GENDER DIFFERENCES IN NAPLAN MATHEMATICS PERFORMANCE

JANELLE C. HILL

Monash University

janelle.hill@monash.edu

The results of the National Assessment Program – Literacy and Numeracy [NAPLAN] tests of Australian students in 2008, 2009, and 2010 were analysed by individual question responses (percentage correct) of females and males. The analysis of Grade 3 and Grade 9 data demonstrate that a decline in the achievement of females is evident and these gender differences become larger as students progress through their schooling.

Introduction

National Assessment Program – Literacy and Numeracy [NAPLAN] testing commenced in 2008. As part of the program, students in Grades 3, 5, 7 and 9 are simultaneously tested using national tests in Reading, Writing, Language Conventions (Spelling, Grammar and Punctuation) and Numeracy (Australian Curriculum, Assessment and Reporting Authority [ACARA], 2010a). As explained by ACARA (2010b), the tests broadly reflect aspects of literacy and numeracy in all States and Territories. The types of test formats and questions are chosen so that they are familiar to teachers and students across Australia. The Victorian Curriculum and Assessment Authority [VCAA] (2010a) affirms that "questions for NAPLAN tests are developed with reference to the nationally agreed Statements of Learning which reflect the core elements of the curriculum documents used in the different states and territories" (p. 2)

Results of the National Assessment Program – Literacy and Numeracy, in 2008, 2009 and 2010, the Program for International Student Assessment [PISA], in 2006 and 2009, and the Trends in International Mathematics and Science Study [TIMMS], in 2007, raise significant concerns with respect to the mathematical achievement of Australian females. The PISA is an initiative of the Organisation for Economic Co-operation and Development [OECD] in Paris and reports on indicators in education (Thomson & DeBortoli, 2008). Similarly, the goal of TIMMS is to provide comparative information on educational achievement across countries to improve teaching and learning in mathematics and science (Thompson, Wernert, Underwood & Nicholas, 2008).

Thomson and DeBortoli's (2008) analysis of the 2006 PISA results showed that Australian males performed significantly better in mathematics than Australian females. This was not the case in PISA 2003 where, although males achieved a mean score

higher than females, the difference was not statistically significant (Thomson & DeBortoli, 2008). Thomson, De Bortoli, Nicholas, Hillman and Buckley's (2010) analysis of the 2009 PISA data showed that Australian males scored ten points higher, on average, than Australian females in the mathematical literacy assessment. They go on to say that was a significant difference. Thomson et al.'s (2008) analysis of TIMMS data presents similar findings. They found that nationally, at the Grade 8 level, there was a substantial and significant gender difference in favour of males. Supporting the findings of the analysis of PISA data (Thomson & DeBortoli, 2006), Thomson et al. (2008) found that this significant gender difference in favour of males, that had not previously been seen in 2003 or 1995, appeared to be due to a significant decline in the average score for females over the 1995-2007 time span. This change in achievement should raise alarm; the fact that there were significant declines in the scores of female students that had not been observed for over a decade suggests that the current education system may not be effectively providing girls with equal opportunities for success in mathematics and so is particularly relevant with respect to the current analysis.

The results of the analyses of data from PISA in 2006 and 2009, and TIMMS in 2006 all point towards a considerable decline in the mathematics achievement of females that has not been observed for approximately a decade. The aim of the present analysis is to determine whether a similar decline in females' mathematics performance is evident in NAPLAN data for all grade levels taking the test (3, 5, 7, and 9) and for the three years 2008 to 2010.

Methodology

For the current analysis, the percentage correct achieved by both females and males on each individual NAPLAN question (provided by VCAA) was utilised. The percentage of correct responses by gender and grade level was compared from each test from 2008, 2009 and 2010. The data from each grade level was then grouped according to its mathematical dimension as described in the Victorian Essential Learning Standards [VELS] (VCAA, 2010b). The four mathematical dimensions of the VELS are: Number; Space; Structure; and Measurement, Chance and Data. The percentage difference in each question, by dimension, was then grouped into intervals of 5 percentage points. Questions for which there was no difference in percentage correct by females and males (ie. 0% difference) have been removed from the 0-4% group and reported separately; and the interval of 1-4% used.

