

Numeracy Development Project Longitudinal Study: Patterns of Achievement

Gill Thomas
Maths Technology Ltd
<gill@nzmaths.co.nz>

Andrew Tagg
Maths Technology Ltd
<andrew@nzmaths.co.nz>

The central aim of the Numeracy Development Projects (NDP) is to improve student performance in mathematics through improving the professional capability of teachers. The aim of the NDP Longitudinal Study is to investigate the longer term impact of the NDP on student achievement in number and in mathematics more generally. This paper reports on the mathematics achievement of students in 26 schools that participated in the NDP prior to 2004. Not surprisingly, the findings suggest that the project has had the strongest impact on the students' performance on number items that are directly related to the NDP. The achievement of students as measured against the Number Framework indicates that, at most year levels, strategy level attainment has increased over time.

Background

From their outset, the Numeracy Development Projects (NDP) have been conceived as a dynamic, evidence-based initiative (Ministry of Education, 2006). The Numeracy Projects have evolved through a series of phases that can be summarised as the formulation phase (pre-2000), the initiation phase (2000–2001), the implementation phase (2002–2006), and the planned sustainability or maintenance phase (post-2007). The focus on the professional capability of mathematics teachers “provided the order and logic for the formulation phase” (Higgins, Parsons, & Hyland, 2003, p. 162). In the formulation phase, the Ministry of Education identified the key issues facing mathematics achievement as:

developing teachers' pedagogical content knowledge; improving teaching quality and confidence; providing resources to support mathematics teaching and learning; making research information more readily available and accessible to teachers; and emphasising the importance of mathematics education prior to school entry. (Higgins et al., 2003, p. 162)

The initiation phase began in 2000 with a national pilot of Count Me In Too for years 1–3 (Thomas & Ward, 2001), and a year 4–6 exploratory study (Higgins, 2001). The findings from these two projects informed the development of the NDP, with the introduction in 2001 of the Early Numeracy Project, the Advanced Numeracy Project, and the Exploratory Study (years 7–10).

The major implementation of the NDP has occurred since 2002, with over 17 000 teachers and 460 000 students estimated to have participated by the end of 2005 (Parsons, 2005). The projects are nationally co-ordinated, with their facilitation contracted to the six major teacher education providers in New Zealand. The Ministry of Education, through a series of evaluations, monitors the quality of the implementation and, by the end of 2005, has published a total of 25 reports and papers on the various components of the NDP.¹

While the focus from 2002 to the present has been on the widespread implementation of the NDP, a number of strategies were established in 2004 to address issues of sustainability. These included increased support for lead teachers, providing training opportunities for teachers new to schools that have already participated in the project, and providing online resources to all schools and teachers.

¹ See www.nzmaths.co.nz/Numeracy/References/reports.aspx for the evaluation reports and papers published by the Ministry of Education.

By 2007, almost all teachers of students in years 1–8 will have had the opportunity to participate in the NDP and the focus of the NDP facilitation contracts will shift to consolidation and maintenance (Parsons, 2005). In this phase, the challenge will be to maintain the strategic focus, quality, and momentum of the NDP while shifting the ownership for ongoing development to the school level.

The NDP Longitudinal Study, which began in 2002, has focused on examining the impact of the projects on the mathematics achievement of students in the years following the implementation of the NDP in the schools involved in the study. The information obtained from the longitudinal schools on the achievement of their students as measured against the Number Framework has helped inform expectations of achievement and progress over time (Thomas & Tagg, 2004, 2005a; Thomas, Tagg, & Ward, 2003). The performance of students in the longitudinal schools on items from the Trends in Mathematics and Science Study (TIMSS) suggests that the NDP has had a positive impact on mathematics achievement at years 4 and 5 and to a lesser extent at year 8 (Thomas & Tagg, 2004, 2005b).

The 2005 Longitudinal Study further investigated the mathematics performance of selected year groups of students and continued to track the achievement of all students as measured against the Number Framework. This paper first reports on the performance of year 4, 5, 6, and 8 students on items from TIMSS, the National Education Monitoring Project (NEMP), and Assessment Tools for Teaching and Learning (asTTle). The second section of the paper reports on the performance on the Number Framework of students at all year levels in 20 of the 26 schools participating in the 2005 Longitudinal Study.

Mathematics Achievement

Sample and Methodology

The aim of the mathematics achievement component of the NDP Longitudinal Study is to investigate the longer term impact of the NDP on student achievement in mathematics more generally.

