# The responses of one school to the 2008 Year 9 NAPLAN numeracy test

**Stephen Norton**
Griffith University
<s.norton@griffith.edu.au>

## Introduction

In May 2008, the National Assessment Program — Literary and Numeracy (NAPLAN) (Ministerial Council on Education, Employment, Training and Youth Affairs, MCEETYA, 2008) conducted nation-wide tests for reading, writing, language conventions, and numeracy. A national assessment program of this kind should provide the stimulus for macro-level (systemic) analytical studies and micro-level (individual school and/or class) analytical studies of the data. This paper is concerned with the latter type of study where numeracy results of Year 9 Queensland students are analysed. The paper provides a critique of the Year 9 numeracy test, ways in which students responded to test questions, and how results might inform teaching mathematics.

The Federal Government wishes to ensure that all students achieve minimum standards in numeracy: band 2 by Year 3; band 4 by Year 5; band 5 by Year 7; and band 6 by Year 9. Hence, the establishment of a nation-wide scale of student achievement across 10 bands that accommodate different curricula across different states. It is intended that those students who are identified as having achieved below the minimum standard for their year level will be given focused intervention and additional support. During September 2008, individual student reports were delivered to parents, and schools also received school reports. The *NAPLAN Summary of the National Report* was released on 12 September by MCEETYA (2008). In order to gain a finer grain analysis of error patterns, the results of a single school, "Suburbia" (pseudonym), are discussed, since the author does not have access to state-wide data. Suburbia is an outer suburban school with a cohort of 100 Year 9 students. The community is relatively economically disadvantaged. The principal feeder primary schools were reported as having 61%, 55%, 52% or 44% of students below the national benchmark for numeracy in Year 7 (Chilcott, 2009).

## Overview and critique of the Year 9 NAPLAN numeracy test 2008

The test is named a "numeracy test." Two descriptors are applied to "thinking involving mathematical ideas:" numeracy and mathematics. Van Groenestijn (2002) suggested that numeracy is essentially the application of mathematics to social, personal and work situations. In this article, numeracy is defined as subset of mathematics, in which mathematical knowledge is required to function effectively in society.

There were two Year 9 NAPLAN tests: the "numeracy non-calculator" test where students were not permitted to use a calculating device; and a "numeracy calculator-allowed" test where students were permitted to use a calculating device.

### Calculator and non-calculator tests

Each of the 2008 papers consisted of 32 questions to be completed within 40 minutes. In reality, only for questions 14 and 22 of the calculator-allowed paper would the use of a calculator be of significant assistance to those students possessing reasonable computation ability. The remaining 30 questions on the calculator-allowed paper could be solved with short written computations or, in some cases, estimation. For example, question 24 required that 25 be substituted into the equation $E = 2T^2$ where $T$ is 25. Students with arithmetic competency can multiply 25 by 25 (= 625) and double that (= 1250). A calculator saves time but is not necessary for students who can multiply two-digit numbers.

An area of mathematics where a calculator might be expected to be of assistance is with computations involving fractions, decimals, and large or many numbers. An analysis of the calculator-allowed paper shows that many such questions can be easily solved using simple written algorithms or mental computation. Two examples follow.

Question 5 in the paper required students to find the number exactly half way between $1\frac{1}{4}$ and $3\frac{3}{4}$. Four choices were given. One way of doing this question is to do a mixed number subtraction. An alternative method is to place both mixed numbers on a number line together with the multiple choice options and either count out the quarters, separating each number, or estimate which option is half way. If a student were to use a calculator, they would have to convert each mixed number to a decimal, do the subtraction of decimals and then convert back to a mixed number. Hence, the availability of a calculator would do little to simplify the problem.

In question 9, students were given the information that there were "14 students in Rina's class on Wednesday. The other 11 were absent." They were asked, "What percentage of Rina's class was absent?" The solution involves finding the total enrolment by adding 11 and 14 to give 25, then expressing the fraction $\frac{11}{25}$ as a percentage. The choice by the item writer of 25 as the denominator obviates any need for the calculator, since the critical part of the question is to find the denominator. If students used a calculator, they would have to convert 0.44 to a percentage, a task no simpler than doing the problem mentally by multiplying both 25 and 11 by 4.

