A Decade of Reform in Mathematics Education: Results for 2009 and Earlier Years

Jennifer M Young-Loveridge
University of Waikato
<jenny.yl@waikato.ac.nz>

This paper reports on the impact of the Numeracy Development Projects (NDP) on the mathematics achievement of students whose teachers participated in the programme in the final year of the initial phase (2009). It also looks back at data gathered in years prior to this and presents an overview of the impact of the NDP on students' mathematics achievement, based on NDP data aggregated over several years (2003, 2005, and 2007) for almost one-quarter of a million students. Analysis of students' gains on the Number Framework for the additive and multiplicative domains shows that students made substantial progress in mathematics as a result of their teachers' NDP professional development. However, the absolute levels on the Framework attained by students were in many cases well short of the numeracy expectations for students at particular year levels stated in *The New Zealand Curriculum* (Ministry of Education, 2007) and in the *Mathematics Standards for Years 1–8* (Ministry of Education, 2009). The implications of these findings for further professional learning and development of teachers in mathematics are discussed.

Introduction

It is now almost a decade since New Zealand's Numeracy Development Projects (NDP) began as part of mathematics education reform. The focus has been on building the capacity of teachers to teach mathematics and, through this process, to raise the mathematics achievement of students across the board (Ministry of Education, 2001). The year 2009 marked the last year that teachers at primary and intermediate levels could participate in phase one of the NDP professional development programme. During this time, a revised curriculum document, *The New Zealand Curriculum,* embedded the Number Framework within the system of expected outcomes that apply to all students in New Zealand schools (Ministry of Education, 2007). This curriculum is now mandatory for all schools in New Zealand. Coinciding with the completion of the initial phase of the NDP was the release of a standards document, *Mathematics Standards for Years 1–8* (Ministry of Education, 2009). The standards document differs from the curriculum document in specifying the expected outcomes for students at the end of each year of schooling rather than for a two-year period (that is, curriculum level). A summary overview of these expectations in poster form shows the correspondence between the new mathematics standards and achievement outcomes in the curriculum, as well as the strategy stages on the Number Framework (Ministry of Education, 2010).

The introduction of standards is a contentious issue. There is some concern within the teaching profession about the impact of the standards on classroom teaching practices in mathematics. Pessimists worry about how the data will be used in the future and the extent to which the standards might inadvertently result in a shift in focus away from improving the quality of mathematics teaching to an overemphasis on assessment comparisons between students, teachers, and schools. Optimists see the potential benefits of making schools more accountable to their communities and the possibility that the standards could work to raise levels of student achievement.

Numeracy professional learning and development now focuses on in-depth sustainability of the NDP approach to teaching mathematics, with regular facilitator support for teachers in the schools involved. Facilitators are also supporting a large number of schools as they consider the implications of the mathematics standards document for their classroom programmes and assessment practices.

The evaluation research undertaken as part of the implementation of the initial phase of the NDP has enormous potential to assist in informing teachers and schools as they begin using the mathematics standards as a basis for reporting students' achievement to parents. A consistent finding has been that students at high-decile schools achieve higher stages on the Number Framework than students at middle-decile schools, who in turn reach higher stages than students at low-decile schools (Young-Loveridge, 2004, 2005, 2006, 2007, 2008, 2009). It is impossible to determine the extent to which these differences are a function of the characteristics of the students who attend schools at various decile levels, result from classroom mathematics programmes differing in quality, or are a consequence of lower expectations by teachers working in lower-decile schools. Ritchie's (2004) analysis of teacher career moves indicates that high-decile schools tend to attract teachers with more experience and qualifications than low-decile schools.

This paper reports on the impact of the NDP on the mathematics achievement of students whose teachers participated in the programme in the final year of the initial phase (2009). It also looks back at data gathered in years prior to this and presents an overview of the impact of the NDP on students' mathematics achievement, based on data aggregated over several years of the NDP (2003, 2005, and 2007). The data of almost one-quarter of a million students was aggregated to produce a summary picture of the impact of the NDP on students' mathematics achievement.

Method

Participants

For the first part of the study (the 2009 cohort), only students with complete data (that is, their teachers had submitted both initial and final assessment data to the Ministry of Education website) on the additive strategy domain and on the place value and basic facts knowledge domains were included in the analysis. (2009 was the last year in which teachers were given an opportunity to participate in the initial phase of the NDP professional learning and development programme.) The 2009 cohort consisted of 1645 year 1–9 students in 18 schools in the North Island. All of the students in years 1–6 of this cohort attended high-decile schools, while those at years 7–9 included a small proportion of students attending low-decile schools (between 7% and 11%), with the remainder at high-decile schools (see Table 1). The cohort included a disproportionately high number of boys (62%), particularly at years 7–9. The majority of students in years 1–6 were of European descent (75% to 87%). Students in years 7–9 included more Māori, Pasifika, Asian, and other ethnicities than for students in years 1–6.

Table 1
Percentages[1] of Students in Each Year Level as a Function of Gender, Ethnicity, and School-decile Band (2009)

Composition	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Overall
Number of students	*85*	*95*	*100*	*95*	*99*	*112*	*336*	*343*	*380*	*1645*
2009										
Gender										
Boys	42	52	57	54	52	58	69	68	64	*62*
Girls	58	48	43	46	49	42	31	32	36	*38*
Ethnicity										
European	87	80	81	77	79	75	51	45	40	*57*
Māori	6	7	7	7	6	7	10	10	34	*14*
Pasifika	1		2	1	1	2	9	10	6	*6*
Asian	1	1	1	5	8	3	17	19	12	*11*
Other	5	12	9	10	6	13	13	16	9	*11*
School-decile band										
Low decile (1–3)							7	8	35	*11*
High decile (8–10)	100	100	100	100	100	100	93	92	65	*89*

The second part of this study included participants with complete data (that is, initial and final) whose teachers participated in their first year of professional development for the NDP in either 2003, 2005, or 2007 (see Table 2). This mega-cohort consisted of close to one-quarter of a million children (n = 240 331). Because of the large numbers of participants in each of these three cohorts (n = 137 121, 54 935, 48 275, respectively), it was possible to look at consistencies from one year to another and to examine patterns for small sub-groups such as Pasifika students attending high-decile schools. (Note that only the percentages for European, Māori, and Pasifika students are included in Table 2 under ethnicity because they were the three largest subgroups within the cohort.)

