

6. Sample Materials

Research Question 5

What examples are available of teacher-assessed tasks and teacher judgment processes in other statewide or international testing programs? What samples of materials can be gathered?

6.1 Teacher-assessed Tasks in Large-scale Testing Programs

There is a variety of large-scale census testing programs that involve teachers assessing tasks both nationally and internationally. These range from teachers being trained centrally to mark test tasks to teachers marking the tasks their own students undertake (see table 1 for examples). Only representative samples of the types of teacher-assessed tasks from the United States of America are discussed here, while all census literacy and numeracy tests from Australian States and Territories are presented. Each is considered below, with an example, and an appraisal based upon the critical positioning from the literature review follows.

6.1.1 Indiana Statewide Testing for Educational Progress (ISTEP+)

The Indiana State Board of Education developed ISTEP+ (Indiana Education Department 2001). One component is a criterion-referenced applied skills assessment for Grades 3, 6, 8 and 10 in English and Mathematics. The applied skills test is marked by trained qualified 'readers', recruited publicly, but are teachers or those who hold a teaching degree. Readers must successfully complete a formal training program and are monitored heavily — unacceptable readers are either retrained or replaced. Figure 1 is an example of a task in Mathematics and the scoring rubric is seen in figure 2. Figure 3 shows a scoring exemplar. While not an extensive open-ended task example, it is representative of the sorts of teacher-assessed tasks that occur in ISTEP+.

II  Use your punchout coins to solve this problem.

Laura found 3 coins in her backpack. Two of the coins are the same and one is different. What is the LEAST amount of money the 3 coins can add up to?

You MUST show your work.

Answer _____ c

Figure 1
Teacher-Assessed Task Mathematics Grade 3 (Indiana Department of Education 2002a, p. 22)

Rubric:	
2 points	Exemplary response
1 point	Correct answer only OR Correct complete process; error in computation
0 points	Other

Figure 2
Scoring Rubric (Indiana Department of Education 2002b, p. 67)

Session 2—Item 11
Score Point 2

This response matches the exemplary response contained in the rubric. The student gives the correct answer of 7 cents and shows the correct work demonstrating the process used. The response receives a Score Point 2.

SCORE POINT 2

II Use your punchout coins to solve this problem.

Laura found 3 coins in her backpack. Two of the coins are the same and one is different. What is the LEAST amount of money the 3 coins can add up to?

You **MUST** show your work.

Answer 7 cents c

Figure 3
Scoring exemplar (Indiana Department of Education 2002b, p. 67)

ISEP+ applied skills test is an example of teachers assessing open-ended tasks. Criteria for marking are rigidly set, and markers must pass scoring examinations before they can be accredited. Teacher involvement is limited to marking only, and most of the teachers involved are not working in classrooms. There would be limited professional development for classroom teachers, and few links to effective pedagogy. There is little use of consistency processes, except formal training of markers.

6.1.2 Kentucky Department of Education Commonwealth Accountability Testing System (CATS)

The CATS (Kentucky Department of Education 2001) was developed through a collaborative process between teachers, parents and education advisors. Students are assessed in national, core content, writing portfolio, and demand writing tests. The writing portfolio is the only part of CATS that is teacher-assessed. The rest of the test, both open response and multiple-choice items, are scored by a test contractor for the state. Teachers, administrators, specialists and members of education organisations undertake assessment design and writing of item and scoring guides. The writing portfolio is assessed at the 4th, 7th and 12th grade. Teachers are trained to use an established scoring rubric (marking criteria) and random student portfolios are reviewed by the state for accuracy in scoring. Classroom teachers are encouraged to collaboratively mark the portfolios. Portfolio results are reported separately from the centrally marked demand writing tests.

Portfolios are a collection of a student's best writing. Students, with their teachers, choose pieces produced in their classes over a period of time. Grade 4 students make four selections of genre: reflective, personal, literary, and transactive. Grade 7 and 12 select the same as Grade 4, but make another selection on a genre of their own choice. Figure 4 shows part of the Parent Guidebook that explains the writing portfolio components and the marking criteria.

WHAT ARE THE REQUIRED PIECES IN THE 7TH GRADE WRITING PORTFOLIO?

The student includes a total of **5** pieces of writing in the portfolio. Any of the following portfolio entries may come from study areas other than English language arts, but a minimum of one piece of writing must come from another subject area.

- **Reflective Writing** in the form of
 - **Letter to the Reviewer** — discussing the student's growth as a writer and reflecting on pieces in the portfolio. *(Student must include one.)*
- **Personal Expressive Writing(s)** in the form of
 - **Personal Narrative** — focusing on one event in the life of the writer
 - **Memoir** — focusing on the relationship of the writer with a particular person, place, animal, or thing
 - **Personal Essay** — focusing on a central idea supported by a variety of incidents in the writer's life
(Student must include one or two.)
- **Literary Writing(s)** in the form of
 - **Short story** • **Poem** • **Script**
(Student must include one or two.)
- **Transactive Writing(s)** for a variety of authentic audiences and purposes in real-world forms (e.g., letter, article, editorial, proposal, brochure, review).
(Student must include one or two.)

