test. 2. States/Territories reject notion of a national test and even this lesser use for equating purposes is probably unacceptable. 3. Retains use of State/Territory tests shown to contain test design and item-writing flaws. 4. Possible security problem with a common test.	Reject on grounds of political unacceptability.
3. State's/Territory's own test with some	1 - 3 As for Nos. 1 to 3 in Proposal 2.	1. Cost of developing additional common items and finding way to insert	Reject on grounds of security.

common items inserted	2. Cheaper than developing a separate test.	<p>them into the various and different tests.</p> <p>2. It is already too late to insert additional items in 1997 State/Territory tests and possibly for 1998.</p> <p>3. Substantial security problem: the items become public once the first test including them is run.</p>	
4. State/Territory tests “cleaned up” and matched to National Profiles.	<p>1. Allows use of each State’s/Territory’s own tests.</p> <p>2. National Profiles provide a basis for equating and comparing.</p>	<p>1. Politically very unacceptable for a national project team to tell States/Territories their tests are flawed and to edit them.</p> <p>2. Not all States/Territories accept the National Profiles.</p> <p>3. As attractive as it seems, it is not easy to relate tests not designed for the purpose to scales of performance descriptors that, themselves, are subject to questions of validity.</p> <p>4. Cost of editing (sometimes substantially) and re-publishing all 6 tests.</p>	Reject on grounds of political unacceptability and cost.
5. Each State’s/Territory’s own tests plus a common test for equating purposes made up from “cleaned up” extracts from tests from all States and Territories.	<p>1. Allows use of each State’s/Territory’s own test.</p> <p>2. An equating instrument made up from selections from all the State/Territory tests may be politically more acceptable.</p> <p>3. Cleaning up the selected extracts into a new common test provides a high quality instrument.</p> <p>4. Facilitates equating and</p>	<p>1. The various State/Territory tests differ as indicated elsewhere and selecting bits for a coherent common test representative of all the tests may prove more difficult and more costly than developing a new test.</p> <p>2. The notion of a national project team “cleaning up” bits of State/Territory tests may be politically unacceptable.</p>	Reject on grounds of test design practicality, probable cost and political unacceptability.

	comparison.	<p>3. The cost of pre-testing and publishing a new common test.</p> <p>4. The new common test would still be a national test and may still be politically unacceptable.</p>	
6. Each State's/Territory's own test plus another State's/Territory's test.	<p>1. Allows use of each State's/Territory's own test. Accepts their material as worthwhile in its own right and sufficiently worthy for another State/Territory to take it.</p> <p>2. Using another State's/Territory's test may be politically more acceptable.</p> <p>3. Cost is minimised since a new test is not required to be developed.</p>	<p>1. The assignment of the second tests needs to be carefully arranged if national equating and comparing is to be possible.</p> <p>2. Different purposes, assumptions, definitions of literacy and test approaches, together with the deficiencies identified in the tests themselves complicate and possibly invalidate comparisons.</p> <p>3. Some States/Territories remain sceptical about the others' tests.</p>	<p>Reject on grounds of differences in underlying philosophies and outcomes invalidating comparisons leading to a lack of a common unifying metric.</p> <p>Recommended as a third choice in the event of any State's/Territory's rejecting the use of the National Profiles.</p>
7. Each State's/Territory's own test plus another State's/Territory's test but with outcomes matched to the National Profiles as a common yardstick.	<p>1 - 3. See Nos. 1 to 3 for Proposal 6.</p> <p>4. The National Profiles provide a common yardstick and basis for test interpretation thus facilitating equating and comparison.</p> <p>5. This limited use of the National Profiles as an expedient for equating purposes may be more acceptable.</p>	<p>1 - 3. See Nos. 1 to 3 for Proposal 6.</p> <p>4. The differing purposes, assumptions, definitions of literacy, test approaches, and the deficiencies in the tests themselves may make any relationship to the National Profiles difficult to realise to different extents.</p> <p>5. Some States/Territories reject the National Profiles.</p>	<p>When all grounds are considered (especially testing principles, ease of equating and political acceptability), this proposal is probably the most practical and acceptable.</p> <p>Recommended as second choice.</p>
8. Each State's/Territory's own test plus another State's/Territory's test but with outcomes matched to the National	<p>1 - 5. See Nos. 1 - 5 for Proposal 7.</p> <p>6. Linkages of statistics are strengthened thus further strengthening the equating model.</p>	<p>1 - 5. See Nos. 1 to 5 for Proposal 7.</p> <p>6. Complexity of design and lack of complete negotiation with the States/Territories.</p> <p>7. Some States/Territories</p>	<p>Has same advantages as proposal 7 plus improved statistical model.</p> <p>Recommended as first choice subject</p>