Table 2
Percentages[2] of Students Who Participated in NDP in 2003, 2005, and 2007

Composition	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Total
Number of students	*17 189*	*18 265*	*19 508*	*18 774*	*18 050*	*18 775*	*12 997*	*11 340*	*2223*	*137 121*
2003										
School-decile band										
Low decile	33	35	36	34	34	34	41	42	56	*36*
Middle decile	37	36	35	38	38	39	43	42	27	*38*
High decile	31	29	29	28	28	28	15	16	17	26
Gender										
Boys	51	51	51	51	51	51	52	51	49	*51*
Girls	49	49	49	49	49	49	49	49	51	*49*
Ethnicity										
European	59	58	58	59	60	59	53	54	49	*58*
Māori	22	21	22	24	23	24	28	28	22	*24*
Pasifika	10	11	10	8	9	8	11	11	22	*10*

1 Percentages are rounded to the nearest whole number.
2 Percentages are rounded to the nearest whole number.

Table 2
Percentages[2] of Students Who Participated in NDP in 2003, 2005, and 2007 – continued

Composition	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Total
Number of students	*4 730*	*5 044*	*5 708*	*6 940*	*8 306*	*8 653*	*6 086*	*5 432*	*4 036*	*54 935*
2005										
School-decile band										
Low decile	15	16	17	21	22	21	13	16	11	*18*
Middle decile	46	45	43	42	39	39	56	55	52	*45*
High decile	39	39	40	37	39	40	32	29	37	*37*
Gender										
Boys	52	52	53	52	52	50	52	53	56	*52*
Girls	49	49	47	48	49	50	48	47	44	*48*
Ethnicity										
European	63	63	63	63	62	63	66	63	66	*64*
Māori	20	19	19	19	19	18	20	21	20	*19*
Pasifika	6	6	7	8	9	8	7	8	5	*7*
Number of students	*2525*	*4166*	*3969*	*4217*	*4360*	*4371*	*7352*	*10104*	*7211*	*48 275*
2007										
School-decile band										
Low decile	26	27	28	28	26	28	19	20	17	*23*
Middle decile	32	32	34	32	31	31	42	44	49	*38*
High decile	42	41	38	40	44	42	39	37	35	*39*
Gender										
Boys	52	52	51	53	52	52	53	52	47	*51*
Girls	48	48	49	48	48	48	48	49	53	*49*
Ethnicity										
European	54	56	56	55	57	58	59	62	60	*58*
Māori	22	23	24	23	22	23	22	21	20	*22*
Pasifika	13	12	11	12	10	10	9	8	8	*10*
Total 2003, 2005, 2007										*240 331*

Results

The results are reported in two sections: the first reports on the results for the cohort of students whose teachers participated in their first year of professional development for the NDP in 2009; the second section reports on analysis of aggregated data from 2003, 2005, and 2007, years in which large cohorts of students and their teachers were involved in the NDP.

2009 Cohort

It is clear from Table 1 (see earlier) that the 2009 cohort was quite atypical compared with larger cohorts in previous years (see, for example, Table 2). Compared with more representative cohorts, the 2009 cohort consisted mostly of students attending high-decile schools, apart from the year 9 students, one-third of whom came from low-decile schools.

[2] Percentages are rounded to the nearest whole number.

Figure 1 presents the proportion of students at each year level who were at various stages on the Number Framework for the additive domain at the end of the year (for percentages, see Appendix A). Figures 2 and 3 present the proportion of students at each year level who, at the end of the year, were at various stages on the Number Framework for the multiplicative and proportional domains, respectively. Consistent with the Ministry of Education's numeracy expectations, approximately three-quarters of the students were at stage 4 on the additive strategy domain by the end of year 2, at stage 5 by the end of year 4, and at stage 6 by the end of year 6 (see Table 3).

The numbers of students at each year level in 2009 were fewer than 100, and all of the students in years 1–6 came from high-decile schools, so it was possible only to analyse the findings as a function of year level.

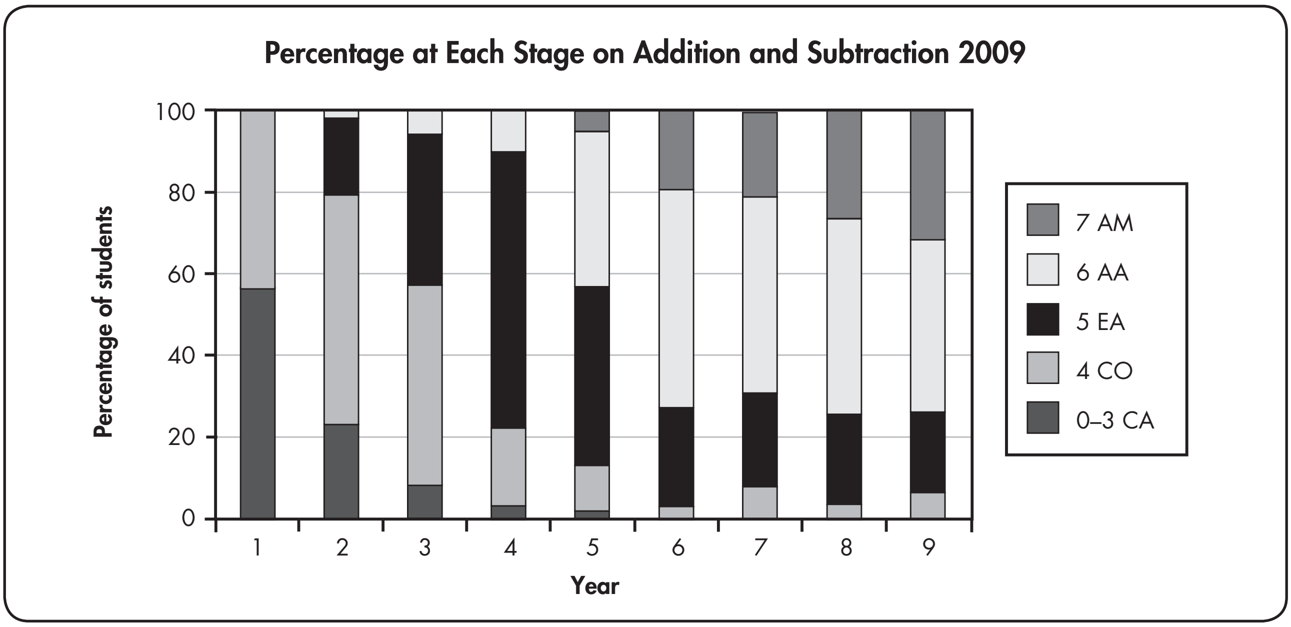

Figure 1. Percentages of students at each stage on the *additive* domain as a function of year level (2009)