With few exceptions, numbers involved in almost all questions in the calculator-allowed paper indicated that the authors had chosen numbers to facilitate simple mental or written computations. The comparable means and standard deviations of the test results from both the calculator-allowed and the non-calculator tests for this school, as well as an analysis of demands of the questions (no-calculator-use mean = 11.51, SD = 5.21; and calculator-use mean = 11.43, SD 5.16), indicated that removing error patterns associated with computation by giving students a calculator had no effect on improving their outcomes. The most likely explanation of these data is that if students did not know the underlying structure, having a calculator did not help them solve the problem. Indeed, the presence of the calculator may have hindered the performance of some students, if, as so often is the case, the presence of a calculator encouraged students to rush to computation before considering the structure of the question.

## Multiple-choice formats

The main format was multiple-choice. Only five of the 32 questions on the calculator-allowed test required written answers (questions 23, 24, 29, 30 and 31). Nine questions on the non-calculator exam required a written answer (questions 4, 7, 16, 18, 19, 20, 23, 29, 32). Over the two tests, about 78% of the questions were multiple-choice, with four choices, and hence subject to students' guessing. Over both papers, guessing might account for about 20% of the correct answers.

Students could also obtain correct answers without necessarily understanding the mathematics in the question, particularly for algebra-based questions. For example, question 11 in the non-calculator test asked students to solve $3x - 5 = x + 1$. This question cannot be solved with "backtracking," which is a numerical rather than an algebraic method. Such questions are generally regarded as a good test of a student's ability to solve for an unknown. However, in a multiple-choice format, students can obtain the correct answer through simple substitution of each distracter. In the multiple choice format, student success in Queensland was 57%. In the same paper, question 23 has the same structure: "Find the value of $b$ in this equation: $5b - 4 = 2b + 17$." The success rate over the State was 29%. The inference is that the almost-double success rate seen in question 11 was due to students substituting the distracters, or guessing.

## Time limits

Students were expected to complete each paper in two separate sessions of 40 minutes duration. This allowed an average of $1\frac{1}{4}$ minutes for each question. Many of the questions would take most students several minutes to read and analyse. Table 1 indicates that fewer students were successful on the later questions. Possible reasons are that either students experienced increasing difficulties with the later questions and/or they rushed to complete each paper in the time allocated. The implication of this finding is that not only do students need to know the material, but must also be proficient in carrying out tasks rapidly.

## Strand coverage

NAPLAN categorises the test items into four strands with approximately equal numbers of questions in each strand:

1. *Number*: eight questions on each paper, 25% of the total;
2. *Algebra*: functions and patterns, eight questions on each paper, 25% of the total;
3. *Measurement, Chance and Data*: eight questions in the non-calculator test and seven questions in the calculator-allowed test, 23% of the total;

4. *Space*: eight questions in the non-calculator test and nine questions in the calculator-allowed test, 27% of the total.

The almost identical strand coverage and the similar structure of the test items further undermine the rationale for two tests.

The choice of strand content in the papers differs from the content identified by some States; for example, the Queensland Studies Authority (QSA, 2004) has five strands: Number; Patterns and Algebra; Measurement; Chance and Data; and finally, Space. The NSW syllabus (Board of Studies, 2002) has five strands: Number; Patterns and Algebra; Data; Measurement; Space and Geometry. In NSW, Chance is subsumed under Number. NAPLAN places Chance with Data and Measurement. By way of international comparison, the TIMSS (Thomson, 2006) international tests use strands of: Number; Algebra; Measurement; Geometry; and Data.

The basis of the NAPLAN classification is by context; for example, non-calculator paper question 1, which is classified as Number, shows students a sequence of a brick wall and a table of ordered pairs. Students are asked to calculate the number of bricks in the next pattern. Since this task involves recognising a growing pattern and extending that pattern, it could be classified in the Algebra strand because, while the context is Number, the process of extending a pattern is essentially an algebraic activity. Similarly, question 8 on the same test is classified by NAPLAN as being in the Space strand since it involves a square with area to be scaled. The same question could be classified as proportional reasoning within the Number strand, since the crux of the question demands an understanding of scale and ratio, irrespective of the context of Space. Alternatively, since there are measurement units involved, this question might be classified as belonging in a Measurement strand.

A further example of the difficulty in classifying questions is seen in question 11 of the non-calculator paper. To complete this question, students need to know that all sides of a square are equal in length in order to form the equation $3x - 5 = x + 1$. Since the geometric component is relatively simple, such a question would be said to primarily test Algebra, and NAPLAN classifies it as such. In some questions, it is less clear which sub-component of the respective questions students would find difficult, and this observation has encouraged me to put forward an alternative classification which is seen in Table 1 and described below.