Profiles as a common yardstick. All recommended combinations administered within all States/Territories using a counter-balanced design.	7. Counter-balanced design controls for practice effect of home test in the equating design.	reject the DART and will not allow all tests to be administered. 8. Cost and complexity of administration and especially central marking. 9. It seems unlikely that all States/Territories would agree to the implementation of this model.	to agreement of all States and Territories.
9. Each State's/Territory's own test plus the National Literacy Survey's (NLS) test and with outcomes matched to the National Profiles as a common yardstick.	1 - 5. See Nos. 1 to 5 for Proposal 7. 6. Uses an existing nationally validated test for equating purposes. 7. The National Literacy Survey test is linked to the National Profiles. 8. Use of the National Literacy Survey test may be more acceptable in some States/Territories than another State's/Territory's test.	1 - 6. See Nos. 1 to 6 for Proposal 8. 7. The National Literacy Survey test requires a 3-week instructional period before its administration and is thus very expensive in terms of teacher time and intrusive on school time (see discussion elsewhere).	Reject on grounds of cost and school time intrusion.

Recommendation

Proposal 8 is recommended provided that the States/Territories agree to the comprehensive counter-balanced design.

The Consortium thought it highly unlikely that such agreement would be forthcoming and consequently offers Proposal 7 as a theoretically defensible and practical alternative recommendation.

Comparison between the State/Territory Tests

The following tables are based on the materials received from the States and Territories when they were asked to provide copies of their test material and supporting documentation. It is accurate and comprehensive to the extent that the material supplied was accurate and comprehensive. In some instances, authorities indicated that the confidential nature of the material precluded their supply to the Consortium.

READING

Feature	Qld (AOL, also known as NET)	NSW-SA (BST)	Vic. (LAP)	WA (MSE) (NB. Only a sample test and supportin g documents for the WA tests were available.	NT (MAP)	Tas-ACT & DART DART (Tas) and DART (standard) (The Tas. 1993 test is related to but not identical with DART. For 1996 DART seemed to be used.)
(Presentat ion) Format	Magazine + Ques. Booklet	Magazine + Ques. Booklet	Traditional script	Stimulus booklet and question paper	“Big” (oversize) book + Ques. Booklet	Newspaper + Ques. Booklet
Time Allocated	(1996) 70 mins. (practice + test)	Informatio n not supplied	45 minutes	75 mins.	No limit - teacher decides	75 mins. (some flexibility)
Literacy defined	No (not in materials supplied)	No (not in materials supplied)	No (not in materials supplied)	No (not in materials supplied. Reference is to “English” which is sub- divided as Reading and Writing.	No (not in materials supplied other than to refer to reading comprehen -sion)	Reading defined as the ability to make meaning from a variety of written text types.
Full Test Specificati ons Available to Consortiu m¹	No	No	No	No	No	No
Local Year Level Focus	6	5	6 (?)	Yrs 3, 7, & 10	Yrs 5 & 7	Yrs 5 and 6

¹ Specifications refer to detailed instructions to itemwriters on how to prepare the tests. In no case was such information available even though most tests provide detailed instructions on administering the tests and, in some cases, on assessing student performance on the tests.