In addition to the 5 pieces of writing, each portfolio must include the following:

- **Table of Contents**
- **Student Signature Sheet** — states ownership of the portfolio and may give permission to use the portfolio for training (optional)

The Scoring Guide

Unlike a grade of A or B, your child's score on the Writing Portfolio can give you information about the characteristics most often observed in your child's writing. When you and your child know what to look for, you also know what needs improvement. The *Kentucky Holistic Scoring Guide*, below, lists the qualities of effective writing under "Proficient," the goal for all Kentucky students.

NOVICE

- Limited awareness of audience and/or purpose
- Minimal idea development; limited and/or unrelated details
- Random and/or weak organization
- Incorrect and/or ineffective sentence structure
- Incorrect and/or ineffective language
- Errors in spelling, punctuation, and capitalization disproportionate to length and complexity of writing

APPRENTICE

- Some evidence of communicating with an audience for a specific purpose; some lapses in focus
- Unelaborated idea development; unelaborated and/or repetitious details
- Lapses in organization and/or coherence
- Simplistic and/or awkward sentence structure
- Simplistic and/or imprecise language
- Some errors in spelling, punctuation, and capitalization that do not interfere with communication

PROFICIENT

- Focused on a purpose; communicates with audience; evidence of voice and/or suitable tone
- Depth of idea development supported by elaborated, relevant details
- Logical, coherent organization
- Controlled and varied sentence structure
- Acceptable, effective language
- Few errors in spelling, punctuation, and capitalization relative to the length and complexity

DISTINGUISHED

- Establishes a purpose and maintains clear focus; strong awareness of audience; evidence of distinctive voice and/or appropriate tone
- Depth and complexity of ideas supported by rich, engaging, and/or pertinent details; evidence of analysis, reflection, insight
- Careful and/or subtle organization
- Variety in sentence structure and length enhances effect
- Precise and/or rich language
- Control of spelling, punctuation, and capitalization

Figure 4

Parent Guidebook on writing portfolio components and marking criteria (Kentucky Department of Education 2002, pp. 6 & 9)

Portfolio assessment is considered to be an effective technique in writing assessment (Freedman 1993). Its use in a large-scale test would contribute to teacher development and to effective writing pedagogy. Teachers are involved in most aspects of test design and in the marking of the portfolios with scoring rubrics. These processes would contribute to teacher development. As teachers are encouraged to share their marking with others, only some consistency processes are used, but it is not known how this sharing contributes to moderating the assessments. As schools are rewarded for increased student performance the stakes for CATS are reasonably high, and if

classroom teachers are involved in selecting samples of writing, and in marking the samples, it could be difficult for them not to offer intervention, or 'mark up' the work. As the portfolio results are reported separately any differences between these scores and writing demand, as they comprise two different writing situations, may be explained, but this issue could be problematic from an accountability position.

6.1.3 Maryland School Performance Assessment Program (MSPAP)

MSPAP (Maryland Department of Education 2001) consists of criterion-referenced performance tests in reading, mathematics, writing, language usage, science and social studies for students in grades 3, 5 and 8. Tests are based on learning outcomes developed by Maryland educators. These specify what students should know and be able to do. The tests emphasise higher order skills to solve problems, make decisions and understand information. It utilises short and extended response items and individual and group performance tasks.

Teachers write the tasks (approximately 140) and mark the tests using state-developed rubrics. Approximately 650 teachers are involved in marking 185 000 tests. All answer books, for a given grade and cluster, are marked at the same time, and at different sites around the state. Scorers mark the open-ended responses and assign a score point on a scan sheet. Quality of scorers' marks is maintained by check sets, accuracy sets, spot checks, and retraining. Figure 5 is an example of part of a task for Grade 3 on 'Deserts'. This part involves writing to persuade and is at the end of a five-day test sequence. All test tasks are related to the theme, and test performance in reading, language use and Social Studies as well as writing. Figure 6 shows the marking criteria and two examples of scored work.

Wednesday, Task 1
Title: Deserts

WRITING PROMPT: WRITING TO PERSUADE

You have heard of someone who is thinking about traveling across the Sahara Desert, the way Geoffrey Moorhouse did. However, this person is not sure whether to take the trip. Write a letter to the traveler to persuade him or her either to go on the trip or to stay home. You may use information from your reading to help support your point of view.

PRE-WRITING

- Think about the problems Geoffrey Moorhouse had on his trip.
- Think about things another person could do differently today to avoid those problems.
- Think about whether or not the traveler should take this trip.

