NATIONAL TESTING OF DATA HANDLING IN YEARS 3, 5 AND 7 IN AUSTRALIA

Steven Nisbet
Faculty of Education, Griffith University, Australia
s.nisbet@griffith.edu.au

This paper provides a critical analysis of the data-handling items contained in the 2008 Australian national Years 3, 5, and 7 numeracy tests from a number of perspectives, namely, congruence with the concept of numeracy, alignment with the Australian National Statements of Learning for Mathematics (Curriculum Corporation, 2006), and alignment with the Statistical Thinking Framework developed by Jones et al. (2001). Overall the test items reveal a limited focus with respect to item content, type of tasks, and method of response. The majority of items limit students to analysing data presented in tables and graphs, with some items requiring interpretation of the data. Gaps in the data-handling skills required by students to answer the questions are identified.

INTRODUCTION

The demands of an evidence-based society include the expectation that schools provide evidence to the community that their students are performing to acceptable standards. Such evidence includes results of large-scale (e.g. national) testing of students at various stages of schooling, particularly in the areas of literacy and numeracy.

The Australian federal government adopted compulsory national testing of literacy and numeracy in 1997 as part of its National Literacy and Numeracy Plan. The purposes of the plan were to (a) identify students at risk, (b) conduct intervention programs, (c) assess all students against national benchmarks, and (d) introduce a national numeracy reporting system (Department of Education, Training, & Youth Affairs, 2000). For the first ten years of national testing, each state and territory of Australia was responsible for writing their own tests, conducting the tests in all schools, and reporting the results to schools and parents. Consequently, the test items were not common across the nation during that time. However, since 2008 the writing of the numeracy tests has been done nationally, and the test items have been based on the National Statements of Learning for Mathematics (Curriculum Corporation, 2006). Up until 2007, in the state of Queensland, the tests were conducted in Years 3, 5 and 7 only, but since 2008, the tests have been extended to Year 9. The tests are currently conducted in all schools in Australia, both government and non-government schools, in May of each year.

The introduction of state-wide and national testing has not been supported widely in the educational community in Australia. Research in Queensland (Nisbet & Grimbeek, 2004) found that teachers had very negative attitudes to state-wide tests. They did not believe that the tests were valid, and hence the results of the tests did not greatly influence their teaching practices. Nor did the teachers use the test results to any great extent to inform their lesson planning, apart from some teachers identifying gaps in their schools' mathematics programs.

National testing has limited validity considering (i) the accepted definitions and purposes of assessment and (ii) the nature of large-scale tests. Firstly, assessment can be defined as the comprehensive accounting of a student's or group of students' knowledge, and should be used as a means to achieve educational goals, and not interpreted as the end of an educational experience (Webb, 1993). The purposes of assessment in mathematics according to Clarke, Clarke and Lovitt (1990) are: (i) to improve instruction by identifying specific sources of student's error that requires remediation, or the specific learning behaviours that might need to be encouraged and developed or discouraged and replaced, (ii) to improve instruction by identifying those instructional strategies that are most successful, (iii) to inform the learner of identified strengths and weaknesses both in knowledge and in learning strategies so that the most effective strategies might be applied where most needed, (iv) to inform subsequent teachers of the students' competencies so that they more readily adapt their instruction to the students' needs, and (v) to inform parents of their child's progress so that they can give more effective support.

Secondly, large-scale numeracy testing in a geographically-spread country as Australia brings with it a number of logistical and budgetry limitations. For instance, the methods of students' responses adopted in the national numeracy tests have been limited to just two: (i) multiple choice ("Colour in the bubble), and (ii) single numerical answers ("Write the number in

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

the box"), which can be marked by computer scanning. With such limited methods of student response available, there are similar limitations on the questions and tasks that can be included in the tests. It is commonly accepted that through assessment, teachers communicate to students which activities and learning outcomes are valued, so that assessment should be comprehensive and give recognition to all valued learning experiences (Clarke, Clarke & Lovitt, 1990). It is suggested here that national testing communicates restricted messages and brings limitations including a bias towards mechanical processes and rules, and away from problem solving and creativity.