The Longitudinal Study began in 2002 with the participation of 20 schools that first implemented the NDP in either 2000 or 2001. Five of the original 20 schools withdrew from the Longitudinal Study at the start of 2004, but the total number of schools in the Longitudinal Study was increased to 31 through the inclusion of 16 schools that first participated in the NDP in 2002. Sixteen of these 31 schools continued to participate in 2005, and 10 new schools were added to the sample. Each year, new schools are randomly selected from a list of schools that completed NDP training in the previous years. The list is stratified by decile to ensure that the sample in the Longitudinal Study closely approximates the national sample and has similar numbers of students in years 1–8. Four of the schools added in 2005 were intermediate schools, to increase the numbers of year 7 and 8 students involved.

Table 1 shows the breakdown of students by year level and decile band. The low-decile band includes decile 1–3 schools, the medium-decile band includes decile 4–7 schools, and the high-decile band includes decile 8–10 schools. As illustrated by Table 1, there is a disproportionate proportion of year 8 students from medium-decile schools in the sample because one of the intermediate schools selected had over 1 000 students.

Table 1
Analysis of Students by Year Level and School Decile Band

Year	Low	Medium	High	Total
4	251	381	310	942
5	270	377	320	967
6	299	405	313	1017
8	298	739	157	1194
Total	1118	1902	1100	4120

Tests for each of the four year levels were developed from items that had previously been used to assess the mathematics achievement of New Zealand students. There were four sources for these items: TIMSS 1995, TIMSS 2003, the 2001 NEMP mathematics assessment, and asTTle. Items from TIMSS 1995 and the NEMP assessment were used for the tests for years 4 and 8. The year 5 test was comprised of items solely from TIMSS 2003, and the only source of items for the year 6 test was asTTle. Table 2 shows the source of the 24 items contained in each test. The items were selected according to two criteria. Firstly, items were selected to ensure all strands of the New Zealand mathematics curriculum were covered (see Table 3). Secondly, the items were selected so that the average score in each test would be 50%, based on the percentage of New Zealand students answering each item correctly in the source assessments. The tests were piloted in a school that was not participating in the Longitudinal Study to check that they took approximately 40 minutes to administer.

Table 2
Source of Items in the Tests

Source	Year 4	Year 5	Year 6	Year 8
TIMSS 1995	14			16
TIMSS 2003		24		
NEMP 2001	10			8
asTTle			24	
	24	24	24	24

Table 3
Analysis of Strand of Items in the Tests

TIMSS Content Category	Year 4	Year 5	Year 6	Year 8
Algebra	4	3	1	4
Geometry	4	4	4	4
Measurement	4	5	2	4
Number	8	8	10	8
Statistics	4	4	7	4
	24	24	24	24

Test scripts were sent to each of the participating schools in July. The classroom teachers administered the tests, following instructions adapted from those used with TIMSS. The tests were sent back to the researchers for marking during August. Once the scripts had been marked, a report was compiled for each of the participating schools. This report included details on the item responses of each student and their overall test score. The schools' average performance by item and overall test was compared with the performance of same-age peers in the assessments from which the items were obtained.

Results

All reporting of results in this section is based on the average percentage of items answered correctly by students in the stated sub-groups. Longitudinal results refer to the 2005 Longitudinal Study unless stated otherwise. For each item, the 95% confidence limits for the difference in mean proportion between the longitudinal sample and the source sample were calculated. This is the criteria used to define significant differences in the results reported below.

Table 4 shows the performance of longitudinal students on the source assessments over the last three years. The first time a year 6 test was administered was in 2005, so there is no comparative data for this year level. The year 4 longitudinal students have performed consistently significantly better overall than the New Zealand students in the source assessments. The year 5 students had stronger performances in 2003 and 2004 than in 2005, while the year 8 students have continued to perform at a similar level to the year 8 students from the source assessments.

Table 4
Average Score by Year Level and Year of Testing

	Year 4	Year 5	Year 6	Year 8
Longitudinal 2003	56%*	59%*	N/A	52%
Longitudinal 2004	56%*	58%*	N/A	53%
Longitudinal 2005	57%*	54%*	47%	51%
Source average	50%	50%	50%	50%

* $p < 0.05$

Low-decile students in years 5, 6, and 8 performed not only lower than the medium- and high-decile schools, but also slightly lower than the 50% average achieved by New Zealand students in the source assessment (see Table 5). In year 4, the low-decile students scored lower than medium- and high-decile students but similarly to students in the source assessments. There is no decile information available on the source assessments, so no comparisons can be made by decile. The differences between high- and medium-decile students were smaller than those between low- and medium-decile students at all year levels.