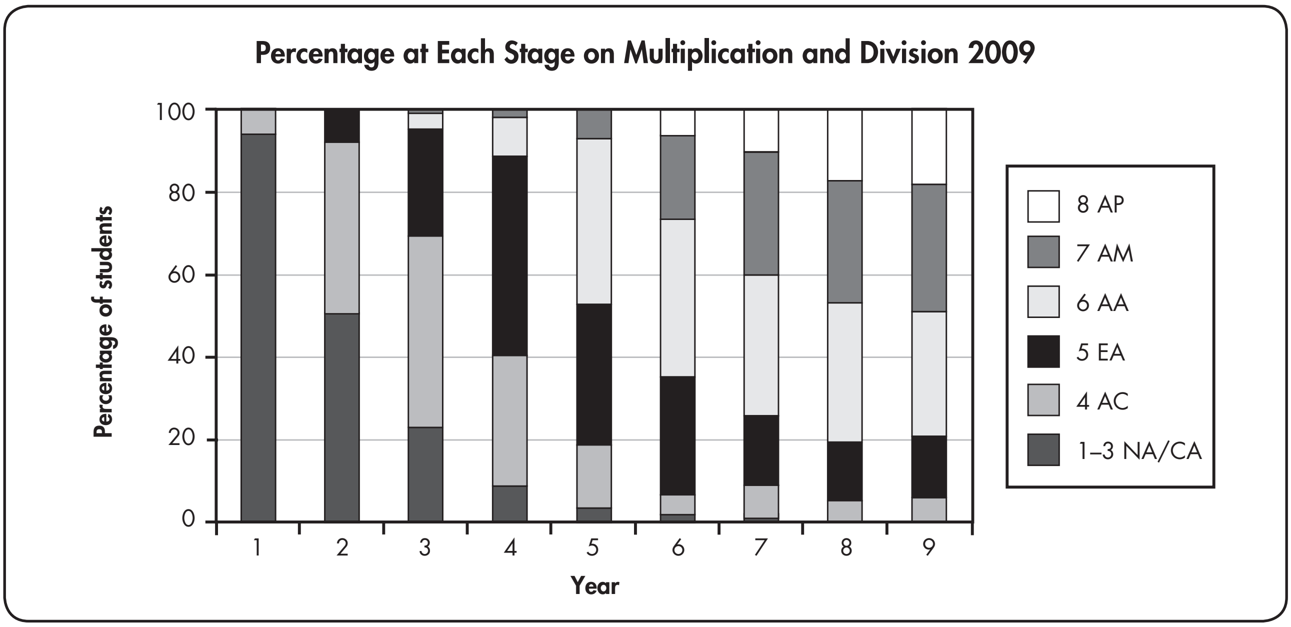

Figure 2. Percentages of students at each stage on the *multiplicative* domain as a function of year level (2009)

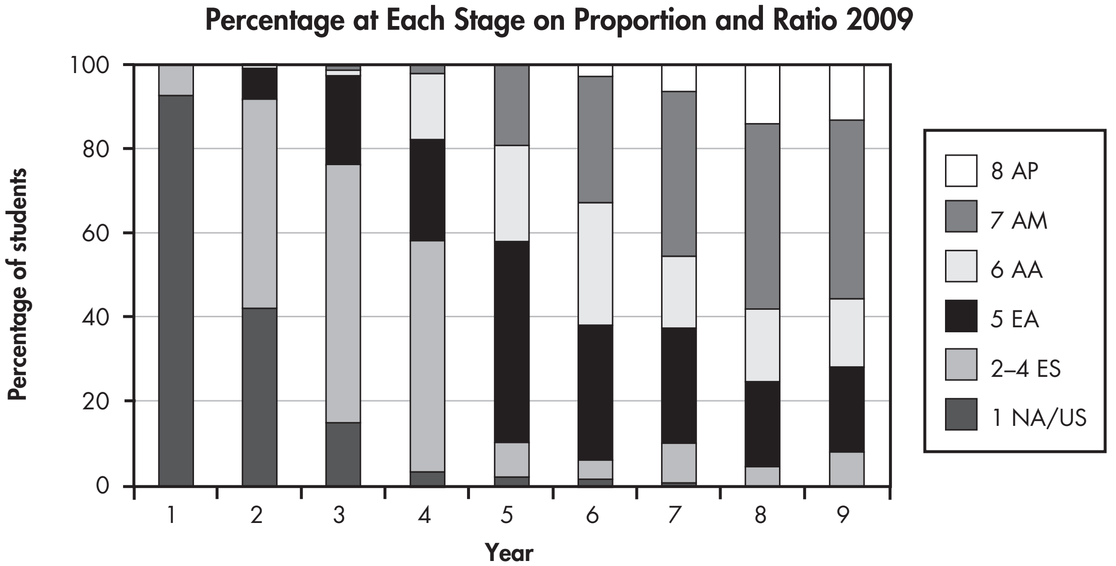

Figure 3. Percentages of students at each stage on the *proportional* domain as a function of year level (2009)

Table 3

Percentages[3] of Students Who Were at or above Particular Stages on the Number Framework as a Function of Year Level and Domain (2009)