Table 1. Summary of strand information on 2008 NAPLAN calculator-allowed and no-calculator-allowed tests combined.

| Strand | % of test | Number of questions |
|---|---|---|
| **Number**: Fraction and percentages reasoning | 12.5% | 8 |
| **Number**: Proportional reasoning | 11% | 7 |
| Other **number** (e.g.,decimals, index, and chance) reasoning | 8% | 5 |
| **Algebraic** reasoning | 28% | 18 |
| **Geometry** and geometric reasoning | 25% | 16 |
| **Measurement** based reasoning | 9% | 6 |
| **Data**-based reasoning | 6.25% | 4 |
| Total number of questions over both tests | 100% | 64 |

From the table it can be noted that there were eight questions based on fractions or percentages, approximately 12.5% of the test.

Based on the above analysis of the 64 questions comprising the numeracy tests, Table 1 indicates that 31% ($n$ = 8 + 7 + 5 = 20) of questions are Number, 28% ($n$ = 18) are based on algebraic reasoning, 25% ($n$ = 16) are associated with geometric reasoning and only six questions relate to Measurement and four to data-based reasoning. This split is similar to the Year 8 TIMSS test (Thompson, 2006): Number 28%; Algebra, 25%; Geometry 15%; Measurement 15%; and Data 16%.

Almost all the questions associated with algebra and geometry in the NAPLAN tests cannot be completed without a good understanding of number computation, most of which is associated with multiplicative thinking.

## An analysis of Suburbia's performance

The recently released report *A Shared Challenge: Improving Literacy, Numeracy and Science Learning in Queensland Primary Schools* (Masters, 2009) indicates that there are challenges associated with the teaching of mathematics in all States, including Queensland. This analysis of Suburbia's performance informs State curriculum design.

Suburbia's results provide an opportunity to examine error patterns in the absence of state-wide data. The results are first examined by comparing student performance on the numeracy test to others in the State and nation according to the NAPLAN banding system (Table 2). The banding conversion weights harder questions more heavily.

Table 2. Proportional achievement of sample school, State and nation by band.

| Band | Suburbia % | Qld % | National % |
|---|---|---|---|
| Band 10 | 1.2 | 4.3 | 8 |
| Band 9 | 0 | 12.3 | 14 |
| Band 8 | 17.6 | 25.2 | 25 |
| Band 7 | 25.9 | 30.2 | 28 |
| Band 6 | 35.3 | 20.2 | 18 |
| Band 5 or less | 20 | 6.6 | 5 |

Band 6 is considered the national minimum standard for Year 9. It is clear that Suburbia's scores are below the State average in numeracy. The level of numeracy at Suburbia, while below Queensland averages, might be representative of a significant proportion of the State's—and indeed the nation's—children. At Suburbia, about one in three students performed at or above the State mean, and about two in three below the State mean. Notably, few students excelled, with only about 19% of Year 9 students achieving Band 8 or better, compared to 42% for the State and 47% for the nation. Conversely, 55% of the school scored at or below the minimum standard of Band 6 compared to 27% for Queensland and 23% for the nation. This lack of representation of students achieving well and/or at a satisfactory level, and a prevalence of those in the lower bands, is reflected in a school mean of 527 compared to the State mean of 573 and a national mean of 582 (scaled scores).

## Gender, streaming and Indigenous issues

School NAPLAN reports indicate that 42% of boys and 61% of girls scored at or above the State mean. It is likely that these are educationally significant gender differences. It is beyond the scope of this paper to examine the gender differences except to note that girls have demonstrated a higher achievement level than the boys, a reversal of earlier findings. A considerable body of literature exists on this topic (e.g., Ethington, 1992; Fennema, 1996; Maccoby & Jacklin, 1994).

Of the 98 students enrolled at Suburbia, seven were Indigenous. One Indigenous student had an overall score above the school mean, one at the school mean and the remaining four had overall scores below the school mean (Indigenous mean = 495; school mean = 527). Interestingly, students classified as having a language background other than English (LBOTE) may be over-represented with scores above the school mean (LBOTE mean = 538 compared with school mean of 527). The small sample sizes samples and the clumping of the data about the mean suggests caution should be exercised before any speculation as to the educational significance of this finding is attempted.