As you write, you may try making a list, a web, or a diagram on the lined paper provided to arrange the ideas you want to share in your letter.

DRAFTING

Use your ideas as you write a first draft of your letter on the lined paper. You will have 30 minutes to plan and to write your first draft.

Thursday, Task 1

Title: Deserts

REVISING

Yesterday you wrote a first draft of a letter. Today you will take 5 minutes to read your draft and think about what you have written. Imagine that you are the traveler reading the letter. Think about the answers to the questions below.

1. Does the letter persuade the traveler to do what is best?
2. Does the letter give reasons that support that advice?
3. Does the letter make sense?

After you have thought about how well your letter answers these questions, you will get some ideas from a partner to help improve your writing.

Thursday, Task 1

Title: Deserts

PEER RESPONSE

You have had the chance to ask yourself questions about how well you have composed your writing. In order to determine if your writing says what you want it to say, it is usually helpful to get someone else to react to your writing. This is called "peer response." You will work with your partner to do your peer response. Your Peer Response Form is on page 36 of your Answer Book.

1. Decide with your partner who will go first.
2. Follow the instructions on the Peer Response Form, and be sure to allow enough time for both of you to read and take notes about the answers to the questions.

Figure 5

Example of part of a task for Grade 3 on 'Deserts' (Maryland State Department of Education 1996a, pp. 30, 34 & 35)

WRITING PROMPT: WRITING TO PERSUADE

You have heard of someone who is thinking about traveling across the Sahara Desert, the way Geoffrey Moorhouse did. However, this person is not sure whether to take the trip. Write a letter to the traveler to persuade him or her either to go on the trip or to stay home. You may use information from your reading to help support your point of view.

0 points

- 0 **Development:** The writer identifies an ambiguous position with little or no relevant personal and/or factual information to support that position; or, the writer fails to identify a position.
- 0 **Organization:** The writer presents an argument that is illogical and/or minimally maintained.
- 0 **Attention to Audience:** The writer does not address the needs and characteristics of the identified audience.
- 0 **Language:** The writer seldom, if ever, uses language choices to enhance the text.

Dear, Ola,
If you go on the trip, you can get lost. You might miss your family or pets. You might get sick.

Score = 0

1 Dear explorers

I do not want to go because I do not want to get in trouble so I'm going to stay home. I read a little paragraph about you maybe meet you but explorers were in it. It must be a good job to be an explorer and hard job to and finish. Long does it take to be an explorer? Was it fun? Is there? I wish I could help you but I'm not going to go but I would like to but I'm not.

Confusing, irrelevant information that is not persuasive.

Score = 0

Figure 6
Marking criteria and two examples of scored work (Maryland State Department of Education 1996b, pp. 19 & 20)

Maryland encourages 'teaching to the test' as the test items are about higher order thinking skills (Maryland Department of Education 2001 — What is MSPAP?, p. 1). Teacher involvement is high, with strong links to pedagogy — especially when teachers are encouraged to teach students how to respond to broader response tasks. MSPAP contributes to professional development because of the large numbers of teachers involved in task development and marking. Consistency issues are dealt with through training of markers, and checking strategies. Maryland has received some publicity recently over its Testing Program (Center for Education Reform Newswire 2002). Some sites' test marks indicated that schools in the area were not able to show required improvements. Parent groups in Maryland have criticised the tests and asked for their suspension while a tool could be found that will help the state comply with the Federal 'No Child Left Behind' Act (House Education and the Workforce Committee 2001). This is an indication of the American large-scale test discourses, and the high stakes that accompany the test processes.

6.1.4 Vermont Statewide Assessment System

Vermont assessments (Vermont Department of Education 2001) include the Vermont Developmental Reading Assessment (DRTA), Written Language Portfolio (WLPA) and Mathematics Problem Solving and Communication Portfolio (MPSCPA) assessments that are teacher-assessed. All results from these tests are reported separately from the centrally computer-marked tests.

The DRTA is a standards-based assessment in reading. It is administered to all Grade 2 students. Teachers mark oral reading for accuracy and retelling for comprehension, and results are analysed centrally. It is reported back to schools the percentage of students who score in the highest two levels — *achieved the standard* and *achieved the standard with honours*. DRTA was adapted from the original Developmental Reading Assessment published by Celebration Press.

The WLPA is a standards-based assessment administered to students in Grades 5 and 8. Students prepare six pieces of writing that have gone through the entire process of draft to final edit. Teachers assess the portfolios using a rubric which scores writing on a two-point scale. The six pieces are about a response to literature, a report/expository piece, a narrative, a procedural piece, a persuasive argument and a personal essay.