Table 5
Average Score by Year and Decile Level

	Year 4	Year 5	Year 6	Year 8
Low decile (1–3)	46%*	39%*	36%*	37%*
Medium decile (4–7)	58%*	58%*	49%*	56%
High decile (8–10)	65%*	62%*	54%*	56%

* $p < 0.05$

Table 6 shows the numbers of number-related items on which longitudinal students performed significantly better or significantly worse than students in the source assessments. In each instance, the numbers in brackets represent the numbers of items where the difference was greater than 10%. The results show that, at all year levels, the longitudinal students performed either better or similarly on 34 of the 38 NDP-related items and either similarly or worse on all 13 number items not directly related to the NDP. The only exception is in year 5, where students performed similarly on the one non-NDP number item. Number items classified as non-NDP included calculations presented in vertical form, calculations involving large numbers, and expressions involving inequalities. Longitudinal students performed either better or similarly on 39 of the 45 non-number (Other) related items. Students at all year levels performed significantly better on 23 out of the 38 NDP-related items and significantly worse on only four items.

Table 6

Performance on NDP, Non-NDP-related Number Items, and Other Items in Relation to Source Assessments

Year	Item type								
	NDP-related			Non-NDP (Number)			Other		
	Declined	Similar	Improved	Declined	Similar	Improved	Declined	Similar	Improved
4		2	7 (4)	3 (3)				5	7 (4)
5	1	6	4 (1)		1		2	7	3 (1)
6	2 (2)	2	6 (3)	2 (2)			4 (4)	4	4 (4)
8	1	1	6 (1)	7 (4)				3	6 (3)
Total	4	11	23	12	1	0	6	19	20

* $p < 0.05$

Performance Highlights

As reported in Table 6, longitudinal students performed significantly better on the majority of items related to NDP topics. These items included both number knowledge items and items involving operating with numbers.

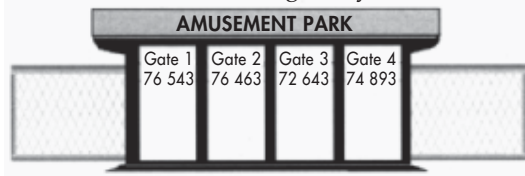
In the year 4 test, the item with the largest improvement was question 12 (Figure 1), in which students were asked to identify the largest of four numbers. Of the longitudinal students, 79% answered correctly, compared with 45% of New Zealand students in TIMSS 1995 (NZ TIMSS 1995). It should be noted that the longitudinal students' performance is similar to that of students internationally in TIMSS 1995 (76%).

12. Which of these is the largest number?			
A. 2735			Percentage
B. 2537		Longitudinal	
C. 2573		NZ TIMSS 1995	
D. 2753		TIMSS 1995	
			79
			45
			76

Figure 1. Item 12

Similarly, question 13a in the year 6 test (Figure 2) required students to identify the smallest of four numbers, in this case, in the context of people passing through amusement park gates. Of the longitudinal students, 83% answered correctly, compared with 76% of New Zealand students in the asTTle norms.

13. These are the gateways to an amusement park. Each gate shows the number of people who have visited during one year.



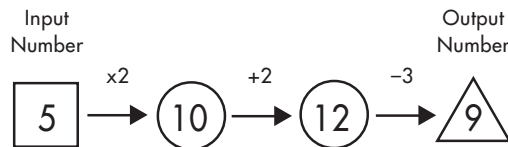
	Percentage
Longitudinal	83
asTTle	76

- a. Which gate did the least number of people use? _____

Figure 2. Item 13

Question 20 in the year 5 test, illustrated in Figure 3, was a number machine from which students were asked to identify the result of a series of operations on a number. Of the longitudinal students, 53% identified the correct response, compared with 45% of New Zealand students and 50% of international students in TIMSS 2003.

20. A number machine takes a number and operates on it. When the Input Number is 5, the Output Number is 9, as shown below.



	Percentage
Longitudinal	53
NZ TIMSS 2003	45
TIMSS 2003	50

When the Input Number is 7, which of these is the Output Number?