Results and discussion

The results of the NAPLAN tests conducted in 2008, 2009 and 2010 for Grades 3 and 9 were analysed in order to investigate what gender differences exist between the performance of students from the highest and lowest grade levels undertaking NAPLAN testing. These tests consisted of between 31 and 35 questions from each of the Mathematics Dimensions as described in the VELS (with the exception of the Grade 3 tests in 2008, which did not include Structure questions). Grade 9 students undertake two NAPLAN tests; one in which they are able to use a calculator and one in which calculator use is not permitted. The number (% in brackets) of each question type is presented in Table 1.

Table 1. NAPLAN question types by dimension: Grade 3 and 9, 2008-10.

Grade		2008	2009	2010
		n	n	n
Grade 3	N[1]	14 (40%)	11 (31%)	13 (37%)
	M[2]	11 (31%)	10 (29%)	10 (29%)
	Sp[3]	10 (29%)	10 (29%)	10 (29%)
	St[4]	0	4 (11%)	2 (5%)
Total		35 (100%)	35 (100%)	35 (100%)
Grade 9 No Calculator	N	8 (25%)	9 (29%)	9 (28%)
	M	8 (25%)	8 (26%)	10 (31%)
	Sp	8 (25%)	7 (23%)	7 (22%)
	St	8 (25%)	7 (23%)	6 (19%)
Total		32 (100%)	31 (100%)	32 (100%)
Grade 9 Calculator	N	8 (25%)	8 (26%)	10 (31%)
	M	7 (22%)	8 (26%)	9 (28%)
	Sp	9 (28%)	8 (26%)	7 (22%)
	St	8 (25%)	7 (22%)	6 (19%)
Total		32 (100%)	31 (100%)	32 (100%)

[1] Number dimension
[2] Measurement, Chance and Data dimension
[3] Space dimension
[4] Structure dimension

The proportions of questions from each dimension remained relatively constant from 2008 to 2010 although there was an increase in the number of Measurement, Chance and Data [MCD] questions in the Grade 9 No Calculator test with a corresponding decrease in the number of Structure questions. Similarly, in the Grade 9 Calculator test, there was an increase in the number of Number and MCD questions in conjunction with a decrease in the number of Structure and Space questions.

When considering all NAPLAN questions (all years and all grade levels) in which a difference existed in the percentage of females and males who correctly answered a particular question, there were fewer questions in which females performed better than males, with males outperforming females at all grade levels. The differences in favour of males were also much larger than those favouring females. Table 2 shows these performance differences for Grades 3 and 9. A percentage difference of 1-4% for females indicates that a higher percentage of females than males answered this question correctly. It must be noted that although there were several questions for which females scored 5-9% better, there was no larger difference than 7% across all years and grade levels; that is, there were a few questions for which 7% more females than males answered correctly but this was the largest difference found for all years and grade levels.

Table 2. Number of questions with differences in percentage correct by grade level and gender; Grades 3 and 9.

		Girls				Boys				Total
Percentage difference	0*	1-4	5-9	10-14	≥15	1-4	5-9	10-14	≥15	
Grade 3 2008	5	10	2	0	0	7	10	1	0	35
Grade 3 2009	4	7	3	0	0	16	4	1	0	35
Grade 3 2010	3	9	3	0	0	15	5	0	0	35
Grade 9 No Calculator 2008	2	4	1	0	0	12	7	5	1	32
Grade 9 No Calculator 2009	5	6	0	0	0	13	6	1	0	31
Grade 9 No Calculator 2010	3	6	0	0	0	9	12	1	1	32
Grade 9 Calculator 2008	3	8	1	0	0	14	5	1	0	32
Grade 9 Calculator 2009	3	7	2	0	0	11	8	0	0	31
Grade 9 Calculator 2010	5	4	1	0	0	13	9	0	0	32

* Questions for which females and males performed equally well.