Analysis of the class results indicated that there were very clear differences between test performances of the classes within the school. The school's Head of Department (HOD) of Mathematics confirmed that the school had a policy of streaming students in Year 9.

While the overall school achievement indicates challenges, it is worth determining if students obtained relatively better scores in particular strands. These findings are shown in Table 3, and indicate that students were challenged across all strands.

Table 3. Student achievement in strands.

| Strand | Mean | SD | % Correct |
|---|---|---|---|
| **Number**: Fraction and percentages reasoning | 2.98/8 | 1.93 | (37%) |
| **Number**: Proportional reasoning | 2.07/7 | 1.67 | (30%) |
| Other **number** (e.g., decimals, index, and chance) reasoning | 1.71/5 | 1.36 | (34%) |
| **Algebraic** reasoning | 5.38/18 | 3.24 | (29%) |
| **Geometry** and geometric reasoning | 6.19/16 | 3.57 | (38%) |
| **Measurement**-based reasoning | 1.57/6 | 2.07 | (26%) |
| **Data**-based reasoning | 1.52/4 | 1.03 | (38%) |

A more detailed summary of the relative performance of students who answered each question correctly is shown in Table 4. In Table 4, the percentage of students in the school who answered each question correctly is compared with the percentage of all students who answered the question correctly in Queensland. The relative difference in correct response is in brackets. The national strand is indicated by the letter beside the question number. For example, "4 m" means question 4, NAPLAN classification Measurement.

With regard to the Number strand, only 49% of students could recognise that 11/4 is equivalent to $2\frac{3}{4}$. The concept of fractions is introduced in Year 4. Similarly, only 8% of students in Year 9 at Suburbia school and 19% of the State's Year 9 students succeeded with the problem: $6.6 \div 0.3 = 22$ (question 21 on the non-calculator test). Question 7 on the non-calculator paper asked students to calculate how many US dollars could be purchased

Table 4. Categorisation and student performance on individual questions.

| | Strand description | % Correct responses | |
|---|---|---|---|
| Number | Fractions and percent reasoning (no calculator) | School | State |
| 14 n | Mixed numbers to improper fractions | 49 | 66 (–17) |
| 19 n | Part of whole as a percentage | 19 | 34 (–15) |
| 20 m | Fraction calculation | 33 | 39 (–6) |
| 21 n | Decimal division | 8 | 19 (–11) |
| 24 m | Proportional reasoning converted to a decimal | 46 | 53 (–7) |
| | Fractions and percent reasoning (with calculator) | | |
| 5 n | Mixed numbers and number line | 58 | 70 (–12) |
| 9 n | Simple percent | 57 | 71 (–14) |
| 14 n | Simple percentage | 43 | 44 (–1) |
| | Proportional reasoning (no calculator) | | |
| 7 n | Money conversions | 40 | 53 (–13) |
| 8 s | Scale | 61 | 75 (–14) |
| 22 n | Reasoning number line | 29 | 32 (–3) |
| 29 n | Proportional reasoning and percent | 12 | 18 (–6) |
| | Proportional reasoning (with calculator) | | |
| 11 s | Proportion and geometry (with calculator) | 36 | 42 (–6) |
| 25 s | Ratio and fractions in geometric setting | 40 | 44 (–4) |
| 28 n | Ratio | 21 | 23 (–2) |
| | Number based reasoning (no calculator) | | |
| 10 m | Probability | 48 | 53 (–5) |
| 25 a | Consecutive numbers | 25 | 28 (–3) |
| 27 n | Index notation | 15 | 33 (–18) |
| | Number based reasoning (With calculator) | | |
| 1 n | Simple decimal subtraction | 71 | 80 (–9) |
| 22 n | Decimal multiplication and subtraction computation | 14 | 26 (–8) |
| Algebra | Algebraic reasoning (no calculator) | School | State |
| 1 n | Recognising and extending a pattern | 69 | 90 (–21) |
| 2 a | Matching data with a graph | 60 | 74 (–14) |
| 11 a | Solving with variables on both sides | 43 | 57 (–14) |
| 13 a | Expanding and expression | 34 | 41 (–7) |
| 23 a | Solving with variables on both sides | 8 | 29 (–21) |
| 26 a | Gradient function | 25 | 45 (–20) |
| 28 a | Substitution including exponent | 25 | 27 (–2) |
| 30 a | Graphing inequalities | 8 | 16 (–8) |

Brackets following the State % correct indicates difference between the school and State % correct.