It has been known for many years that statistics items in traditional pencil-and-paper mathematics tests have shortcomings; they test skills in isolation from the problem context, they do not test whether or not students understand how statistical measures are interpreted, and they fail to assess students' ability to communicate using statistical language (Garfield, 1993). Similarly, it has been known that national testing programs have tended to result in classroom assessment moving away from authentic formative practices and towards techniques closely aligned to the national test format (Stiggins, 1999). Teachers feel compelled to spend time preparing their students to master the skills included on the tests. It is therefore against this background and a degree of scepticism on the part of the author that this critical analysis is conducted.

THE NATURE OF THE AUSTRALIAN NUMERACY TESTS

The Australian national numeracy tests are based on a broad definition of numeracy, viz. 'Numeracy is the effective use of mathematics to meet the general demands of life at school, at home, in paid work, and for participation in community and civic life' (MCEETYA Benchmarking Task Force, 1997, p. 4). Hence the tests cover the strands of mathematics which most people would meet in daily life, i.e., number, measurement, geometry, chance, and data. The Year 3 test contains 35 items, and the Year 5 test contains 40 items. There are two tests at the Year 7 level, one which is done without a calculator (32 items), and another in which a calculator is allowed (also 32 items).

The numeracy tests are required to be aligned with the National Statements of Learning for Mathematics for Years 3, 5, and 7 (Curriculum Corporation, 2006). For data handling, the Statement of Learning (SOL) for Year 3 refers to students organising data, presenting data, recognising variation in data sets. The Year 5 SOL refers to students collecting data, presenting data in tables and graphs, and describing variation in data. The Year 7 SOL refers to students identifying discrete and continuous data, representing data in two-way tables, using frequency and relative frequency, choosing suitable measures of location (mean, median & mode), describing distributions of sample data, identifying variation in data, and making informal inferences.

The method of response varies across items; some items are answered by multiple choice, in which students have to colour in with pencil one of the small 'bubbles' placed under four alternatives, and some items are open response, in which students write their answers in a small box (30mm x 12mm) provided on the page. Each paper has a set of three practice items to assist students become familiar with how to answer the questions.

Another feature of the tests is the use of 'link items', i.e., items which are common to two year levels, e.g., the Year 3 and Year 5 tests, or the Year 5 and Year 7 tests. For instance, seven of the 32 items in the Year 7 non-calculator test are linked to the Year 5 test, and another seven items in the Year 7 non-calculator test are linked to the Year 9 calculator test. These items are inserted to compare performance across grades levels. Typically, the percentage of students who answer a link item correctly in Year 7 is greater than the percentage of students who answer it correctly in Year 5. Similarly, the percentage of students who answer a link item correctly in Year 9 is greater than the percentage of students who answer it correctly in Year 7.

Of the total items in the national numeracy tests, the topic of data-analysis accounted for between 10% and 16% of the items. In comparison to the other strands of the mathematics syllabus, data has a smaller percentage of items than number and geometry, and has approximately the same percentage as measurement, but a greater percentage than probability. See Table 1.

ANALYSIS OF TEST ITEMS

The data-handling items are analysed below in relation to the following criteria: definition of numeracy and use of real-world contexts; type of stimulus material; type of task; alignment with the National Statements of Learning (Curriculum Corporation, 2006); and alignment with the Statistical Thinking Framework (Jones et al., 2001).

Table 1. Number and percentage of items in each strand of syllabus

	Year 3	Year 5	Yr 7 non-calc	Yr 7 (calc)	Yr 7 overall
Statistics	5 (14%)	4 (10%)	5 (16%)	3 (9%)	8 (12.5%)
Number & pattern	14 (40%)	16 (40%)	13 (41%)	10 (31%)	23 (36%)
Geometry	11 (31%)	12 (30%)	8 (25%)	11 (34%)	19 (30%)
Measurement	4 (13%)	6 (15%)	4 (13%)	6 (19%)	10 (16%)
Probability	1 (3%)	2 (5%)	2 (6%)	2 (6%)	4 (6%)
Total items in test	35	40	32	32	64

Analysis of items in relation to the definition of numeracy and use of contexts

The definition of numeracy that underlies the construction of the Year 3, 5, 7 and 9 numeracy tests was adopted by the Australian Ministerial Council on Education (comprising all state, federal and territory ministers of education) when national numeracy benchmarks were set in 1997, and reads as follows: 'Numeracy is the effective use of mathematics to meet the general demands of life at school, at home, in paid work, and for participation in community and civic life' (MCEETYA Benchmarking Task Force, 1997, p. 4).