NAPLAN 2008–2010 Grade 3

When comparing the performance of Grade 3 students, Tables 3 and 4 show that there were no differences in the performance of females and males for 5 questions in 2008, 4 questions in 2009 and 3 questions in 2010. As the number of test questions in each year remained constant at 35, these results represent a small decline in the number of questions in which females and males performed equally well.

Table 3. Number and type of questions with percentage differences in favour of females - Grade 3.

Dimension	Number					Measurement, Chance and Data					Space					Structure				
% difference	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15
Grade 3 2008	2	1	1	0	0	2	2	1	0	0	1	7	0	0	0	-	-	-	-	-
Grade 3 2009	2	1	1	0	0	1	3	2	0	0	1	2	0	0	0	0	1	0	0	0
Grade 3 2010	1	4	0	0	0	1	1	2	0	0	1	3	1	0	0	0	1	0	0	0

Table 4. Number and type of questions with percentage difference in favour of males - Grade 3.

Dimension	Number					Measurement, Chance and Data					Space					Structure				
% difference	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15
Grade 3 2008	2	4	5	1	0	2	2	4	0	0	1	1	1	0	0	-	-	-	-	-
Grade 3 2009	2	7	0	0	0	1	3	1	0	0	1	5	2	0	0	0	1	1	1	0
Grade 3 2010	1	6	2	0	0	1	4	2	0	0	1	5	0	0	0	0	0	1	0	0

Grade 3 females appeared to do well in Space questions. In 2008, 7 of 12 Space questions were answered correctly by a higher percentage of females than males. This result was not repeated in 2008 or 2009. Instead, males outperformed females in Space questions in these years. Of the questions for which a higher percentage of females than

males answered correctly, the difference in percentage was no greater than 7%. In contrast, there were two questions in which the difference favouring males was greater than 10%. Differences in performance, although small, are already obvious at the Grade 3 level, with males outperforming females in all dimensions.

NAPLAN 2008 – 2010 Grade 9 (No calculator)

In the Grade 9 No Calculator test results it can be seen that there were very few questions with no percentage difference in the performance of females and males. Tables 5 and 6 show that far fewer questions were more likely to be answered correctly by females than by males. Females outperformed males by more than 5% in only one MCD question. Males outperformed females for most questions, especially those from the Number and MCD dimensions.

Table 5. Number and type of questions with percentage differences in favour of females: Grade 9 No Calculator.

Dimension	Number					Measurement, Chance and Data					Space					Structure				
% difference	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15
Grade 9 No Calculator 2008	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	3	0	0	0
Grade 9 No Calculator 2009	0	0	0	0	0	0	2	0	0	0	3	1	0	0	0	2	3	0	0	0
Grade 9 No Calculator 2010	1	2	0	0	0	0	2	0	0	0	0	1	0	0	0	2	1	0	0	0

Table 6. Number and type of questions with percentage difference in favour of males: Grade 9 No Calculator.

Dimension	Number					Measurement, Chance and Data					Space					Structure				
% difference	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15
Grade 9 No Calculator 2008	1	2	2	2	0	1	1	2	2	1	0	4	3	1	0	0	5	0	0	0
Grade 9 No Calculator 2009	0	6	2	1	0	0	4	2	0	0	3	2	1	0	0	2	1	1	0	0
Grade 9 No Calculator 2010	1	4	2	0	0	0	2	5	1	0	0	2	4	0	0	2	1	1	0	1

In two questions, (one MCD and one Structure), males dramatically outperformed females. In these questions, a difference of greater than 15% existed in the percentages of males and females who correctly answered. These results can be contrasted to the results from Grade 3 students. In Grade 3, a much higher number of questions were more likely to be answered correctly by females than can be observed in the Grade 9 data, demonstrating the decline in achievement of females as they progress through their schooling.

NAPLAN 2008 – 2010 Grade 9 (Calculator)

Again, in the Grade 9 Calculator test results, there were few questions for which there was no difference in performance by gender, with results being similar to those from the Grade 9 No Calculator data. As shown in Tables 7 and 8, questions for which a higher percentage of females than males answered correctly comprise a greater proportion than those from the Grade 9 No Calculator data. Females appear to do best in Structure type questions. Interestingly, there was a reduction in the number of Structure questions from 2008 to 2010. It is clear that in Grade 9, both males and females do better on the NAPLAN Calculator tests than on the Non Calculator tests, but males dramatically outperform females in both test types.