| | Strand description | % Correct responses | |
|---|---|---|---|
| Alg. cont. | Algebraic reasoning (with calculator) | School | State |
| 2 a | Simple substitution | 60 | 80 (–20) |
| 3 a | Describing relationships from a graph | 60 | 80 (–20) |
| 8 a | Substitution fraction format | 33 | 61 (–28) |
| 12 a | Describing a relationship, non linear | 35 | 42 (–7) |
| 18 a | Substitution with a decimal value | 12 | 48 (–36) |
| 21 a | Developing an equation | 24 | 32 (–8) |
| 23 n | Extending a pattern | 26 | 49 (–23) |
| 24 a | Substitution including exponent | 10 | 30 (–20) |
| 27 a | Developing an equation | 15 | 37 (–22) |
| 31 a | Solving | 5 | 13 (–8) |
| Geometry | Geometry and geometric reasoning (no calculator) | | |
| 3 s | Angles in a triangle – solve | 38 | 30 (–8) |
| 5 s | 3D spatial visualisation | 52 | 62 (–10) |
| 6 s | Co–ordinate application | 72 | 79 (–7) |
| 9 m | Geometric reasoning and proportion | 51 | 68 (–17) |
| 15 s | 3D spatial visualization | 41 | 58 (–17) |
| 17 s | 2D and 3D geometric reasoning | 31 | 42 (–11) |
| 32 s | 2D knowledge and reasoning | 5 | 15 (–10) |
| | Geometry and geometric reasoning (with calculator) | | |
| 6 s | Geometric reasoning nets | 75 | 83 (–8) |
| 7 s | Simple symmetry | 81 | 87 (–6) |
| 10 s | Co–ordinate geometry | 33 | 59 (–26) |
| 13 s | Analytic geometry | 27 | 51 (–24) |
| 15 m | Application of pie | 36 | 47 (–11) |
| 16 s | Analytic geometry | 21 | 31 (–10) |
| 17 m | Geometric reasoning | 48 | 63 (–15) |
| 20 s | Geometric definitions | 40 | 61 (–21) |
| 26 n | Co–ordinates and direction | 40 | 52 (–22) |
| Measurement | Measurement reasoning and logic (no calculator) | | |
| 12 s | Measurement logic | 58 | 72 (–14) |
| 16 m | Using a protractor | 38 | 66 (–28) |
| 18 m | Measurement of volume and proportion | 27 | 46 (–19) |
| | Measurement reasoning and logic (with a calculator) | | |
| 19 m | Time calculation | 46 | 45 (+1) |
| 29 n | Distance logic | 6 | 24 (–18) |
| 30 m | Area of net | 2 | 5 (–3) |
| Data | Data based reasoning (no calculator) | | |
| 4 m | Stem and leaf | 58 | 53 (+5) |
| 31 m | Mean median mode | 25 | 24 (+1) |
| | Data based logic (with calculator) | | |
| 4 m | Concept of mean | 56 | 75 (–19) |
| 32 m | Mean median mode | 13 | 31 (–18) |

if they had 50 Australian dollars, and AU$1 buys US$0.80. Only 43% of the school and 53% of the State managed to compute the correct answer. These suggetexamples indicate the challenges students face in the learning of fractions, percent, and proportional reasoning.

Students also struggled with Algebra; for example, the correct response rate to question 23 (What is the value of $b$ in this equation? $5b - 4 = 2b + 17$, non-multiple-choice) was 8% for the school and 29% for the State. Students also experienced difficulty with substitution, with 25% correct when substituting with an exponent; 33% correct when "3" was substituted with the unknown variable occurring in both the numerator and denominator of an expression; and 12% correct when the substituted number was a decimal (3.75), even when a calculator was at hand.

In the Geometry strand, results were similar to other strands. Only 38% of the students correctly answered the multiple-choice question where they were required to find an unknown angle in a triangle, given that one angle was 35° and the other a right angle (indicated by the conventional symbol). The few questions in which most students did achieve satisfactory results were those to which students in the middle primary years might be expected to give correct responses: simple symmetry, 81% correct (question 7); and reasoning with nets, 75% correct (question 6, both in the calculator-allowed paper).

One question was relatively well done in the Measurement strand: 46% for the school (1% above the State) succeeded in finding the hours and minutes between 3.27 am and 2.16 pm (question 19, multiple-choice in the calculator-allowed paper). Encouragingly, the Year 9 NAPLAN results in 2008 showed that 58% of Suburbia students correctly answered a stem and leaf question (5% above the State average) and 25% correctly identified the correct response in a multiple-choice question on mean and mode.