- A. 11
- B. 13
- C. 14
- D. 25

Figure 3. Item 20

Item 18 in the year 6 test (Figure 4) asked students to identify the temperature 9° hotter than negative 5° . Of the longitudinal students, 52% did so correctly, compared with 29% of students in the asTTle norms. It is encouraging to note that longitudinal students are able to operate with negative numbers in this context.

18. The temperature was -5° . It rose 9° .
What is the temperature now? _____ $^{\circ}$

	Percentage
Longitudinal	52
asTTle	29

Figure 4. Item 18

Longitudinal students performed particularly strongly on items related to fractions, an area in which students in New Zealand had performed poorly in TIMSS 1995 (Garden, 1997). Question 7 in the year 4 test asked the students to write a fraction that was larger than two-sevenths. Of the longitudinal students, 63% did so correctly, compared with 38% of New Zealand students and 41% of international students in TIMSS 1995. In 2004, 58% of year 4 longitudinal students answered this question correctly. Question 9 in the year 5 test required the students to find one-third of 600, with 44% of longitudinal students able to do so, compared with 34% of New Zealand and 49% of international students in TIMSS 2003. Question 6 in the year 8 test asked students to identify the largest of four fractions ($\frac{5}{6}$, $\frac{5}{7}$, $\frac{5}{8}$, $\frac{5}{9}$). Of the longitudinal students, 73% were able to do so, compared with 56% of students in NEMP 2001.

Performance Concerns

Items on which longitudinal students performed poorly included those too difficult to calculate mentally. It is a cause for concern that many students, particularly at the higher year levels, have no strategy for dealing with large or more complex calculations.

As shown in Figure 5, the numbers presented in question 18 in the year 4 test are too large and, with three columns adding to greater than 10, too complicated to be readily added using a mental strategy. Written forms are not introduced until the higher stages of the Number Framework, so typically from stage 6, few year 4 students would be expected to answer this question correctly. This result is consistent with the findings from the 2004 Longitudinal Study, in which 38% of year 4 longitudinal students correctly answered the same question (Thomas & Tagg, 2005b).

18. Add $\begin{array}{r} 6971 \\ +5291 \\ \hline \end{array}$

A. 11 162

B. 12 162

C. 12 262

D. 1 211 162

	Percentage
Longitudinal	34
NZ TIMSS 1995	47
TIMSS 1995	67

Figure 5. Item 18

Question 15 in the year 6 test asked students to give the answer to the number sentence $5024.6 - 2975.8 = \square$. Of the longitudinal students, 15% answered correctly, compared with 39% of students in the asTTle norms. The numbers of digits in this sum are again too large to be solved mentally.

Three items in the year 8 test (Figures 6–8) required students to carry out number operations where the numbers were too large to be easily calculated mentally. While written forms do not appear until the later stages of the Number Framework, it would be hoped that by year 8, students would have been introduced to techniques for calculating sums of this type. The difference is made more significant by the fact that in the two items from TIMSS 1995, the New Zealand performance was already significantly lower than the international average.

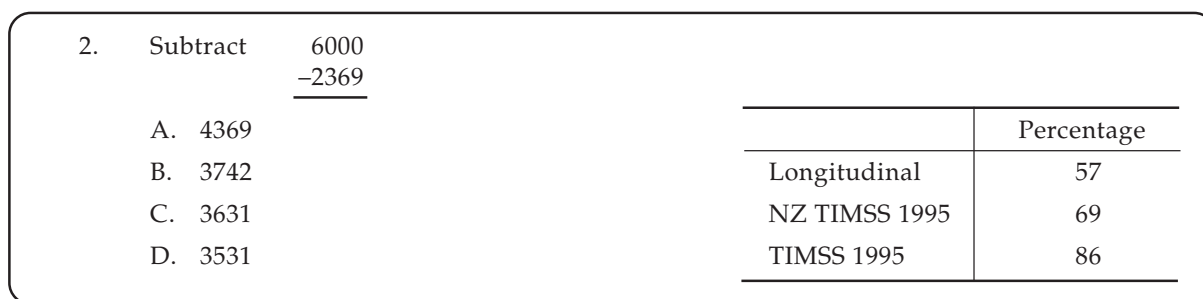


Figure 6. Item 2



Figure 7. Item 5



Figure 8. Item 22

Another area in which longitudinal students performed poorly in relation to students in the source assessments was in problems involving inequalities.