The MPSCPA, aligned to state framework of standards, is administered to all students in Grades 4, 8 and 10. The portfolio is a compilation of the students' best problem-solving work on assigned, complex tasks. In addition to maths work, students are asked to describe how they approached the problem. Teachers mark the portfolio by rating seven areas for Grades 4 and 8, and five areas for Grade 10, on a six-point scale. Areas assessed are approach and reasoning, connections, accuracy of the solution, mathematical language, representation and documentation in Grades 4 and 8, and approach and reasoning, execution, observation and extensions, mathematical communication and presentation in Grade 10. Sample portfolios are selected for re-scoring for state data, but schools use the locally scored portfolios for their school reports.

Figure 7 shows a marked task from the Mathematics portfolio, while figure 8 gives some of the criteria used for marking. The marked task shows that the student has identified an underlying mathematical concept — Level 2 'connections' criteria — second bullet point.

Problem Solving Criteria			
Approach and Reasoning			
START HERE →			
Level 1	Level 2	Level 3	Level 4
<ul style="list-style-type: none"> Approach wouldn't work or No approach evident 	<ul style="list-style-type: none"> Approach would lead to solving only part of the problem¹ or reaching a partial solution or Approach would work but there is some flaw in the reasoning 	<ul style="list-style-type: none"> Approach worked or would work for solving the problem, and reasoning, if evident, is not flawed <p>(Note: Use of a formula is an approach that worked or would work.)</p>	<ul style="list-style-type: none"> Approach worked, and at least one of the following 3 additional aspects of good problem solving is evident: <ul style="list-style-type: none"> Justifying the application of a known formula or rule used to solve all or part of the problem or Making a formula or rule used to solve all or part of the problem or Describing verification of her/his solution²
Connections			
START HERE →			
Level 1	Level 2	Level 3	Level 4
<ul style="list-style-type: none"> Response stopped without including a mathematically relevant observation with respect to her/his solution 	<ul style="list-style-type: none"> Made a mathematically relevant observation about her/his solution or Identified an underlying mathematical concept or pattern in her/his solution or Solved the problem and (s)he recreated³ the problem and found a new solution or Solved the problem and (s)he used a different mathematical process to solve the same problem 	<ul style="list-style-type: none"> Related this problem to a similar problem⁴ in a real world phenomenon by expressing the mathematical relationship(s) or Analyzed the relationship among elements in her/his solution or among similar or different mathematical topics in her/his solution or Tested and accepted and/or rejected an hypothesis or conjecture about her/his solution or Identified a formula or rule, while solving the problem, that worked or would work in solving all or part of that problem 	<ul style="list-style-type: none"> Solved the problem, discovered a general rule⁵ about the solution², and demonstrated understanding of the generalization either through expansion of the derivation, or through application to more than one other case or Solved the problem, and (s)he extended her/his solution to a more complicated situation or Evaluated the reasonableness or significance of her/his solution
Solution			
START HERE →			
Level 1	Level 2	Level 3	
<ul style="list-style-type: none"> No work is present or No part of the solution² is correct or Some work is present, but the work doesn't support the answer given 	<ul style="list-style-type: none"> The solution² is correct for only part of the problem, and there is work to support those correct parts) or The solution² contains mathematical errors which lead to an incomplete or incorrect answer 	<ul style="list-style-type: none"> The answer is correct, and the work in the solution² supports the answer 	
<p>¹ Would: An approach that would work for solving the problem addresses all aspects of the mathematical situation presented in the task. An approach that would work may contain mathematical errors, an incorrect solution, or may be incomplete.</p> <p>² Part of the Problem: Within a problem, there may be several mathematical components that need to be addressed, or there may be multi-parts, if not all of the mathematical components of the problem are addressed, or not all of the parts of the problem are addressed, then the student only found an approach to solve part of the problem.</p> <p>³ Solution: All of the work that was done to solve the problem, including the answer.</p> <p>⁴ Recreated: The student substituted different numbers in the same problem and found another solution, or used the same procedure in a different circumstance.</p> <p>⁵ General Rule: A rule that can be used no matter what the numbers in the problem are, either expressed in algebraic notation or in words.</p>			

Figure 8
Some marking criteria for Mathematics Problem-Solving and Communication Portfolio Tasks

Fairtest (2001) states that Vermont has nearly a model testing system. The assessment burden is reasonable as are the stakes. Teacher involvement is high due to the local marking of the three tests. Teachers are trained to use the portfolio rubrics. Marking meetings are held across the state twice yearly. Fairtest states that the portfolios were also intended to improve teaching. Independent reviews by the RAND Corporation confirmed that this intent was met; pedagogy links are high. Other grade areas include portfolio assessments, but as the program is comprehensive, selection of writing samples and problem-solving activities may dominate the curriculum in the years the test is administered. To assist with consistency, teachers are encouraged to collaboratively mark the portfolios, but Fairtest has stated that reliability of the assessments will need

continued attention. To help in this, the Vermont Department of Education has reduced the marking criteria from five points to three.