## Discussion

### Implications for test writers

The similarity of balance of Strands, structure of questions, and, in particular, the use of relatively simple numbers in the calculator-allowed test suggest that there is an argument for having one non-calculator-allowed test, rather than two separate tests. There is little evidence that removing error patterns in computation processes by allowing the use of calculators advantaged students.

Well-constructed multiple-choice questions have been reported to be highly reliable in assessing student competency, including mathematics (e.g., Bridgemen, 1991; Hopkins, George & Williams, 1985). Norcini, Swanson, Grosso and Webster (2009) found that written solutions to patient-management problems and multiple-choice responses in the same domain essentially measured the same thing. On the other hand, Berg and Smith (2006), in their study of students' abilities to construct and interpret line graphs, found that multiple-choice and free-response instruments showed significant discrepancies. These authors concluded that in some instances, multiple-choice questions were not a valid measure of abilities (Berg & Smith, 2006).

Analysis of the 2008 Year 9 NAPLAN results indicates that the predominance of multiple-choice formats added an element of variability to the test which became increasingly important among students who struggled with mathematics. Guessing or substitution could account for about 60% of the marks of a middle-performing student at the school, and, in some cases, for all of the marks. The effect of multiple-choice questions in inflating achievement marks was seen on those questions where students performed relatively poorly on questions which required them to write an answer, compared to similarly structured questions where there was a multiple-choice option.

The classification of questions by NAPLAN according to *context* rather than the *critical thinking* required to answer the questions has implications for teaching allocated to particular mathematics strands. The NAPLAN classification suggests that 16 questions are Number; it could be argued that 20 questions are Number. There are also inconsistencies between the way that NAPLAN and different States classify the strands. These differences might affect the allocation of time for the study of particular topic domains. This is particularly important for questions that essentially involve fractional thinking (fractions, decimals, proportion, and percent). By classifying questions as Measurement rather than Number questions (e.g., 9, 10, 18, 20 and 24 on the non-calculator test), the true extent of the importance of fractional thinking is obscured.

## Implications of the test results for Suburbia's work program

The results of this paper indicate that there needs to be an emphasis on teaching the fundamental concepts of mathematics. This consideration may be particularly important given the context of the school, which has inherited mathematically-disadvantaged students from primary feeder schools. The data indicate a need to focus on the students' learning basic number facts, mental strategies and algorithms for fractions, percent, and decimals, rates and ratio, linear algebra, and plane geometry for about 75% of the learning time. An analysis of Suburbia's work program below illustrates that there is a mismatch between the study time allocated to particular strands and the test weighting for these strands.

The benchmark levels in the "Junior Mathematics Work Program" of Suburbia are consistent with the *Mathematics Years 1 to 10 Syllabus* (QSA, 2004) and the *Essential Learnings* curriculum (QSA, 2007). Neither of these documents suggests the proportion of learning time to be spent on particular concepts. This has implications for Suburbia's programs for Year 8 and 9, which are analysed below.

The school's work program defined "numeracy" as: "Numeracy within the mathematics classroom is identifying the mathematics in a context relevant to the student" (Suburbia Work Program, 2008, p. 7). Consistent with this definition, the program was organised about integrated projects, each of five weeks' duration. The projects are briefly described below. However, only the first two units of the Year 9 program are described, since the NAPLAN tests are conducted in early May and subsequent work becomes irrelevant to student performance.

Number-learning underpins each of the domains shown in Table 5 and the projects offered opportunities for students to learn critical number concepts. However, there is a lack of alignment with the weighting of NAPLAN tests; for example, the data project represents 10% of Year 8 and 9 study, but 6.5% of the NAPLAN tests. Geometry and measurement projects (Units 3, 5, 6 and 7) take up almost half of the learning time but are 34% of NAPLAN questions. Algebra projects occupy 20% of Year 8 and 9 learning time but 28% of the NAPLAN test. This disjunction in terms of

Table 5. Suburbia's Year 8 and 9 Work Program.

| **Unit** (all durations 5 weeks) | **Key domains** |
|---|---|
| Year 8 | |
| 1. What does the school think about? | Data collection and display |
| 2. Fundraising for a charity | Computations in financial contexts some algebra |
| 3. Inside a beehive | Aspects of number, measurement, and geometry |
| 4. Chance | Chance |
| 5. Time and location | Measurement of time and geometry in mapping contexts |
| 6. Designing a container | Measurement and geometry |
| 7. Forensics | Geometry and measurement |
| 8. Linear functions | Algebra; introduction to variables, expressions and solving linear equations. |
| Year 9 | |
| 1. Measurement | Perimeter and area |
| 2. Linear equations | Algebra; writing, graphing and solving linear equations |

the time allocated is, however, relevant to the needs of the students who come to the school.