It was surprising that only 45% of year 4 longitudinal students were able to correctly identify that 442 is greater than 436 (Figure 9), as number sequence is a key focus of the NDP and 79% of year 4 students from the NDP 2003 were able to order numbers up to 1000 (Higgins, 2004). It can only be assumed that the students were unfamiliar with the symbols for greater than and less than.



Figure 9. Item 21

In item 20 of the year 8 test (Figure 10), students were asked to find an inequality equivalent to one given. Of the longitudinal students, 28% were able to do so, compared with 42% of New Zealand students and 36% of international students in TIMSS 1995.

20. $\frac{x}{2} < 7$ is equivalent to

A. $x < \frac{7}{2}$

B. $x < 5$

C. $x < 14$

D. $x > 5$

E. $x > 14$

	Percentage
Longitudinal	28
NZ TIMSS 1995	42
TIMSS 1995	36

Figure 10. Item 20

Achievement on the Number Framework

Sample and Methodology

This component of the Longitudinal Study investigates the longer-term impact of the NDP on student achievement in number. More specifically, it examines the stages that year 1–8 students achieve on the strategy domains of the Number Framework over time. This information helps inform benchmarks or targets set by schools to monitor the achievement of their students.

The 26 schools participating in the 2005 Longitudinal Study were asked to enter the strategy stage results of their students on the online database by the end of November 2005. Not all schools were able to do this, so the deadline was extended and schools were given the option of providing their results to the researchers as an Excel spreadsheet. Twenty of the 26 participating schools provided the requested information for inclusion in the analysis reported in this section.

Table 7 shows a breakdown of students by year level and school decile band. The largest anomaly is at year 8, where there were no student results from high-decile schools. This is due to the only high-decile intermediate school in the sample failing to submit results.

Table 7

Analysis of Students by Year and School Decile Band

Year	Low	Medium	High	Total
1	319	347	346	1012
2	251	227	277	755
3	217	241	292	750
4	281	221	325	827
5	252	192	300	744
6	296	240	278	814
7	92	745	39	876
8	358	690		1048
Total	2066	2903	1857	6826

Some schools do not collect information on all three strategy domains, so a decision was made to report on strategy results by Global Strategy Stage (GloSS). A student's GloSS is usually

determined by using the GloSS assessment forms² but for the purposes of this study, the GloSS result for each student was determined by taking the highest strategy stage reported for that student across the three strategy domains. The GloSS results of students from schools that participated in the NDP in 2005 were calculated in the same way.

Four hundred questionnaires were sent to all teachers in the 2005 longitudinal schools, with an estimated³ 60% return rate. The questionnaire included items on numeracy assessment, student achievement, and numeracy targets.

Teachers' Perspectives on Numeracy Achievement

The large majority (94%) of teachers reported tracking the numeracy strategy stages of students in their class, with 70% indicating that their school had developed targets for student achievement related to the NDP. Table 11 summarises the percentage of students that the teachers believed were "on track" to achieve the school numeracy targets. As shown by Table 8, 70% of teachers believe that at least 70% of the students in their class were on track to achieve the school numeracy targets.

Table 8
Percentage of Students Perceived by Teachers as Achieving Numeracy Targets

Percentage of students achieving targets	Number of teachers	Percentage of teachers
Less than 50	11	9
50–59	11	9
60–69	15	12
70–79	23	19
80–89	26	22
90–100	35	29
Total	122	100

Student Numeracy Achievement

Figure 11 shows the GloSS result of longitudinal students over the four years (2002–2005) of the Longitudinal Study. As a point of comparison, results from the schools that first participated in the NDP in 2005 are included. It is pleasing to note that the general trend in years 2, 4, and 6 is one of improving achievement in the longitudinal schools, and that, at those three year levels, the students in the 2005 Longitudinal Study performed better than the students in the NDP in 2005.

The achievement of students in the longitudinal schools on the Number Framework provides evidence for describing the levels of performance that can reasonably be expected from students in schools in the years following their participation in the NDP. As illustrated by Figure 11, approximately 80% of year 2 longitudinal students reach at least stage 4 (advanced counting), approximately 70% of year 4 students reach at least stage 5 (early additive), and approximately 70% of year 6 students reach at least stage 6 (advanced additive).

² Available from www.nzmaths.co.nz/Numeracy/Other%20material/GLOSS.aspx

³ It is not possible to calculate an exact response rate because the number of questionnaires distributed was based on an estimated class size of 23 students, ensuring that more than sufficient questionnaires were sent to each school.