All data questions in the tests were set in real-world contexts, however the contexts were mostly related to home life and past-times. Graphs presented for analysis were set in contexts such as: animals on a farm, favourite songs, suburbs where students lived, UV index at various times during the day, currency conversion, and mode of transport to school. Tables of data presented for analysis were set in contexts such as: books, exercise programs, sport, television, movie programs, mobile phone bills, car colours and types of DVDs. Although it can be seen that the contexts do relate to the students' 'real world' and hence would be easily understood by students, some of the data in the tables and graphs seem to be contrived and lack authenticity. Examples of authentic contexts include the UV index at various times of the day (Year 7, non-calculator test, Item 9) and movie programs (Year 5, Item 28). Less authentic examples include type of DVDs (Year 7, calculator test, Item 29) and Manu's exercise program (Year 3, Item 12).

Analysis in relation to type of stimulus material

The stimulus material for each item was either a table of data or a graph, with more items overall with the data presented in a table (59%) than in a graph (41%). See Table 2.

Table 2. Number of items according to type of stimulus material

	Year 3	Year 5	Year 7 (non-calc)	Year 7 (calc)	Total
Table of data	3	2	3	2	10 (59%)
Graph	2	2	2	1	7 (41%)
Sub-total	5	4	5	3	17

The types of graphs became more sophisticated going from Year 3 to Year 5 to Year 7. The graphs presented in Year 3 include one bar graph and one pictograph. In Year 5, there was one bar graph and one pie graph. In Year 7 there were three line graphs and one double bar graph.

Analysis in relation to type of task

The data-handling tasks typically required the students to identify one piece of data in a table or graph, compare or use two or more pieces of data in a table or graph, or link information from a table and a graph. The Year 7 tests had more of the type requiring students to compare or use two or more pieces of data in a table or graph than the Year 3 or Year 5 tests. See Table 3.

Analysis of items in relation to the Year 3 Statement of Learning for Mathematics

The Year 3 Statement of Learning (SOL) for data handling (Curriculum Corporation, 2006) is as follows: 'Students know that some questions and issues, including statements and questions they have created themselves, can best be answered by collecting data. They work out suitable ways to do this, including the use of technology to access existing data, and explore the best ways of organising it. They present the data in ways that assist its interpretation.' (p.6).

Table 3. Number of items in test according to type of task and year level of test

Type of task	Yr 3	Yr 5	Yr 7 (non-calc)	Yr 7 (calc)	Total
Identify one piece of data in a table	2 (Q10, 12)	1 (Q8)		1 (Q29)	4
Compare or use two or more pieces of data in a table	1 (Q18)	1 (Q28)	3 (Q18, 22, 24)	1 (Q4)	6
Identify one piece of data in a graph	1 (Q31)	1 (Q36)			2
Compare or use two or more pieces of data in a graph	1 (Q29)	1 (Q17)	2 (Q9, Q32)		4
Link information from a table and a graph				1 (Q1)	1

Clearly, there is an emphasis in the Statement of Learning (Year 3) on students being proactive in data collection, representation and interpretation, rather than responding to data presented to them. This makes for a mismatch with the items written for the Year 3 test requiring students merely to analyse and interpret other peoples' data sets, as presented in tables and graphs.

The items in the Year 3 test followed a common pattern; data were presented in a table or graph, and a question was then asked pertaining to the data where students had to identify one piece of data, e.g. 'How many push-ups did Manu do on Wednesday?' (Item 12), or compare or combine two or more pieces of data, e.g. 'How many books did Rani and Sam read altogether?' (Item 10), and 'In which game did Kim score 3 more goals than she scored in Game 2?' (Item 18). Hence it can be concluded that the Year 3 test items assess the Year 3 SOL in a limited way, focussing only on analysing data, and ignoring data collection, representation and interpretation. See the Appendix for two items from the Year 3 test.