Stage and domain, 2009 final	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9
STRATEGIES									
Number of students	*85*	*95*	*100*	*95*	*99*	*112*	*336*	*343*	*380*
Additive domain									
Stage 4+	44	77	92	97	98	100	99	100	100
Stage 5+	0	21	43	78	87	97	92	97	93
Stage 6+	0	2	6	11	44	73	69	75	74
Stage 7+	0	0	0	0	5	20	21	27	31
Multiplicative domain									
Stage 4+	6	50	77	92	97	98	99	100	100
Stage 5+	0	8	31	60	82	94	92	95	94
Stage 6+	0	0	5	12	48	65	75	81	80
Stage 7+	0	0	1	2	7	27	40	47	49
Proportional domain									
Stage 5+	0	7	24	42	90	97	90	96	92
Stage 6+	0	1	3	18	42	64	63	76	72
Stage 7+	0	0	1	2	19	35	46	58	56

Aggregation of 2003, 2005, and 2007 Cohorts

Data from students with both initial and final data in 2003, 2005, or 2007 were aggregated to create a mega-cohort. Almost one-quarter of a million students between years 1 and 9 were included in this analysis. Tables 2 and 4 shows that the composition of this mega-cohort was very close to the composition of the New Zealand population, with approximately 30% of the students attending low-decile (1–3) schools, 30% attending high-decile (8–10) schools, and 40% attending middle-decile (4–7) schools. If anything, there were slightly more students at high-decile schools and slightly fewer at middle-decile schools for students in years 1–6. At the intermediate level (years 7 and 8), that pattern

[3] Boxes indicate the expected stage at particular year levels in *The New Zealand Curriculum*

was reversed. Only at year 9 were there disproportionately fewer students from low-decile schools, balanced by more students from middle-decile schools. Hence, this mega-cohort could be considered to be representative of the total population.

Table 4

Percentages[4] of Students at Schools within Each Decile Band as a Function of Year Level (aggregated for 2003, 2005, 2007)

School Decile	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9
Low (1–3)	29	30	31	30	30	30	29	29	22
Middle (4–7)	38	37	37	38	37	37	46	45	46
High (8–10)	33	33	33	32	34	33	26	26	32
Number of students (246 165)	*24 931*	*27 947*	*29 720*	*30 576*	*31 475*	*32 526*	*27 286*	*27 998*	*13 706*

The Additive Domain

Overall

Figure 4 shows the percentages of students at each stage on the additive domain as a function of year level (see Appendix B for percentages). Table 5 shows the percentages of students who reached or exceeded stages 4 through to 7 on the strategy domains. It is clear from Table 5 that the actual percentages of a large representation sample of students whose teachers had completed one year of the NDP fell some way short of the Ministry's numeracy expectations (Ministry of Education, n.d.) and the mathematics standards (Ministry of Education, 2009, 2010). For example, just over half (57%) of the students were able to count on (stage 4) by the end of year 2. Almost two-thirds (63%) were able to use a limited range of additive part–whole strategies (stage 5) by the end of year 4. Just over one-third (38%) could use an extensive range of additive part–whole strategies (stage 6) by the end of year 6. A little over one-third (36%) could use an extensive range of multiplicative part–whole strategies (stage 7) by the end of year 8. It was interesting to note that a greater proportion of year 6 students were beginning to use a limited range of multiplicative part–whole strategies (stage 6) compared with those who could use an extensive range of additive part–whole strategies (stage 6; 53% compared with 38%).

Table 5

Percentages[5] of Students Who Were At or Above Particular Stages on the Number Framework as a Function of Year Level and Domain (2003, 2005, 2007)

Year Level	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9
Additive domain									
Number of students	*24 931*	*27 947*	*29 720*	*30 576*	*31 475*	*32 526*	*27 286*	*27 998*	*13 706*
Stage 4+	19	57	84	94	97	98	98	98	99
Stage 5+	2	14	41	63	75	84	86	90	92
Stage 6+	0	1	5	14	25	38	46	58	61
Multiplicative domain									
Number of students	*23 106*	*27 376*	*29 955*	*31 021*	*31 946*	*33 051*	*28 149*	*28 445*	*14 708*
Stage 4+	6	34	68	86	92	95	96	96	98
Stage 5+	1	7	26	50	67	79	83	88	91
Stage 6+	0	1	7	21	37	53	60	70	74
Stage 7+	0	0	1	3	9	19	25	36	41

4 Percentages are rounded to the nearest whole number.

5 Boxes indicate the expected stage at particular year levels in *The New Zealand Curriculum*.

Table 5 – continued

Percentages[5] of Students Who Were At or Above Particular Stages on the Number Framework as a Function of Year Level and Domain (2003, 2005, 2007)

Year Level	**Y1**	**Y2**	**Y3**	**Y4**	**Y5**	**Y6**	**Y7**	**Y8**	**Y9**
Proportional domain									
Number of students	*23 116*	*27 391*	*29 910*	*30 994*	*31 931*	*33 025*	*28 122*	*28 417*	*14 698*
Stage 5+	0	5	21	42	60	72	77	84	88
Stage 6+	0	1	4	14	29	44	51	62	66
Stage 7+	0	0	1	3	9	19	25	36	47

Figure 5 (and Table 6) shows the average initial stage on the Number Framework for the additive domain and the gain from initial to final assessment for each year level (see Appendix C for averages for subgroups). On average, students in year 1 began the project halfway between stage 1 and stage 2. By year 9, the average initial stage was a little above stage 5. It is clear from Appendix C that there is a great deal of consistency from one cohort to another in the first couple of years of school. However, in 2005 and 2007 the average stage on the Framework from about year 3 was higher than in 2003. It is likely that this increase can be explained by the addition of another stage (stage 7) on the Framework for the additive domain, making it possible for students who could add and subtract with fractions and decimals as well as whole numbers to be awarded a higher stage. This increase became more pronounced at higher year levels as more students reached stage 7 on the Framework for the additive domain (see Appendix C). It is important to note that more than half of the mega-cohort consists of students in the 2003 cohort, and at that time, the Number Framework only went to stage 6 on the additive domain and to stage 7 on the multiplicative domain. This would have effectively lowered the mean stage slightly for students at higher year levels for whom a ceiling effect was probably operating in 2003.