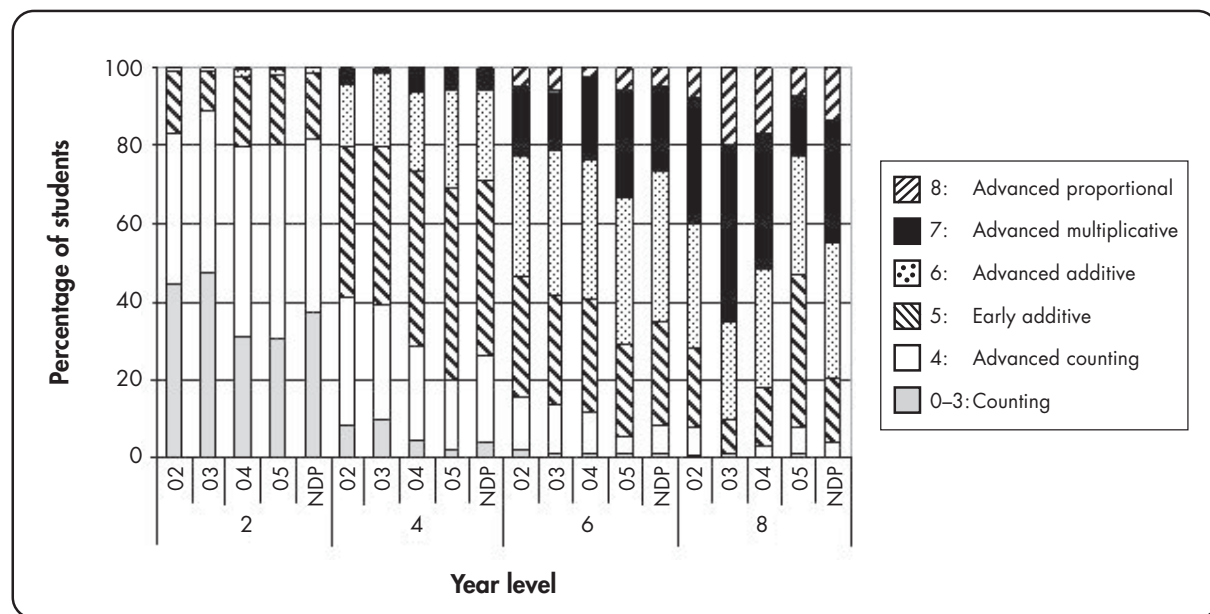


Figure 11. GloSS of Longitudinal Students 2002–2005 and NDP 2005 Students

The results of year 8 students have fluctuated over the years of the Longitudinal Study, with the highest achievement levels being in 2003 and the lowest in 2005. In 2005, the year 8 students in the 2005 NDP outperformed the year 8 longitudinal students. (Note that no strategy data was received for year 8 students in high-decile schools.)

Figure 12 shows the GloSS stage of year 8 students from the 2005 Longitudinal Study, categorised according to the year that their school first participated in the NDP. This figure illustrates the impact that the inclusion in the 2005 Longitudinal Study of one low-decile and one medium-decile intermediate school has had on the results for year 8 longitudinal students. While the sample of schools who first participated in 2003 performed significantly worse than the sample from the 2005 NDP, the sample of schools who first participated in 2000–2002 performed slightly better than the 2005 NDP students.