Analysis of items in relation to the Year 5 Statement of Learning

The Year 5 Statement of Learning (SOL) for data handling (Curriculum Corporation, 2006) is as follows: 'Students use a range of ways of collecting data including surveys, observations, and experiments. They choose tables or graphs including technology-generated graphs to present the information. They use these to support statements or predictions they have made, or to convince them that additional data is required. They look for and describe expected or unexpected variation within sets of data they use.' (p.10). As with the Year 3 SOL, the emphasis in the Year 5 SOL is on students being proactive with data collection, representation and interpretation, rather than responding to data presented to them. However, the items written for the Year 5 numeracy tests require students merely to analyse other peoples' data sets, as presented in tables and graphs. The pattern in the Year 5 tests was that data were presented in a table or graph, and a question was then asked pertaining to the data, where students had to identify one piece of data, e.g., "How many students live in Scanlon" (Item 36), or compare two pieces of data, e.g., "Are there more goats than cows" (Item 20). Again, there is limited focus in the tests, with no reference to data collection, data representation or variation.

Analysis of items in relation to the Year 7 Statement of Learning

The Year 7 Statement of Learning (SOL) for data handling (Curriculum Corporation, 2006) reads: 'Students identify data as discrete or continuous, and use a variety of representations including two-way tables to summarise sample data obtained from a given population. They use frequency, relative frequency and choose suitable measures of location (mean, median, mode) as summary statistics to describe the distribution of sample data from a given context.' Again, the emphasis in the Year 7 SOL is on students being proactive with data summarisation, and representation, more than describing data presented to them. However, the items written for the Year 7 numeracy tests require students merely to analyse other peoples' data sets, as presented in tables and graphs. The 2008 tests had very little emphasis on data interpretation and variation within sets of data. The pattern in the Year 7 tests was that data were presented in a table or graph, and a question was then asked pertaining to the data, where students had to use two pieces of data, e.g., “How many people under the age of 20 had a monthly mobile phone bill of less than $30?” (Item 22, non-calculator test), or identify a piece of data and make a calculation, e.g., “What percentage of the cars is red?” (Item 24, non-calculator test). One item did require the students to calculate a measure of centre, namely, “What was the average time for Mick’s walks with his dog?” (Item 4, calculator test). The most difficult question was Item 32 (non-calculator test) with two graphs for converting currencies (Australian dollars to British pounds and Australian dollars to Brunei dollars). Students were asked to convert 50 British pounds to Brunei dollars.

Analysis in relation to the Statistical Thinking Framework (Jones et al., 2001)

The Statistical Thinking Framework was developed by Jones and his colleagues at Illinois State University after extensive research, development and validation with elementary school pupils. The purpose of the framework is to describe and predict students’ statistical thinking. The framework consists of four constructs (Describing Data Displays, Organising and Reducing data, Representing Data, & Analysing and Interpreting Data), which are described at four levels (Level 1 –Idiosyncratic, Level 2–Transitional, Level 3–Quantitative & Level 4–Analytical).

The construct *Describing Data Displays* relates to a student’s reading and describing displays such as graphs and tables, understanding the conventions used, recognising correspondences between two displays of the same data, and evaluating the effectiveness of different displays. *Organising and Reducing Data* refers to grouping and ordering data, being able to explain the basis for the organisation of the data, recognising when data reduction occurs, and being able to calculate measures of typicality and spread. *Representing Data* relates to being able to complete or produce displays of data i.e. various types of graphs, and includes the ability to produce a display which shows some reorganisation of the data. *Analysing and Interpreting Data* refers to being able to make comparisons, reading between the data, making inferences with the data (i.e. reading beyond the data), and recognising what a display does not say about the data.

This analysis was conducted to determine which constructs were tested in the 2008 numeracy tests, and what levels of thinking were expected of students. It revealed that the great majority of items (16 out of 17 items i.e., 94%) tested the construct ‘Analysing and interpreting data’, and of those 16 items, 11 of them (69%) related to analysing data and only 31% related to interpreting data. There were no items relating to describing data and representing data, and only one item relating to organising and reducing data. See Table 4.