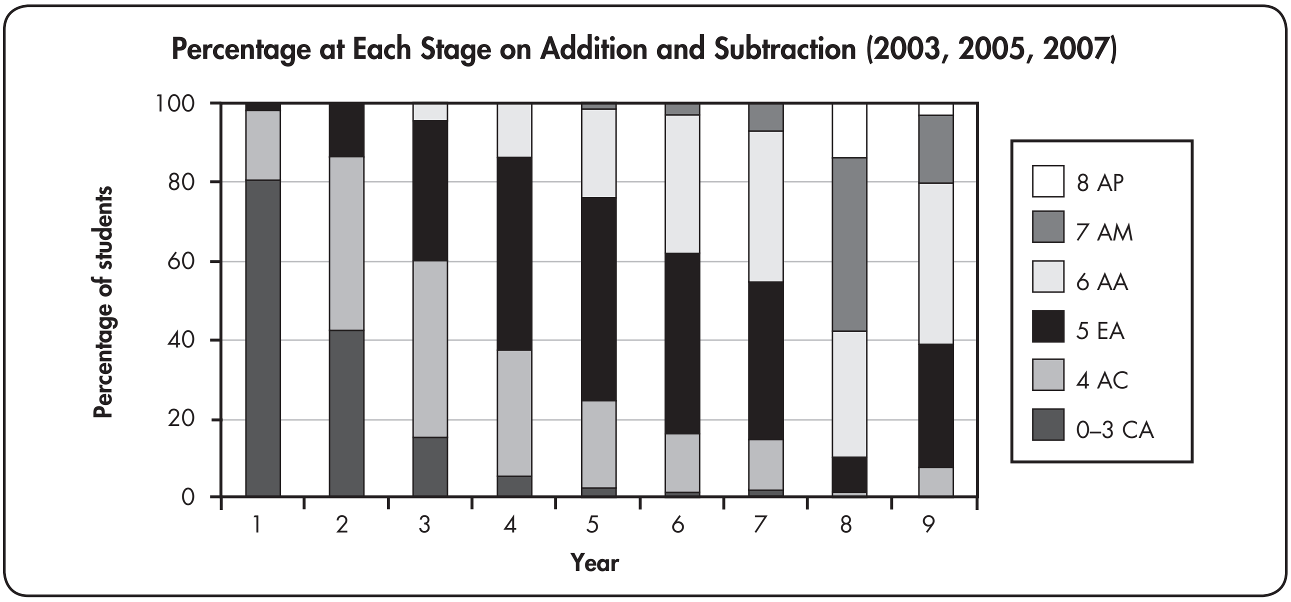

Figure 4. Percentages of students at each stage on the *additive* domain (2003, 2005, 2007)

5 Boxes indicate the expected stage at particular year levels in *The New Zealand Curriculum.*

Table 6
Average Stage[6] on the Additive Domain of the Number Framework as a Function of Year Level (averaged for 2003, 2005, 2007)

Additive Domain	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9
Initial stage	1.45	2.44	3.46	4.16	4.54	4.81	4.97	5.16	5.18
Gain	1.06	1.01	0.76	0.53	0.46	0.47	0.42	0.44	0.40
Final stage	2.51	3.45	4.22	4.69	4.99	5.27	5.39	5.60	5.59

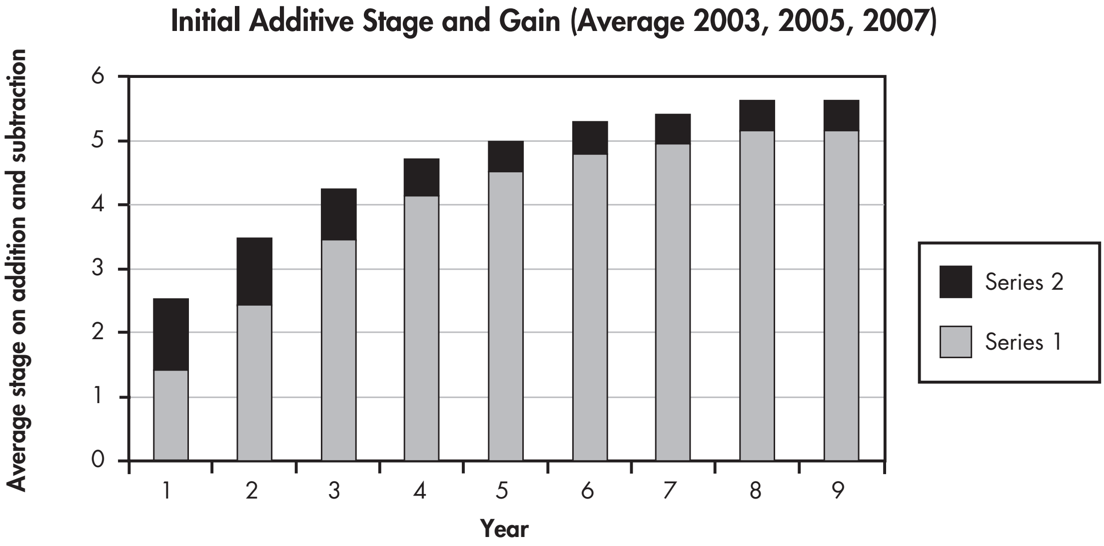

Figure 5. Average stage on the *additive* domain as a function of year level (averaged for 2003, 2005, 2007; Series 1: Initial stage; Series 2: Gain to final stage)

Gender

Appendix C shows the difference in average initial stage for boys compared with girls. Initially, the difference was very small and favoured girls in the first year of school. However, from the second year at school, boys gained a slightly higher average stage, and this average difference gradually increased from 0.01 to 0.19 by year 9. This could be explained by the Math-Fact Retrieval hypothesis, which shows that boys are able to access number facts more quickly and effectively than girls (Geary, 1999; Royer, Tronsky, Jackson, & Marchant, 1999a; Royer, et al., 1999b; Wigfield & Byrnes, 1999).

School decile

Students at high-decile schools reached higher initial stages on the Number Framework for addition and subtraction than those at middle- and low-decile schools (see Appendix C). By year 9, students at low-decile schools had not quite reached stage 5, on average, whereas those at high-decile schools were almost halfway between stage 5 and stage 6. The average advantage for high-decile over middle-decile schools was almost one-quarter of a stage (0.23) and just over half a stage (0.52) for high-decile over low-decile schools.