No. of Reading Items	(1995) 45 + 9 lines of proofrdg. (1996) 37	(1994) 46 + 3 pages of cloze and interlinear (1995) 45 + 3 pages of cloze and interlinear (1996) 46 + 3 pages of cloze and interlinear	(1996) 58	(Sample Yr 7) 20	(1995) 121 (Teacher determines where a student starts and finishes.)	(Tas.) 42 DART A 22 DART B 27
Itemtypes (Not all in each column appear in each of that State's tests)	MCQ (ques. form and sent. completion) MCQ Pictures/ Diagrams Sequencing	MCQ (ques. form and sent. completion) MCQ Pictures/ Diagrams Pict. Interp. Interlinear Banked Cloze	MCQ Banked Cloze Sequencing	(Available info. was not complete and not all listed below were viewed.) MCQ Cloze Open-ended questions Open-ended sentence completion	MCQ Banked Cloze Cloze Open-ended Questions Open-ended sentence completion Sequencing Sent.-Pict. match	MCQ Cloze Open-ended Questions

	Proofrdg.	Proofrdg.				Showing meaning by drawing Re-write in changed form
Item Focus	General Compreh. Inference Register/ Style Punctuation Grammar Spelling Word choice Purpose of Text/Part	General Compreh. Gist Inference Grammar Word Choice	General Compreh. Gist Inference Genre Grammar Spelling Purpose of Text/Part	General Compreh. Inference Reason/ Explanation	General Compreh. Inference Spelling Word meaning	General Compreh. Inference Grammar Word meaning Phonics Alphabetic Order
How scored	Centrally	Centrally	Centrally (?)	Locally	Locally by teachers	(DART) Locally by teacher using marking guide
How results reported	Information not supplied	Information not supplied	Percentage at CSF levels. LAP reporter disk supplied to schools.	Performance scales using National Profiles, detailed statistical results and matched against State profile.	Numerical score related to Stages (5 to 7) and to Territory-wide profiles	Numerical score related to Profile Levels and short descriptive statements related to tests
Relation	No	No	Related	Related	Said to be	Related

					"Everyday" texts	TV Program Newspaper Ad.
Item performance Statistics available to Consortium	No	Yes	No	No	Some information available	No
Analysis of Statewide Outcomes Viewed	Yes	Yes		Yes	Yes	Yes
Mode of Statement of Results	No information supplied	No information supplied	LAP reporter disk - Profile of performance descriptors	Profile using perform. descriptors	Scores related by IRT and classical statistics to stages (pre-stage 5 to Beyond Stage 7)	Numerical score related to Profile Levels and descriptive statement derived from test
Background knowledge assumed other than general Australian	Slight use of names and themes from other cultures but posed no obvious cultural problems.	Some use of names and themes from other cultures but posed no obvious cultural problems.	Nil	Some use of names and themes from other cultures but posed no obvious cultural problems.	Some use of names and themes from other cultures but posed no obvious cultural problems. One non-urban test also used Aboriginal and African themes in addition to same as urban.	nil
Performance/Proficiency Scale(s) available to Consortium	No	No	CSF derived from National Profiles	Yes, related to National Profiles	Yes, eventually National Profiles	Levels from National Profiles and descriptors based on tests

WRITING

Feature	Qld	NSW-SA (not available)	Vic. (Test viewed for 1995, other document- ation for 1996)	WA	NT	Tas-ACT & DART
(Presentation) Format	Stimulus statement and instruction	n/a ²	(1995) Context statement and instruction	Unclear from documen- tation viewed but all students in Yrs 7 and 10 were set the same task. (a narrative and a letter)	Folio of 3 writing scripts	Stimulus statements and pictures etc together with instruction to write.
Time Allocated	(1996) 55 mins. (include. 25 mins. discussion)	n/a	Up to 90 minutes writing in class for teacher marked tasks plus 20 minutes for centrally marked tasks	80 mins. (narrative)& /or 80 mins. (letter)	No limit set	2hr40m (including discussion)
Full Specifications Available to Consortium	No	n/a	No (1996) Yes	No	No	No
No. of Items	1	n/a	(1995)1	2 (?)	3	3 with optional editing and poster in DART
Itemtypes	Report (possibly including narrative and description)	n/a	(1995) Newsletter report or article (possibly including narrative and description)	Narrative Letter	3 different genres and 3 different subject areas (one genre and one subject area is common across the sample of schools)	Essay Story Narrative Editing
How scored	Centrally	n/a	(1996) Both centrally and	Centrally by a team of	By teachers using "frame-	Locally by teacher using