As noted earlier, the feeder schools have high proportions of students classified as being below the national numeracy benchmark (44% to 61%, Chilcott, 2009). The Year 9 results for 2008 indicate that the school has been unable to remediate the lack of knowledge of basic skills of students from the feeder schools. In late 2008 it was debated among mathematics Heads of Department in the district whether the Year 8 and Year 9 work programs ought to focus on developing fundamental number competency. The Head of Department of Mathematics at Suburbia explained that most of the students were "so disengaged that [he was] very reluctant to change to a more formal, and potentially less contextual, and thus less engaging, program." In the context of many students' prior mathematics learning, attempts to take account of students' attitude are understandable.

Suburbia's work program contains statements about pedagogy based upon assumptions about how students learn. It reflects the philosophy and recommendations suggested by the State syllabus documents which advise that students should become "active investigators" (QSA, 2004, p. 3), and that a "learner centred approach" be adopted (QSA, 2004, p. 10). The school work program recommends working by "investigation … in teams or groups working cooperatively and collaboratively towards shared goals." Not all authors agree that the extensive use of investigations is an effective use of student learning time (e.g., Kirschner, Sweller & Clark, 2006). Stacey (2003) found that curriculum documents of some States over the past decade showed less emphasis on computational skill and algebra procedures, but a greater emphasis on students obtaining deep understandings of concepts frequently in contextual settings. This did not necessarily translate to improved test results. There is, therefore, a danger that too much time be spent on investigations where a few concepts might be explored in depth and not enough time spent learning and understanding mathematical structures and processes.

### Implications of the test results for State- and school-based assessment

In Queensland, schools write their own assessment items for the majority of junior mathematics assessment. In this school, half of all assessment in Years 8 and 9 focuses on “extended tasks.” Typically, these involve students working in groups to investigate aspects of mathematics in a particular context, and writing a report. For example, the 2009 Year 9 task requires students to “analyse data to make rainfall predictions and investigate how rainfall is measured. They then use local rainfall data to choose an appropriate storage tank and plan how to make effective use of the collected water” (QSA, 2009, P. 2). This type of assessment is different from the NAPLAN tests and thus it is recommended that the school consider some alternative assessment formats.

## Concluding remarks

Analysis of the 2008 NAPLAN tests suggests that further research into the categorisation of test items according to strand is warranted. The current classification has the effect of under-representing the importance of the Number strand, particularly those concepts associated with fractional thinking. Error-pattern analysis suggests that there is merit in pursuing further research in regard to the use of multiple-choice questions. The data indicated that this format tended to inflate student marks, particularly among those students with limited mathematical knowledge. The test writers might consider the use of additional short answer questions.

Learning is a complex activity influenced by the interaction of variables including: cultural expectations (e.g., Eccles et al., 1993; Harrison & Huntington, 2000); student attitudes and beliefs (e.g., Ethington, 1992; Fennema, 1996; Thomson & Fleming, 2003); teacher knowledge of subject and pedagogy (e.g., Stacey & Chick, 2004; Thomson & Fleming, 2003); curriculum documents, including support material and textbooks (e.g., Groves, Mousley & Forgasz, 2006); and work programs as a manifestation of curriculum intentions. Of these variables, it is the implementation of the work program that is most relevant to this paper. Clearly there is a need for synergy between national expectations as reflected in NAPLAN tests and individual school work programs. In this case, the data indicate a case for greater focus on teaching the fundamental concepts of Number. There is also evidence that the school should explore ways to accommodate expectations of State curriculum documents and testing processes and the demands of NAPLAN testing. It is possible that a reduction in report-writing assessment for mathematics is warranted.

The very poor results of many students in basic facts on the NAPLAN tests and the results published by Chilcott (2009) indicate that there is a need for Suburbia to collaborate with the primary feeder schools in regard to the development of critical Number understanding. A further consideration is that few secondary teachers are trained to teach middle primary school mathematics concepts. This suggests that there is a need to conduct professional development for secondary teachers on how to diagnose, and then remediate, misconceptions associated with the middle and upper primary school Number strand.