Ethnicity

Students of European backgrounds reached higher initial stages on the Number Framework for addition and subtraction than those of Māori and Pasifika ancestry (see Appendix C). By year 9, European students were well into stage 5 (mean = 5.29), while those of Māori and Pasifika descent

6 Average stage is rounded to two decimal places.

were towards the end of stage 4 (means = 4.94 and 4.83, respectively). The average advantage for European students over Māori was just over one-third of a stage (0.37), while that for European over Pasifika students was just under half a stage (0.48). The direction of the differences as a function of ethnicity are consistent with those found in other studies of mathematics achievement, such as the Third International Mathematics and Science Study (TIMSS) (Garden, 1996, 1997, 1998) and the National Education Monitoring Project (NEMP) (Crooks & Flockton, 2002a, 2002b; Flockton & Crooks, 1998; Flockton, Crooks, Smith, & Smith, 2006). However, a major difference between this NDP research and other comparisons is that the effect sizes (standardised differences between groups) are smaller for the NDP than for the international comparisons based on group paper-and-pencil tests. Typically, the effect size for the European-Māori comparison on international studies is about three-quarters of a standard deviation and that for the European-Pasifika comparison is approximately one standard deviation (see Young-Loveridge, 2006). Analysis of ethnicity differences on initial Number Framework stage show them to be much smaller (approximately 0.17 for European–Māori comparisons and 0.29 for European–Pasifika comparison; see Young-Loveridge, 2006). The advantage of the individual orally administered NDP assessment over the paper-and-pencil group assessment is that the questions are read out by the teacher, so the mathematics is immediately accessible to the student. This contrasts with written mathematics assessment tasks, which may not be accessible to poor readers. Māori and Pasifika students have, on average, lower levels of achievement than European students, so they are likely to be more disadvantaged by the literacy demands of a written test. It has been suggested that the reason for the substantial reduction in ethnicity differences for NDP assessment is that the individual diagnostic interview by the student's own teacher provides a much more reliable (albeit more time-consuming) assessment of their mathematics than a paper-and-pencil test to an entire class (Young-Loveridge, 2006).

Ethnicity and school decile

The impact of ethnicity was examined separately for students from high- and low-decile schools. Students from all three major groups (European, Māori, Pasifika) reached higher stages on the Number Framework for the additive domain when they attended a high-decile school rather than a low-decile school. For all three groups, the advantage was about one-third of a stage, with the biggest difference found for Māori students (high-decile advantage = 0.38, on average). European students at high-decile schools reached the highest stage (mean = 5.42 at year 9), whereas Pasifika students at low- and high-decile schools were the lowest (mean = 4.72 and 4.81, respectively). Research on teacher movements in relation to the decile ranking of schools indicates that there is a tendency for more experienced teachers to be in schools with higher decile rankings (see Ritchie, 2004). Low-decile schools tend to have teachers with less teaching experience, on average. This, combined with the greater transience levels for students (and teachers), and lower expectations for students' achievement, may compound the disadvantage experienced by low-decile students.

Gain in Stage on the Additive Domain

Overall

The magnitude of the gain in stage on the addition and subtraction domain varied as a function of year level (see Appendix D). In the first two years of school (years 1 and 2), students gained one whole stage on the Number Framework, on average. In year 3, the gain was about three-quarters of a stage. Between year 4 and year 6, the gain was approximately half a stage. Between years 7 and 9, the gain was between one-third and half a stage. A possible reason for the change in size of gain across the primary years is that students can progress through the early (lower) stages on the Framework far more quickly and easily than they can through the later (upper) stages. What this indicates is that the steps on the Framework are not of equal size, and this is borne out by a statistical analysis (of logits; see Ward & Thomas, this volume).

Gender

Although boys began at a slightly higher initial stage on the Number Framework than girls, their gain was virtually identical, with boys gaining fractionally more than girls initially, and then from year 6, girls gained very slightly more than boys (see Appendix D). The differences ranged from 0.05 to –0.07 but averaged out to zero across all year levels.

School decile

There were very tiny differences in gain between students at schools differing in terms of decile and no consistent pattern across year levels (see Appendix D). On average, there was absolutely no difference in the size of the *gain* as a function of school decile even though there were clear differences in terms of initial stages.

Ethnicity

Students from all ethnic backgrounds made similar gains on the Number Framework for the additive domain, and any tiny differences cancelled each other out when averaged across year level (see Appendix D). One exception to this appeared to be year 7 and 8 Pasifika students in 2003, who made zero gains after their teachers' year of involvement in the NDP professional learning and development programme. However, in 2005 and 2007, the gains for Pasifika were approximately one-third of a stage on the Framework.

Ethnicity and school decile

The analysis of gains as a function of ethnicity and school decile revealed that in 2003, Pasifika students attending low-decile schools actually went backwards, losing, on average, one-tenth of a stage on the Number Framework for addition and subtraction. This exacerbated the high-decile advantage, which was between half and almost three-quarters of a stage. However, this aberration vanished in 2005 and 2007, suggesting that it may have been the result of something unusual happening in the particular intermediate schools that participated in the NDP in 2003 rather than a particular vulnerability for Pasifika students at the intermediate (year 7–8) level. It is also possible that later Schooling Improvement projects for low-decile schools had an advantageous impact on students' numeracy levels (for example, the Manurewa Enhancement Initiative; see Young-Loveridge, 2005).

Effect Sizes

An analysis of the gains was made in terms of effect size for each of the years 2003, 2005, and 2007, and then the average across these years was calculated (see Appendix E). The advantage of using effect size is that it standardises the magnitude of the difference between students' initial Number Framework stage and the final stage at the end of the year of NDP professional development. According to several writers (for example, Fan, 2001; Schagen & Hodgen, 2009), effect sizes of 0.2, 0.5, and 0.8 are considered small, medium, and large, respectively. However, Hattie (2009) argues that for educational outcomes, effect sizes of 0.40 and 0.60 should be considered medium and large, respectively. He argues further that quite small effect sizes may be important in some circumstances. According to Hattie, an effect size of 1.0 means that students receiving a particular intervention programme would outperform 84% of students not receiving that programme. Hattie suggests setting the bar at 0.40 in terms of a desirable effect size and using that as the benchmark to judge effects in education. He asserts that "any innovation, any teaching programme, and all teachers should be aiming to demonstrate that the effects on students' achievement should exceed $d = 0.40$" (p. 249; d means effect size). This corresponds to students receiving intervention outperforming 66% of students not receiving intervention.