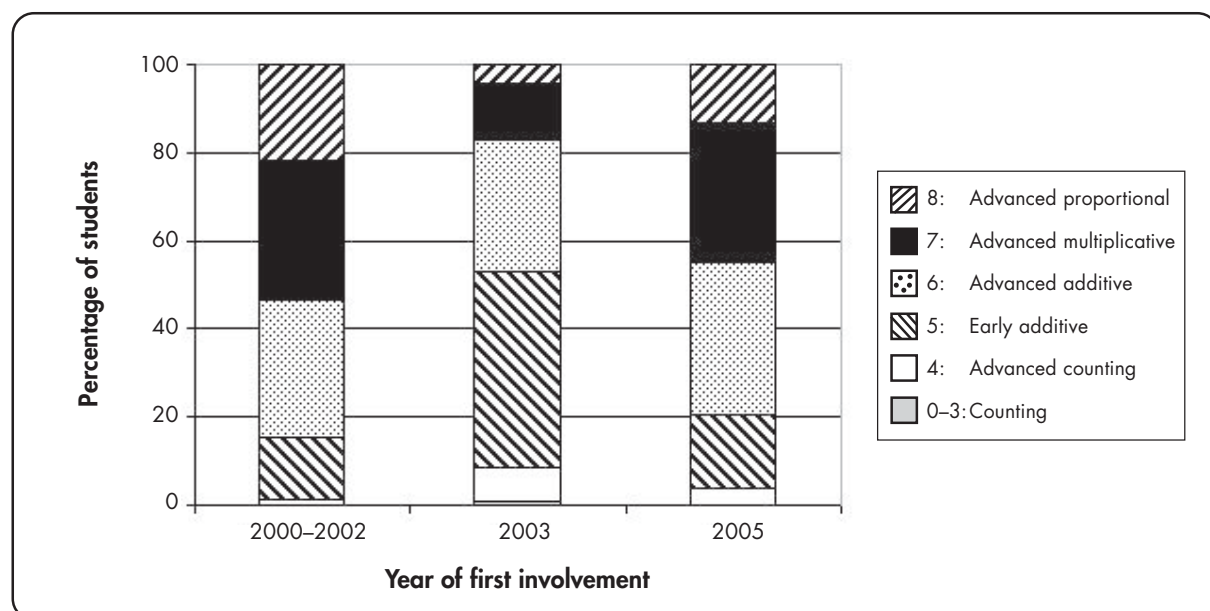


Figure 12. GloSS of year 8 students by first year of school's numeracy PD (no high-decile students in 2005)

Concluding Comment

The performance of students in the longitudinal schools on the mathematics tests is mixed. The year 4 longitudinal students performed on average 7% better than the New Zealand students in the source assessments. The year 5, 6, and 8 longitudinal students performed very similarly to the students in the source assessments, with year 6 longitudinal students rating 3% worse than those students. The year 5 and 8 longitudinal students' average overall test scores were not significantly higher than those of the New Zealand students in the source assessments, although they outperformed them on 8 and 12 of the 24 questions respectively.

The performance of the longitudinal students as measured against the Number Framework continues the trend of previous NDP longitudinal findings of improving achievement over time. The reasonably stable patterns of achievement for years 2, 4, and 6 students provide evidence for establishing benchmarks of numeracy achievement for the strategy domains.

References

- Garden, R., Ed. (1997). *Mathematics and science performance in the middle primary school: Results from New Zealand's participation in the Third International Mathematics and Science Study*. Wellington: Research and International Section, Ministry of Education.
- Higgins, J. (2001). *An evaluation of the years 4–6 Numeracy Exploratory Study*: Exploring issues in mathematics education. Wellington: Ministry of Education.
- Higgins, J. (2004). *An evaluation of the Advanced Numeracy Project 2003*: Exploring issues in mathematics education. Wellington: Ministry of Education.
- Higgins, J., Parsons, R., & Hyland, M. (2003). The Numeracy Development Project: Policy to practice. In J. Livingstone (Ed.), *New Zealand annual review of education* (pp. 157–174). Wellington: Victoria University of Wellington.
- Ministry of Education (2006). *Book 1: The Number Framework*. Wellington: Ministry of Education.
- Parsons, R. (2005, March). *Numeracy Development Project: Scope and scale*. Paper presented at the Numeracy Development Project Reference Group, Wellington.
- Thomas, G., Tagg, A., & Ward, J. (2003). *An evaluation of the Early Numeracy Project 2002*: Exploring issues in mathematics education. Wellington: Ministry of Education.
- Thomas, G. & Tagg, A. (2004). *An evaluation of the Early Numeracy Project 2003*: Exploring issues in mathematics education. Wellington: Ministry of Education.
- Thomas, G. & Tagg, A. (2005a). Evidence for expectations: Findings from the Numeracy Project longitudinal study. In *Findings from the New Zealand Numeracy Development Project 2004* (pp. 21–34). Wellington: Ministry of Education.
- Thomas, G. & Tagg, A. (2005b). The impact of the Numeracy Development Project on mathematics achievement. In *Findings from the New Zealand Numeracy Development Project 2004* (pp. 35–46). Wellington: Ministry of Education.
- Thomas, G. and Ward, J. (2001). *An evaluation of the Count Me In Too Pilot*: Exploring issues in mathematics education. Wellington: Ministry of Education.