Table 4. Items and levels of thinking required according to construct

Construct		Items and levels of thinking required				Totals
		Year 3	Year 5	Y7 (non-cal)	Y 7(cal)	
Describing data displays						0
Organising & reducing data					Q4 – level 3	1
Representing data						0
Analysing & interpreting data	Analysing data	Q10 – level 2 Q11 – level 1 Q18 – level 2 Q20 – level 2 Q31 – level 1	Q8 – level 1 Q20 – level 2 Q36 – level 1	Q9 – level 2 Q22 – level 2 Q32 – level 4		11
	Interpreting data		Q28 – level 3	Q18 – level 3 Q24 – level 3	Q1 – level 4 Q29 – level 3	5

Despite the great imbalance of constructs tested, the levels of thinking expected of students follow a reasonable progression as one moves from Year 3 to Year 5 to Year 7. In the Year 3 test, students are required to operate at levels 1 and 2. In the Year 5 test, students are required to operate at levels 1, 2 and 3, and in the Year 7 test, students are required to operate at levels 2 to 4.

CONCLUSION

Overall, the data-handling test items focus on a limited range of knowledge and skills compared to what is expected to be taught in schools according to the National Statements of Learning (Curriculum Corporation, 2006) and the constructs identified by the Statistical Thinking Framework (Jones et al., 2001). The items require students to respond to other people's data only, and respond in limited ways, i.e., mostly to analyse a given data set (in a table or graph), and sometimes, to interpret the data. There is no scope for students to be 'producers of statistics' by collecting, organising, reducing, and representing data. Nor is there much emphasis on variation in data. Clearly, if schools place too much emphasis on the style of items contained in national tests, and teach to the test (Stiggins, 1999), there will be an unfortunate narrowing of the curriculum.

REFERENCES

Clarke, D., Clarke. D., & Lovitt, C. (1990). Changes in mathematics teaching call for changes in assessment alternatives. In T. Cooney & C. Hirsch (Eds.) *Teaching and Learning Mathematics in the 1990s, NCTM 1990 Yearbook.* Reston, VA: National Council of Teachers of Mathematics.

Curriculum Corporation (2006). National Statements of Learning for Mathematics. Melbourne, VIC: Author.

Department of Education, Training, & Youth Affairs, (2000). *National Literacy and Numeracy Plan.* Canberra, ACT: Author.

Garfield, J. (1993). An authentic assessment of students' statistical knowledge. In N. Webb & A. Coxford (Eds.) *Assessment in the Mathematics Classroom. NCTM 1993 Yearbook* (pp. 187-196). Reston, VA: National Council of Teachers of Mathematics.

Jones, G., Langrall, C., Thornton, C., Mooney, E., Perry, B., Putt, I., & Nisbet, S. (2001). Using students' statistical thinking to inform instruction. *Journal of Mathematical Behavior, 20*, 109-144.

MCEETYA Benchmarking Task Force (1997). *National Report on Schooling in Australia.* Melbourne, VIC: MCEETYA.

Nisbet, S., & Grimbeek, P. (2004). Primary teachers' beliefs and practices with respect to compulsory numeracy testing. In I. Putt, R. Faragher & M. McLean (Eds.) *Mathematics Education for the Third Millennium. Proceedings of the 27th Conference of the Mathematics Education Research Group of Australasia. Townsville, June 2004*. MERGA.

Stiggins, RJ. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practices, 18*(1), 23–27.

Webb, N. (1993). Assessment for the mathematics classroom. In N. Webb & A. Coxford (Eds.) *Assessment in the Mathematics Classroom. NCTM 1993 Yearbook* (pp. 1-6). Reston, VA: National Council of Teachers of Mathematics.

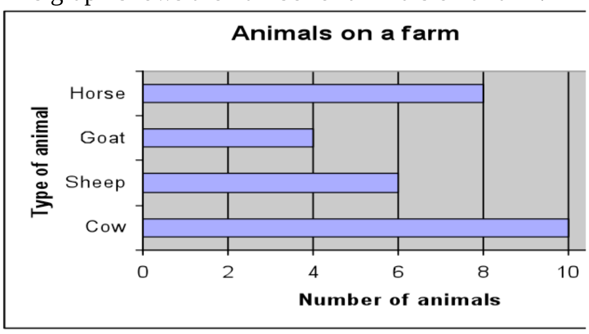
APPENDIX – Sample items from Year 3 test

This table shows the number of books read by some children.

Name	Number of books
Rani	2
Alicia	10
David	2
Jim	3
Sam	7

How many books did Rani and Sam read altogether?

This graph shows the number of animals on a farm.



Which statement is true? There are more goats than cows. There are more horses than cows. There are fewer sheep than goats. There are fewer sheep than horses.