The average effect sizes calculated in the present study for the gain in stage from the initial assessment at the beginning of the school year to the final assessment at the end of the school year were substantial:

approximately one standard deviation in the first two years of school and more than half a standard deviation in the following years of primary schooling (see Appendix E). In the intermediate years, effect sizes were closer to half a standard deviation and in year 9, were between one-third and half a deviation. All the effect sizes for the primary years (1–6) were well in excess of 0.40. At the intermediate level (years 7–8), the only effect sizes not to exceed 0.40 were those for low-decile students at year 7 (d = 0.39) and for Pasifika students at both years 7 and 8 (d = 0.36 and 0.37, respectively). When ethnicity was considered separately for each decile level, it was only the low-decile Pasifika students in years 7 and 8 (d = 0.34 and 0.32, respectively) for whom the effect size did not exceed the 0.40 level. At the year 9 level, boys and Māori students at high-decile schools had effect sizes less than 0.40 (d = 0.37 and 0.36, respectively).

Initial stage on the Multiplicative Domain

The analysis of data for the multiplicative domain was complicated by the fact that not all students were given the opportunity to solve multiplication and division problems and some were only given this chance at the end of the year but not the beginning. Hence a decision was made to include data only from students who had been assessed on the multiplicative domain at the beginning and end of the year of professional development. Very few year 1 students were given the opportunity to solve multiplication and division problems, so the analysis was confined to students in years 2–9 (see Appendix C).

Overall

On average, students in year 2 began on the multiplicative domain halfway between stages 3 and 4, while those at year 9 almost reached stage 6 (mean = 5.85) (see Table 6, Figure 6, and Appendix C).

Table 7
Average Stage on the Multiplicative Domain of the Number Framework as a Function of Year Level (averaged for 2003, 2005, and 2007)

Multiplicative Domain	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9
Initial stage	3.41	3.79	4.35	4.80	5.21	5.44	5.74	5.85
Gain	0.63	0.75	0.71	0.71	0.68	0.63	0.62	0.52
Final stage	4.04	4.54	5.06	5.51	5.89	6.07	6.36	6.36

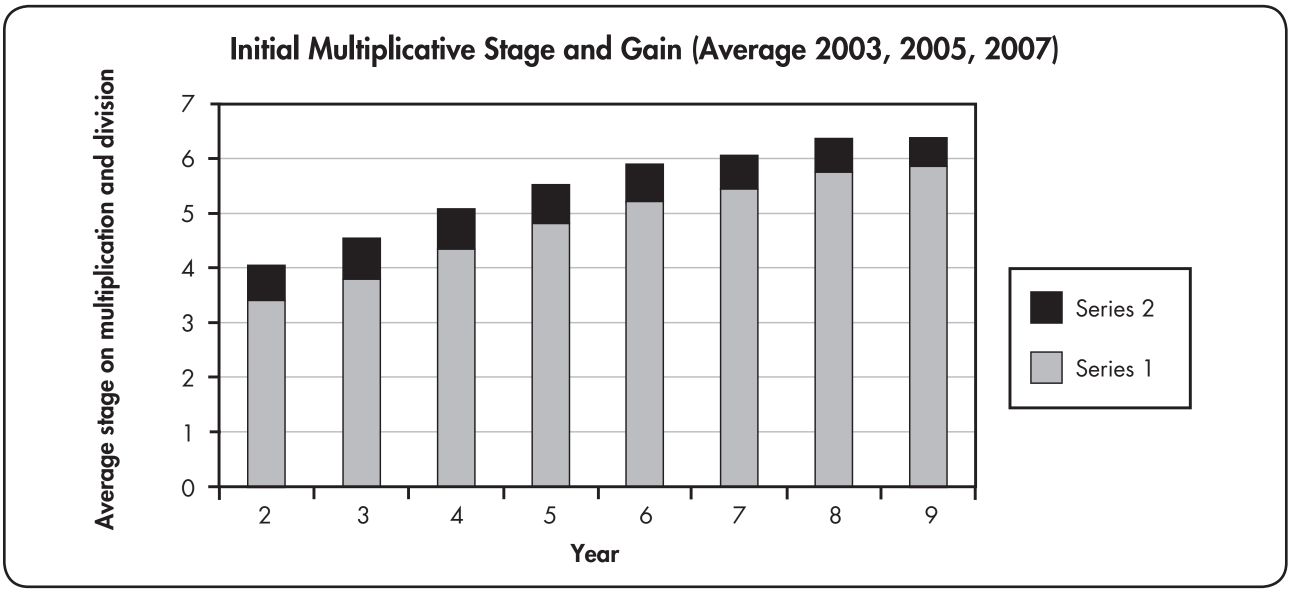

Figure 6. Average stage on the *multiplicative* domain as a function of year level (averaged for 2003, 2005, and 2007; Series 1: Initial stage; Series 2: Gain to final stage)

Gender

Boys did consistently better than girls on the multiplicative domain at all year levels and for all three cohorts (see Appendix C). On average, the boys were one-fifth of a stage higher on the Number Framework for multiplication and division than were girls.

School decile

As for the additive domain, students from high-decile schools reached the highest Number Framework stages on the multiplicative domain, with an advantage over those from middle-decile schools of almost one-quarter of a stage (0.24), while the high-decile advantage over low-decile was more than half a stage (0.54), on average (see Appendix C).

Ethnicity

European students reached stage 6, on average (mean = 6.05), whereas Māori and Pasifika students reached partway between stage 5 and 6 (mean = 5.52 & 5.33, respectively) (see Appendix C).

Gain on the Multiplicative Domain

Overall

On average, students annually gained between half and three-quarters of a stage on the Number Framework for multiplication and division (see Appendix D). The highest gains were for students in years 3–6 and the lowest at year 9.