6.1.5 Australian Capital Territory Assessment Program 2001

The speaking component (Australian Capital Territory Department of Education and Community Services 2001) is the only component of the ACT tests in Literacy and Numeracy that is teacher-assessed, and only for Grades 3 and 5. Teachers assess on a four-point scale for content and performance — Grade 3, 1–4; Grade 5, 2–5. ACT provides a video for professional development when marking students' talk. Teachers introduce the task, and students talk about the task with a partner — instructions also include reminders about formal speaking. Figure 9 shows these instructions. Each student in the class then speaks for one minute while the teacher scores from a marking guide. The second task involves rehearsal time before the formal talk. Grade 5 are asked to prepare the talk before rehearsal with a partner. Palm cards and prompts are encouraged. Teachers are encouraged to mark in pairs, but this is not mandatory.

YEAR 3 - SPEAKING

ACT Assessment Program 2001 Year 3 SPEAKING



TASK 1: THE BEST TIME I EVER HAD

Speaking on the topic *The Best Time I ever had* involves three steps:

- Introduction
- Interview with a partner
- Individual presentations.

REQUIREMENTS

Time approximately 1 to 1.5 hours

Materials copies of *The Best Time I ever had* stimulus sheet (one for each student)

Students will need

- adequate space to work with a partner
- one *The Best Time I ever had* stimulus sheet for each student
- pencil
- palm-sized paper or card

For marking teachers will need

- 2B pencil
- Marking Guide
- Student Record sheets for 2001 ACT Assessment Program Year 3

Figure 9
An example of speaking task administration requirements

**ACT ASSESSMENT PROGRAM
2001 SPEAKING
STUDENT RECORD SHEET
YEAR 3**

INSTRUCTIONS:

- Use pencil only, preferably 2B
- Do NOT use any other ink color
- Please mark as fully
- Mark to show marks

← **CONFIDENTIAL!**

Please MARK LIKE THIS:

STUDENT ID: STUDENT NAME:

SCHOOL NAME:

TASK 1			TASK 2		
NON-PARTICIPATION	CONTENT	PERFORMANCE	NON-PARTICIPATION	CONTENT	PERFORMANCE
<input type="checkbox"/>					
<input type="checkbox"/>					
<input type="checkbox"/>					

Figure 10
Student record sheet and criteria

For this component the classroom teacher is the task administrator and the task marker. Figure 10 shows the record sheet and marking criteria. There are some pedagogy links as the test is close to real-life public speaking situations and strategies. There is a small amount of Professional Development through a video to support marking requirements, criteria to mark speaking on content and performance for a range of spoken texts, and room for moderation. Consistency is achieved through clear instructions, criteria understanding, practice marking with video support, and marking in pairs with moderation processes only encouraged. Speaking assessment does not contribute to National Benchmark data.

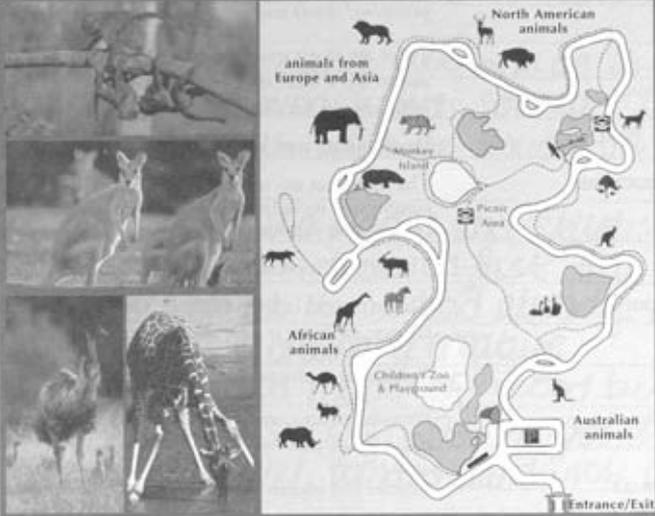
6.1.6 New South Wales

The Primary Writing Assessment (PWA) is administered to Years 3 (New South Wales Department of Education and Training 2001a) and 5 (2001b) students, at around the same time as the centrally marked multiple-choice Basic Skills Test. Students complete two writing tasks — one literary, one factual. Teachers mark these using set criteria for each task. The English Language and Learning Assessment (ELLA) is administered to Years 7 and 8 (New South Wales Department of Education and Training 2001, School Assessment and Reporting Unit 2001). The writing task for this program is teacher-assessed. The extended response task for the Statewide Numeracy Assessment Program (SNAP) (New South Wales Department of Education and Training 2001c) for Years 7 and 8 is teacher-assessed. Teachers who nominate for ‘in school’ markers receive two days of classroom release — one for training (after administration of the test) and one for marking the tasks. Figure 11 shows an example of the teacher-assessed writing task with figure 12 showing marking criteria.