The results of this analysis of one school’s 2008 Year 9 NAPLAN report has given the school a wealth of data on which to reflect. The analysis provides background to further investigation in that school. The reflections may be of assistance for teachers in other schools who may have similar problems and wish to analyse their MCEETYA reports and plan their responses to the national testing processes.

## References


Berg, C. & Smith, P. (2006). Assessing students' abilities to construct and interpret line graphs: Disparities between multiple-choice and free response instruments. *Science Education, 78*(6), 527–554.

Board of Studies, NSW (2002). *Mathematics K–6 syllabus 2002*. Sydney: Board of Studies.

Bridgeman, B. (1991). Essays and multiple-choice tests as predictors of college freshman GPA. *Research in Higher Education, 32*(3), 319–332.

Chilcott, T. (23 May 2009). Primary schools report. *Courier Mail*.

Eccles, J., Adler, T., Futterman, R., Goff, S., Kaczala, C., Meece, J., Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco: W. H. Freeman.

Ethington, C. (1992). Gender differences in a psychological model of mathematics achievement. *Journal for Research in Mathematics Education, 23*(2), 166–181.

Fennema, E. (1996). Scholarship, gender and mathematics. In P. F. Murphy & C. V. Gipps (Eds), *Equity in the classroom: Towards effective pedagogy for girls and boys* (pp. 73–80). London: Falmer Press.

Groves, S., Mousley, J. & Forgasz, H. (2006). *Primary numeracy: A mapping, review and analysis of Australian research in numeracy learning at the primary school level.* Canberra: Department of Education, Science and Training.

Harrison, L. & Huntington, S. (2000). *Culture matters: How values shape human progress*. New York: Basic Books.

Hopkins, K., George, C. & Williams, D. (1985). The concurrent validity of standardized achievement tests by content area using teacher's ratings as criteria. *Journal of Educational Measurement, 22*(3), 177–182.

Kirschner, P; Sweller, J. & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constuctivist, discovery, problem-based, experimental and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Maccoby, E. E. & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford Palo Alto, CA: Stanford University Press.

Masters, G. (2009). *A shared challenge: Improving literacy, numeracy and science learning in Queensland primary schools*. Camberwell: Australian Council for Education Research.

Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA). (2008). *Numeracy non-calculator and calculator allowed National Assessment Program Literacy and Numeracy*. Carlton: Curriculum Corporation.

National Numeracy Review Panel & National Numeracy Review Secretariat. (2007). *National numeracy review: Background paper June 2007*. Canberra: Commonwealth of Australia.

Norcini, J., Swanson, L., Grosso, L. & Webster, G. (2009). Reliability, validity and efficiency of multiple choice questions and patient management problem item formats in assessment of clinical competence. *Medical Education, 19*(3), 238–247.

Queensland Studies Authority. (2004). *Mathematics Years 1–10 syllabus*. Spring Hill: Author.

Queensland Studies Authority. (2007). *Essential learnings and standards*. Spring Hill: Author.

Queensland Studies Authority. (2009). *Information statement: QCATs 2009 suggested reporting comments*. Accessed from http://www.qsa.qld.edu.au/assessment/3163.html.

Stacey, K. (2003). The need to increase attention to mathematical reasoning. In H. Hollingsworth, J. Lokan & B. McCrae (Eds), *Teaching mathematics in Australia: Results from the TIMSS 1999 video study* (pp. 119–122). Melbourne: Australian Council for Educational Research.

Stacey, K. & Chick, H. (2004). Solving the problem with algebra. In K. Stacey, H. Chick & M. Kendal, (Eds), *The future of teaching and learning of algebra. The 12th ICMI study* (pp. 1–20). Boston: Kluwer.

Thomson, S. (2006). *Australian student's achievement in TIMSS 2002 mathematics cognitive domains. TIMSS Australian monograph No. 9.* Melbourne: Australian Council for Educational Research.

Thomson, S. & Fleming, N. (2003). *Summing it up: Mathematics achievement in Australian schools, TIMSS 2002.* Melbourne: Australian Council for Educational Research.

Van Groenestijn, M. (2002). *A gateway to numeracy. A study of numeracy in adult basic education* [CD-ROM]. Utrecht, The Netherlands: Press, Universiteit Utrecht.