² n/a means here "not available to or sighted by the Consortium" though it may or may not exist.

			in schools	teachers	works” or criteria provided, moderated at school and system levels.	marking guide with performance descriptors
How results reported	No information available	n/a	(1996) Against descriptive statements in the <i>Curriculum and Standards Framework</i> giving levels of performance in 4 categories: strategies, texts, contextual understanding, and linguistic structures and features	Performance scales using National Profiles, detailed statistical results and matched against State profile.	Against criteria in four “aspects” of subject matter, or- ganisation, language use and - conventions.	Against descriptive statements in categories of “Content/ Context” and “Language”. These then related to Profile Levels.
Relation to National Profiles	No information available	n/a	(1996) Not used but the <i>Curriculum and Standards Framework</i> can be related to the Profiles	Related	Not used but presumably working towards National Profiles as for Reading.	Related
Analytic /Global	No information available	n/a	(1995) n/a (1996) Both but using performance descriptors within categories	Both, with analytic rating focussing on <i>presentation, punctuation, spelling, grammar, form of writing, vocab., content, elaboration of ideas, sequencing, syntax.</i>	Both, with analytic rating focussing on categories referred to above.	Both but mainly global
Integrated with Reading	No	n/a	No	No	No but some R items contain much Writing	No
Administr.	Yes	n/a	n/a	Yes	n/a	Yes

Instructions						
Genres	Narrative Newspaper report	n/a	(1995) Narrative Description Newsletter report or article	Narrative Letter	Folio of 3 pieces of writing in 3 genres one of which is common to all students/schools	Expressive Personal Narrative Expository Poster
Item performance Statistics	No	n/a	n/a	Overall, not individual items	n/a	No
Analysis of Statewide Outcomes Viewed	Not available to Consortium	n/a	n/a	Yes	n/a	Yes
Mode of Statement of Results	Not available to Consortium	n/a	Profile from CSF on LAP reporter disk	Profile using performance descriptors related to National Profiles	Report of rank ordering based on moderated teacher assessments.	Performance descriptors related to Profile levels.
Background knowledge assumed beyond general Australian	Nil	n/a	(1995) nil	nil	Nil	nil
Performance/Proficiency Scale(s)	Not available	n/a	(1995) n/a (1996) Yes, based on Victoria's <i>Curriculum and Standards Framework</i> , relatable to National Profiles	Yes, related to National Profiles and examples of student writing	Yes, set criteria as above and eventually National Profiles	Profile levels and performance descriptors

APPENDIX IV

RELATIONSHIP BETWEEN THE STATE/TERRITORY TESTS AND THE NATIONAL PROFILES FOR ENGLISH

Introduction

Only some of the tests and related documentation viewed made explicit reference to the National English Profiles. However, it should, in principle, be possible to relate test outcomes and possibly performance on individual items to levels in the National Profiles. In other words, since the National Profiles are essentially performance descriptors, it should be possible to relate overall results on any of the tests and possibly individual items to levels on the National Profiles so that they could provide a common yardstick by which to interpret all test results and by which, if required, comparisons could be made.

DART and the Tasmania-Australian Capital Territory Tests

Reference is made to the National English Profiles (*English- A Curriculum Profile for Australian Schools*) in the design of these tests and the reporting and interpretation of the results.

The documentation supporting both tests is similar and makes statements such as the following:

The DART materials were developed to match the National English Profile, in which the three strands are Reading and Viewing, Speaking and Listening, and Writing.

While DART refers to English, it addresses literacy as understood in Tasmania.