D. MARKING WRITING TASK TWO: ZOO EXCURSION

Writing Task Two

Below is a map of the Western Plains Zoo. Visitors to the zoo can ride bikes, drive or walk around the zoo and look at the animals. The animals are not kept in cages. They can move around freely at a safe distance from the people.



Imagine your class went on an excursion to this zoo.
Write a recount of the excursion for the school newsletter.
Use the photographs and the map to give you some ideas about the things you might have seen.

- You should begin your recount of the excursion with an introduction.
- You should include details about the things you saw.
- You should write in sentences and use paragraphs to help organise your writing.
- You should pay attention to spelling and punctuation.

4

Figure 11
Example of the PWA teacher-assessed writing task

1. TASK LINKS TO THE CURRICULUM

Text purpose – to recount a school excursion

This task requires students to produce a piece of writing that recounts an excursion for the school newsletter. Students are asked to provide an introduction for readers, and then describe what they saw on the excursion. The purpose is for students to write about their observations on a school-based excursion to members of the school community.

Documenting educational experiences and events is part of all KLAs. The focus of the assessment of students responses is on their ability to present a logical organised text with elaborations about what they observed.

2. TASK-SPECIFIC MARKING CRITERIA

The specific marking criteria, and marks to be allocated for each criterion in Writing Task Two, are represented by the columns marked 1 to 15 on page 3 of the test books. These are identified below and expanded with examples in the table which follows.

Non-Attempt	TP										WL						
	1	2	3a	3b	3c	4	5	6	7	8	9	10	11	12	13	14	15
<input type="radio"/>	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰
	①	①	①	①	①	①	①	①	①	①	①	①	①	①	①	①	①
M																	
S																	
<input type="radio"/>																	
<input type="radio"/>																	

Writing Task Two — Zoo Excursion

Text	1	=	Text Function
Processes	2	=	Theme
	3a	=	Text Structure - Orientation
	3b	=	Text Structure - Elaboration
	3c	=	Text Structure - Conclusion
Text	4	=	Paragraphs
Features	5	=	Pronouns
	6	=	Conjunctions
	7	=	Sentence Structure
	8	=	Tense
Sentence	9	=	Clause Pattern
Level	10	=	Agreement
	11	=	Prepositions
	12	=	Articles/Plurals
	13	=	Punctuation
Word Level	14	=	Verb Form
	15	=	Spelling

Figure 12
PWA writing task marking criteria

Teachers are formally trained to mark the tasks, which are assessed against criteria. There are pedagogy links as the tasks are both closed and open-ended, and linked to syllabus documents. It is suggested that from the test results, teachers can adjust teaching programs to meet the needs of individual students. Professional development occurs through training of markers, and through their subsequent conversations upon returning to schools. Consistency is only achieved through formal training and markers being able to contact a coordinator by phone during the marking process. Only PWA and ELLA are used to contribute to national benchmark data.

6.1.7 Northern Territory Multi-level Assessment Program

The reading, spelling and numeracy tests are teacher-assessed. The Common Writing Task (CWT) is assessed centrally (Northern Territory Government 2001). Trained teachers mark the common writing task, and the criteria for marking are published in the administration guide. Teachers can mark their own students' work; complete the optional marking table on the cover of the test booklet, before returning scripts for central marking. The marking procedure is to assign a numerical score (based upon learning outcomes) against criteria of subject matter; ideas and vocabulary, and textual features; generic structure, cohesion, punctuation and spelling. Figure 13 shows the optional marking table on the tear-off front page of the student test booklet.

Reading item responses are reasonably precise. Students work through multi-levelled reading stimulus, recording responses using a variety of response techniques. Reading material is levelled against NT assessment profiles, and scores indicate which level students are working within. Descriptors explain reasons for each response in terms of learning outcomes. Numeracy responses are varied, but reasonably precise. Teachers mark with a key for right or wrong responses. Items are explained by referring to syllabus document page numbers. Scores indicate which level within which students are working.