Table 7. Number and type of questions with percentage differences in favour of females: Grade 9 Calculator.

Dimension	Number					Measurement, Chance and Data					Space					Structure				
% difference	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15
Grade 9 Calculator 2008	1	1	0	0	0	2	1	0	0	0	0	1	0	0	0	0	5	1	0	0
Grade 9 Calculator 2009	0	2	0	0	0	2	1	1	0	0	0	1	0	0	0	1	3	1	0	0
Grade 9 Calculator 2010	0	2	1	0	0	0	0	0	0	0	4	0	0	0	0	1	2	0	0	0

Table 8. Number and type of questions with percentage differences in favour of males: Grade 9 Calculator.

Dimension	Number					Measurement, Chance and Data					Space					Structure				
% difference	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15	0	1-4	5-9	10-14	≥15
Grade 9 Calculator 2008	1	5	1	0	0	2	1	2	1	0	0	6	2	0	0	0	2	0	0	0
Grade 9 Calculator 2009	0	3	3	0	0	2	0	4	0	0	0	6	1	0	0	1	2	0	0	0
Grade 9 Calculator 2010	0	2	5	0	0	0	5	4	0	0	4	3	0	0	0	1	3	0	0	0

Conclusion

The results of this analysis of NAPLAN data support the findings of Thomson and DeBortoli (2006), Thompson, Wernert, Underwood and Nicholas (2008) and Thomson, De Bortoli, Nicholas, Hillman and Buckley (2010). It is evident that the mathematics achievement of females in Australia is on a decline and it appears that gender differences in favour of males become larger as students progress through their schooling. As stated by Forgasz (2008), 'gendered patterns from the past are still evident in the context of contemporary mathematics education in Australia' (p. 13). The issue of gender differences in mathematics achievement needs to be brought to the

forefront of educational research before such differences become even more pronounced.

References

Australian Curriculum, Assessment and Reporting Authority (2010a). *National Assessment Program – Literacy and Numeracy*. Retrieved 1 September, 2010, from http://www.naplan.edu.au/home_page.html

Australian Curriculum, Assessment and Reporting Authority. (2010b). *National Assessment Program – Literacy and Numeracy: Tests.* Retrieved 1 September, 2010, from http://www.naplan.edu.au/tests/tests_landing_page.html

Forgasz, H. (2008). Stars, compass, and GPS: Navigating currents and charting directions for mathematics education research on gender issues. In M. Goos, R. Brown & K. Makar (Eds.), *Navigating currents and charting directions* (Proceedings of the 31st Annual Conference of the Mathematics Education Research Group of Australasia, pp. 5–15). Brisbane: MERGA.

Thomson, S. & DeBortoli, L. (2008). *Exploring scientific literacy: How Australia measures up: the PISA 2006 survey of students' scientific, reading and mathematical literacy skills.* Camberwell, Vic: ACER Press.

Thomson, S., De Bortoli, L., Nicholas, M., Hillman, K. & Buckley, S. (2010). *Challenges for Australian education: Results from PISA 2009.* Retrieved 16 January, 2011, from http://www.acer.edu.au/documents/PISA-2009-Report.pdf

Thomson, S., Wernert, N., Underwood, C. & Nicholas, M. (2008). *TIMSS 2007: Taking a closer look at mathematics and science in Australia*. Retrieved December 5, 2010, from http://research.acer.edu.au/timss_2007/2

Victorian Curriculum and Assessment Authority. (2010a). *Ministerial Council on Education, Employment, Training and Youth Affairs Reporting Guide NAPLAN*. Retrieved 1 December, 2010, from http://www.vcaa.vic.edu.au/vcaa/prep10/naplan/schools/VICReportingGuide2010.pdf

Victorian Curriculum and Assessment Authority. (2010b). *Structure of the mathematics domain.* Retrieved 15 December, 2010, from http://vels.vcaa.vic.edu.au/maths/structure.html