Each student's estimated level of achievement on a strand is described in terms of the kinds of skills and understandings typically displayed at that level - example: a student with a score of 21 on Reading is estimated to be achieving at Level 4 of the English Profile.

DART offers core assessments in viewing, listening, speaking and writing. [These are the strands in the English Profiles.]

DART also offers four supplementary measures:

Small Group Discussion...
Role Play...
Poster...
Editing...

The DART manual provides a detailed description of the test, its design and its relationship to the National English Profiles. It lacks, however, detailed test specifications or an indication of how the test items and scores are related to the Profiles, though one surmises that one part of the process is the matching of item descriptions against elements of the Profile descriptors. The manual also makes statements such as the following about the relationship with the Profiles:

DART provides an estimate of a student's level of achievement on each of five strands: Viewing, Reading, Listening, Speaking and Writing. Each estimate is based on the student's recorded performance on a set of DART activities. These activities address skills and understandings identified in the English Profile for Australian Schools...

DART provides a form that can be used to display a student's performances on all five strands of the English Profile simultaneously.

The manual provides a table relating scores on the DART to levels on the English Profiles.

Western Australian tests

No discussion could be found in the materials relating the tests to the National English Profiles though an acknowledgment at the foot of two pages suggested that the National Profiles were used for outcome statements in Reading and Writing.

Outcomes are expressed in terms of the Western Australian Student Outcome Statements. However, the manual indicates that

The performance profiles map on a continuum the skills used by students in responding to the assessment task. Skills and understanding are mapped to show those skills that are relatively less difficult through to those that are more difficult.

An illustrative table shows scores matched against short behavioural descriptions such as *completes a familiar idiomatic expression* or *draws comparison between general and specific classifications*. The manual does not indicate how the match between scores and descriptors is carried out. It should, in principle, be possible to relate these behavioural descriptions to the National English Profiles.

Queensland and New South Wales

There was no reference to the National English Profiles in the material that was made available to the Consortium. Nevertheless, the general statement above is applicable: it should be possible to relate overall outcomes and possibly performance on individual items to the Profiles. This was reinforced by a separate statement received from Education Queensland (the State Department of

Education) comparing reading outcomes from the Queensland syllabus and the National Profiles. Though not explicitly stated, the implication was that the State test (Aspects of Literacy) assessed the syllabus outcomes and so could, in turn, be related to the National Profiles. The statement from Education Queensland matched National Profile levels and performance descriptors with the (draft) Queensland Student Performance Standards while acknowledging some differences.

Northern Territory

There was no significant discussion of the National English Profiles or of the relevance of the tests or outcome statements to them in the material made available to the Consortium. However, one statement did indicate that it was intended to work towards the use of the Profiles. Again, it is reasonable to anticipate that the Stages referred to in the Northern Territory materials, the overall outcomes, and the individual items, could be related to the profiles.

Victoria

The test materials viewed for 1995 made no reference to the National English Profiles. However, discussion with one of the developers of the Victorian materials indicated that the individual reading items had been related quite precisely to the levels in the Profiles. No information was available on writing prior to 1996 but extensive documentation was viewed in relation to the assessment of writing in 1996. Assessment is based on the performance descriptors in the Victorian *Curriculum and Standards Framework* but these can be related to the National Profiles without much difficulty.

APPENDIX V

INDICATIVE COMPARATIVE ANALYSIS OF TWO TEST ITEMS: WHAT DO STUDENTS HAVE TO DO, KNOW AND VALUE IN ORDER TO DO A TEST?

In this section, an analysis is reported showing a comparison between two test items which, on the surface, appear to measure similar 'literacy' outcomes. There are two purposes of the analysis: The first is to provide explanatory documentation as to why 'item comparison' is difficult between test batteries through a demonstration of the different notions of 'literacy' operating in tests across the nation.

The second purpose of the analysis is to make explicit that, while the Project recommends procedures for establishing equivalences between State/Territory tests, it cautions that future work needs to be done to inform the ways in which differing assessed aspects of literacy may be used in national comparison.