The Northern Territory Board of Studies
Year 3 030211
Literacy Tasks, MAP 2001

Student Background Information
(To be completed for every student including those who did not do the test)

Teacher's use only

School Code
Student No
ESL
Absent
Exempt
Special Needs

School Name _____
Student Name _____
Outstation Name (if applicable) _____
Date of Birth _____
Day Month Year

1. Are you a boy or girl? Boy Girl
2. Are you an Aboriginal or Torres Strait Islander? Yes No
3. Does everyone at home speak to you in English? Yes No
4. How often do you speak English at home? Never Sometimes Usually Always
5. What is the main language spoken at home? _____
6. In which country were you born? _____
7. In which country was your father born? _____
8. In which country was your mother born? _____
9. Do you need English language support in an ESL program? Yes No

Teacher's use only

Please tick if this is a special needs student	<input type="checkbox"/>		
	Writing	Reading	Spelling
Please tick if this student was absent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Please tick if this student is exempt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Please attach explanation for exemptions			

OPTIONAL Common Writing Task: Teacher's Assessment

Subject Matter	Textual Features	Spelling
Quality of Ideas <input type="checkbox"/>	Cohesion <input type="checkbox"/>	Spelling in Context <input type="checkbox"/>
Vocabulary <input type="checkbox"/>	Punctuation <input type="checkbox"/>	
	Generic Structure <input type="checkbox"/>	

Based on school program and assessment measures indicate where on the continuum this student is currently working in Reading and Writing.

English (NT Profile) Reading Level
Writing Level
ESL (NT Profile) Reading Level
Writing Level
First Steps Reading
Writing
English NT Curriculum Framework Band Reading
Writing
Other Reading
Writing

Northern Territory Government

Figure 13
Front page of the NTAP student Year 3 literacy test booklet showing marking table for the optional common writing task teacher's assessment

Teachers would have to engage with marking criteria if taking the option to mark the CWT. They would have to indicate where the student is currently working within school programs based upon NT profile levels, First Step phases, or NT curriculum bands. There are obvious links here to NT syllabus, which indicates the program provides some positive professional development. Teachers are encouraged to let students know the results as soon as they have completed marking, and to identify beneficial teaching points for the remainder of the school year. There was little evidence of consistency or moderation processes, except for contacting a relevant government officer for further explanation. The CWT is double marked, where both marks and the teacher optional mark are not disclosed. A moderator rectifies any discrepancies. There is no data available regarding correlation between teacher and marker assessment, but anecdotal evidence suggests it is high. In 2002, the Northern Territory Government is offering

professional development for teachers in marking the writing task. All tests contribute to national benchmark data.

6.1.8 South Australia

All tests are centrally marked in South Australia, but they are investigating using teacher assessments and judgment. In 2001 South Australia published a product termed Calibrated Assessment Tasks (Out of Print). This was designed for teachers to work individually with assessment tasks, and then to trial their own judgments against criteria and exemplars in the package. This has been discontinued as the tasks were limited, their use was dubious, and there was little durable change to teaching practice or assessments that could have been collected. This year South Australia will trial a process of professional development where teachers will come together to discuss assessment tasks, assessment criteria against the Curriculum Frameworks and Accountability Framework. They will trial the tasks and meet for moderation processes. This trial will be documented as a possible path for State collection of data on student performance using broader tasks and teacher judgments.

6.1.9 Victoria Achievement Improvement Monitor (AIM)

Teacher-assessed writing and mathematics occurs as part of the AIM program in years 3 and 5 (Victorian Curriculum and Assessment Authority 2000). Students write on a topic that fits in with normal classroom activities, or selected from three that have stimulus material supplied. Students discuss the stimulus, prepare a draft (using all resources in the classroom, including discussion with peers) and revise their work. Final versions are written individually and within a set time. Maths involves students performing applied mathematics tasks, with a timed teacher demonstration and timed task completion.

Writing is marked against criteria that relate to the Victorian outcomes approach Curriculum Standards and Framework support documents and accompanying elaborations. Teachers are given exemplars to explain criteria, and trial pieces to practise marking. Teachers assess mathematics by marking responses with a numerical score if responses are right or wrong. Maths tasks are hands on; encompass explanation and a variety of responses can be given to the tasks. Figure 14 gives an example of teacher-assessed mathematics tasks.

2. Look at the spinner:

If you were to spin a paperclip on this spinner, which number would it be most likely to land on?

Use the spinner with the paperclip as your teacher showed you. Spin it 20 times and record your result for each spin.

Number	Tally	Total
1		
2		
3		
4		
5		
6		

(b) Draw a bar graph (column graph) of your tally table.

(c) A pizza shop sells several different types of pizzas. Below is a list of the most popular types of pizza and how many of each type were sold in one week.

(i) Round each total to the nearest 10 units and write this value in the rounded amount column.

Type of pizza	Number sold in one week	Rounded amount
Vegetarian	80	
Supreme	187	
Hawaiian	118	
Mexican	108	
Chicken	90	

(ii) Using the rounded amounts and the key given, draw a pictograph of the number of each type of pizza sold. Please be sure to put a title on the graph and label each axis.

(d) What number did your paperclip land on most often? _____

(e) Would you expect this result? _____

Why? _____

STOP YEAR THREE STOP YEAR FIVE CONTINUE WORKING.

Teacher Assessment

Figure 14
Example of AIM teacher-assessed mathematics tasks

Writing prompts are a little lacking in breadth, purpose or audience requirements. Because of the 'process' focus, this task links to writing pedagogy. As the marking criteria are related to Victoria Curriculum and Standards Frameworks documents, this would assist in teacher professional development. Consistency in marking would be an issue, as moderation is not overemphasised. The purpose was for the teacher-assessed tasks could also be questioned, as the tasks are often too similar to standardised components.

6.2 Appraisal Conclusions

Discussed earlier, four stances were drawn from the literature, which were used to form a critical position to critique the samples of test materials and the above testing programs. Questions, based upon the stances, were asked of the materials and programs to see what they had to offer. Questions asked were:

- what kinds of teacher involvement are there in the program
- what are the implications the program has for pedagogy
- are teachers professionally developed by the program, and
- what consistency processes are used to assist in validity and reliability of marking?

There was a variety of ways teachers were involved in these tests. They ranged from teachers being formally trained to mark tasks, to teachers being involved in planning and developing tasks and marking tasks that their own students undertook. In programs where the stakes were higher, as in budgetary rewards for test results, formal training of teachers as markers resulted. Table 1 shows some of these trends.

Positive pedagogy links were reported when testing programs were related to local syllabus documents, effective assessment practices, or if the results could directly affect changes to teaching and learning programs for individual students. A number of the programs were about testing to outcome statements or performance standards, with marking criteria developed from these. 'Teaching to the test' was encouraged in these instances.

Professional development occurred when teachers were involved in task writing or marking. If teachers were only involved in administration of the test, gains for professional development was dubious. Where there was not a positive pedagogy link, professional development was limited to understanding only the test processes, which in itself, is helpful for teachers in understanding the Testing Program. The test assisted further in teacher development of effective assessment practices if these practices were valued in the Testing Program.

There were very few moderation practices occurring with the tests that were reviewed. Often there were chances for teachers to meet and share the marking process, but these were only encouraged, and not seen as vital to the process, nor was it stated how this sharing contributed to moderating the marks. Mostly, training of markers, especially when perceived test stakes were higher, was the way consistency was realised. In the collected sample it appeared that the more reliant the test stakeholders are in valid, reliable and objective assessments (as when performance data is used to report to National Benchmarks), the more test contractors turn to formal training of markers and central task construction (see table 1).

6.3 Conclusion

What examples are available of teacher-assessed tasks and teacher judgment processes in other statewide or international testing programs? What samples of materials can be gathered for analysis?

There are a variety of ways in which teachers assess test tasks beyond central marking. More extended test responses need more extensive marking processes and some of these include teachers marking their own students' tasks. While Indiana uses central marking processes, Maryland uses local marking of their multiple task and cross-curricular tests at different sites around the state. Both train markers rigidly in understanding marking rubrics and marking processes. Kentucky and Vermont include classroom teachers marking their own students for writing portfolio in Kentucky and writing and Mathematics portfolio in Vermont. Year-level teachers concerned attend out of school sessions in training to use assessment rubrics and criteria. Portfolio results are reported separately from writing demand, or in the case of Vermont, Mathematics multiple choice and short answer, test results.

Australian States also show a similar variety of teacher-assessed tasks. Australian Capital Territory employs classroom teacher assessment for their speaking test, as does Victoria in a writing and Mathematics tasks. Both present teachers with marking criteria and exemplars, in video and print format respectively. New South Wales uses local site marking for their teacher-assessed tests in writing and extended Mathematics response tests, while the Northern Territory offers teachers an option to mark their demand writing task before it is centrally marked. Western Australia uses teacher-assessed tasks in their sample tests of English and Mathematics, but their literacy and numeracy census tests are machine scanned. Tasmania is not using teacher-assessed tasks at the present time and South Australia is trialling teacher-assessed tasks along with processes of consistency of teacher judgment as a way to collect statewide performance data. All use standardisation processes and training of teachers to enhance validity and

reliability of test marking. There was little use of consistency processes in as far as assessments being moderated through formal procedures or informally through discussion with other markers.

Sample teacher-assessed tasks in testing programs from the USA offer the Queensland Testing Program some avenues to explore for the broadening of the program. As Maryland's cohort is similar to Queensland's the processes used to administer and assess their open-ended tasks could be worthy of investigation. Portfolio assessments similar to Kentucky and Vermont could be adapted for trial in Queensland to include extended response literacy and numeracy tasks related to Queensland syllabuses. The Speaking Test from Australian Capital Territory could also be adapted for the Queensland testing context, as could the processes Victoria uses for teacher-assessed extended mathematics response and teacher-assessed writing tasks.

The processes New South Wales use to locally mark their tests could be further investigated for their use in local site marking. Teacher assessment of demand writing tasks resembling the Northern Territory processes could also be used in Queensland. While not broadening the coverage, both would allow for more teacher involvement in the Queensland Testing Program.