This is not to say that national equivalences information is not useful and important. The research team believes that projects such as this may in fact help clarify the nature of such comparisons.

It should be noted clearly that the analysis is not intended to be critical of the items or tests used as examples, nor in any way to suggest that the tests are not valid or useful. The comparison is made only to illustrate the differing skills and outcomes required and/or implied by seemingly similar test items, aimed at measuring similar skills for a similar target group.

Surface feature comparisons

Two tests items, each from a different test battery, are used for the analysis:

1. The Northern Territory Multilevel Assessment Program 1995 for Urban Schools 'Reading Comprehension' test (Item 75), and
2. The Basic Skills Testing Program 'Aspects of Literacy' 1996 Year 5 used by New South Wales and South Australia (Item 36).

The test items both relate to passages of text, each divided into paragraphs. Both passages form part of test batteries specifically designed to measure reading comprehension in the target group of students aged approximately 10 years of age. Each item requires a short answer. In the case of the MAP test it is one word. In the case of the BST test it is checking the space in a multiple-choice item.

For this brief analysis, these were the only comparisons for which an attempt was made to equate items. It should be noted that in other items responses required varied enormously. In the case of the MAP test, many items test writing incidentally with 12 of the 121 items requiring a paragraph of writing involving explicit skills such as 'imagine', 'list', 'why?', 'give three reasons', 'describe', and 'write out a paragraph which indicates...'. These skills, in tandem with the knowledge required to answer some questions, are clearly of a significant and complex order, both in terms of quality and quantity. In contrast, the BST reading comprehension test is almost entirely constructed of multiple-choice items. No question at all requires a written response involving words. Responses are made by filling in the spaces required, presumably to promote electronic marking of scripts.

This analysis has chosen to ignore the test conditions of each test battery. These conditions are not viewed as being relevant to the purpose of the analysis, which is to help answer the question, 'What do students have to do, know and value in order to complete these items?'

Multilevel Assessment Program (MAP) 1995 (Urban) Reading Comprehension

Item 75 of this Reading test is based on a reading passage which is approximately 1500 words long. It is called 'The Inuit'. It is spread across five columns of text, three of which have around 12 words per line, and two of which have about 16 words per line. There is a map at the start, and three black and white photographs interspersed through the text. Fourteen questions are connected with the passage out of a total of 121 for the whole test.

Item 75 reads as follows:

*Give another word for 'cached,' in paragraph four of The Seasonal Cycle.
(The word must fit grammatically in the text.)*

A single line appears below the question, implying that there is a one-line space only for the answer.

First, it should be noted that there is an apparent error in the punctuation of the question, in that the quotation marks around the word 'cached' enclose the comma after that word. Most commonly accepted rules of punctuation would put the comma outside the quotation marks. It is also against accepted grammatical rules to have a comma in that location at all. Mention is made of this to alert readers to the many difficulties associated with test item construction. We in no way imply that such a typographical error is likely to influence test results, though of course it is a possibility. It does however raise the issue of 'whose grammar' or which grammatical system should be used as a benchmark when question 75 specifically requires the student to know and make reference to grammatical fit.

Here is a list of skills, knowledge and values which it is likely that a student would have to know in order to answer the question either at all, or correctly:

Knowledge. (Know that, what...)	Skills. Be able to...	Values & attitudes
<ol style="list-style-type: none"> 1. that 'another' word means a word with the same meaning, not just any other word. This also means knowing about school activities such as synonym exercises. 2. what a paragraph is 3. that the words 'The Seasonal Cycle' refer to a sub-heading in the 1500 word passage 4. understand that even though the words 'The Seasonal Cycle' are not distinguished by quotation marks, italics or other means except capitalisation of first letters, that nevertheless these words form a group intended to refer to a sub-heading title 5. what a particular grammar system is in order to ensure the chosen answer 'fits' in the text 6. that the text referred to for the answer to have a grammatical fit is the text of the original passage, not that of the answer book 7. that the correct place to insert the answer is on the black line 	<ol style="list-style-type: none"> 1. read and understand a 1500 word prose passage 2. read the question 3. understand the question (see also 'knowledge' items in left column which have to be 'known' before they can be enacted as skills) 4. count groups of words ('four' paragraphs) 5. write (form the answer word with a pen or pencil) 6. identify synonyms 	<ol style="list-style-type: none"> 1. perseverance to read to the end of 1500 words under test conditions 2. value tests 3. value reading prose about Inuit

This list is not intended to be comprehensive, assuming as it does that students know and do basic material such as form and recognise upper and lower case letters of the alphabet and so on. It is only meant to be indicative.

Basic Skills Testing (BST) Program

The passage is called 'Tropical Paradise' and is about a boy returning home, purposely avoiding his parents, who are working in a fish-and-chip shop. Question 36 of the 1996 'Aspects of Literacy' test is a multiple-choice item. It is reproduced below:

36. What did Keith think was wrong with Mum and Dad?

- ☐ They were hot.
- ☐ They were late.
- ☐ They had seen Keith.
- ☐ They didn't like swimming.

To answer this question correctly, students must be able to interpret, differentiate and discriminate between items of information in this particular section of the story:

"They [Keith's Mum and Dad] were looking at each other. Dad was saying something to Mum, pointing at her with a piece of fish, and Mum was saying something back, waving the chip scoop at him.

Even at that distance, Keith could see that Dad's mouth was droopier than a palm frond and that Mum's forehead had more furrows in it than wet sand when the sea was a bit choppy.

Keith's stomach knotted even tighter.

Another argument.

Poor things. Stuck in a fish-and-chip shop all day in the heat. Anyone'd get a bit irritable standing over a fryer all day with this poxy sun pounding down nonstop."

The question asks 'What did Keith think was wrong with Mum and Dad?' There are two good answers here. One would respond to the parents' argument, and formulating this answer would suggest students showed some evidence of feelings, empathy, and perhaps appreciated that life had its complex moments. The other response, the one students 'have' to choose as 'right' relates to the reason for his parents' argument. It is doubtful if the question phrasing 'what is wrong with' in fact means the same as asking by implication what the actual cause of the problem might be. It seems that being hot may not be the best answer to a question about 'wrongness'. Be that as it may, the question does force a student to choose a simple rather than challenging answer to the test item.

The table below presents a list similar to the one above:

Knowledge. (Know that, what...)	Skills. Be able to...	Values & attitudes
1. that 'wrong' refers to 'being hot' rather than to Mum and Dad fighting 2. what a paragraph and other written conventions are 3. how standard conversational text is	1. read and understand a 200 word prose passage 2. read the question 3. understand the question (see also 'knowledge' items in left column which have to be 'known' before they can be enacted as skills)	1. perseverance to read the 200 word passage under test conditions 2. value tests 3. value reading prose about this topic (table continues)

formatted in print 4. that third person text can legitimately carry heavily idiomatic first-person expressions such as 'anyone'd' and 'poxy sun pounding down nonstop'.	4. fill in a small circle to indicate correct response with a pen or pencil	
--	---	--

Conclusion

It can be seen at a glance that the skills, knowledge and values required of the BST item are of a qualitatively and quantitatively different order from the MAP test. The implications of this are not related to which is 'right' and which is 'wrong', but to the fact of the differences.

Statistically, it is important to emphasise that comparisons of literacy tests across the nation can be made, will be made, and perhaps even should be made. The point to consider, however, is that what is being compared does not necessarily equate, nor does it seem to measure like notions of literacy.

Meanings for the term 'literacy' vary, and there is ample reliable research evidence to show that what teachers teach when they teach literacy varies from one location to another, and what is tested will vary accordingly. National English curriculum profiles are one way of showing how the reporting of literacy outcomes can in some way be compared.

LIST OF ATTACHMENTS

1. ACT test materials
2. NSW test materials
3. NT test materials
4. Qld test materials
5. SA test materials
6. Tas. test materials
7. Vic. test materials
8. WA test materials