Gender

Boys and girls made very similar gains on the multiplicative domain, with virtually no difference when differences were averaged across the three cohorts and over all year levels (mean = 0.01) (see Appendix D).

Decile

There was a small difference, with students from high-decile schools making slightly higher gains than those from middle-decile schools (mean difference = 0.04) and those from low-decile schools (mean difference = 0.09) (see Appendix D).

Ethnicity

European students made slightly higher gains than those of Māori (mean difference = 0.04) and Pasifika descent (mean difference = 0.07). However, these differences were very small and indicate that all ethnic groups benefited equally from participation in the NDP (see Appendix D).

Effect Sizes

An analysis of the gains on the multiplicative domain was made in terms of effect size (see Appendix E). Effect sizes on the multiplicative domain were not quite as large as for the additive domain. However, all were at least three-quarters of a standard deviation in years 2 and 3, while those for students in years 4–8 were at least half a standard deviation. Those for students in year 9 were between one-third and half a standard deviation. None of the effect sizes calculated was below Hattie's (2009) "hinge-point" of 0.40. However, it is important to remember that this analysis only includes students who were assessed on the multiplicative domain at the beginning and end of the year of NDP professional development. Students who were not initially assessed on the multiplicative domain (because the only strategy domain included in Form A is the additive domain) but by the end of the year had made sufficient progress to be assessed on Form B or C (which include both the multiplicative and proportional domains) were excluded from the analysis because of missing data.

Discussion

One of the notable findings of this analysis was the remarkable consistency across the three cohorts in initial stage on the Number Framework and the gain in stage on the Framework for the two domains examined. When the large sample size (almost one-quarter of a million students) is taken into account, it is clear that the findings have considerable weight. Although there were some differences between various subgroups (based on ethnicity and school decile level) in terms of initial stage on the Number Framework, it was encouraging to see that all subgroups made very similar gains in stage. When the effect sizes were examined in relation to Hattie's (2009) 0.40 hinge point, it was clear that the NDP professional development programme has produced substantial gains in terms of progress on the Framework. Teachers of students in the early years can be justifiably proud of the effect sizes of approximately one standard deviation that they produced on the additive domain. According to Hattie (2009), this corresponds to acceleration of between two and three years of education. He suggests that "teachers average an effect of $d = 0.20$ to $d = 0.40$ per year on student achievement" (p. 16). In his own work, the yearly effect in reading, mathematics, and writing from years 4 to 13 is $d = 0.35$. He argues that "teachers should be seeking greater than d = 0.40 for their achievement gains to be considered above average and greater than $d = 0.60$ to be considered excellent" (p. 17). It is clear that most of the gains from NDP professional development were above average, and many of these gains were excellent. Teachers have a lot to be proud of.

The next big question is, are these gains enough when considered in relation to desirable levels of achievement? Whether the data is examined in terms of percentages of students reaching or exceeding particular levels on the Number Framework or in terms of the average stage attained on the Framework, the picture is consistent. Students after the first year of the NDP professional development programme are not yet reaching the levels thought to be needed for them to reach an acceptable level of achievement by year 12 (the fourth year of secondary schooling) (see Ministry of Education, 2007, 2009, n.d.).

For example, data from year 6 students has shown that by the end of the first year in which their teachers participated in NDP professional development, fewer than half of the students at this level were at stage 6 or higher by the end of the year (between 41% and 49% over the period 2005–2008) (see Young-Loveridge, 2005, 2006, 2007, 2008, 2009). This is well short of the majority of students (75% to 80%) expected to be at stage 6 or higher by the end of year 6 (Ministry of Education,2007, 2009, n.d.). It should be noted that the proportion of year 5 students at stage 6 was considerably lower (between 22% and 34% over the period 2005–2008). Likewise, the percentages of year 8 students reaching stage 7 by the end of their teachers' first year of NDP were considerably less than half (between 28% and 42% over the period 2005–2008). An analysis of changes in students' end-of-year achievement levels for NDP schools that submitted data in two consecutive years (2007 and 2008) shows little, if any, gain from the first to the second year of the NDP, except for students in year 2 (see Young-Loveridge, 2009).

These findings may raise important questions for some people about the levels at which the expectations are set. However, there is clear evidence that the standards need to be set high, even though it may be some time before the majority of students meet the standards set for their particular year levels. The decision about what standard to expect at a particular year level is based on several important sources of evidence. One is research evidence showing that students need to be multiplicative thinkers if they are to engage meaningfully with instruction in algebra at secondary school (Brown & Quinn, 2006; Lamon, 2007; Wu, 2002). Hence, the evidence clearly shows that students need to be at stage 7 or higher on the Number Framework by the end of year 8.

Another reason for having the expectations at the current level comes from *The New Zealand Curriculum* (Ministry of Education, 2007). Unlike previous curriculum documents, the latest curriculum document is based on research evidence about progressions in students' thinking rather than on a collection of isolated topics to be covered by teachers. This curriculum also reflects recent reforms in mathematics education that have led to prioritising conceptual understanding over procedural knowledge and skills. The curriculum overall is designed to support the development of learners in being "confident and creative, connected, and actively involved" (Ministry of Education, 2007, p. 4) – in other words, innovative problem solvers who can confidently think through the solutions to new problems using the tools they have acquired. It is no coincidence that both *The New Zealand Curriculum* and the *Mathematics Standards* have the expectation that students at the end of year 8 need to be multiplicative thinkers. Although the national standards were written later, they were designed to align with the curriculum document (see Ministry of Education, 2007, 2009).

Another important source of evidence for expectations is the curriculum documents of other education systems, which reveal that the expectations for students at the end of year 8 in New Zealand to be advanced multiplicative thinkers (stage 7 on the Number Framework) are set somewhat later than in other jurisdictions. For example, in the United States, students at grade 5 (year 6) are expected to be "developing an understanding of and fluency with addition and subtraction of fractions and decimals" and in the following year (grade 6 [year 7]), "developing an understanding of and fluency with multiplication and division of fractions and decimals" as well as "connecting ratio and rate to multiplication and division" (NCTM, n.d.). These expectations are at least stage 7 on the New Zealand Number Framework – an expectation for the end of year 8. At grade 7 (year 8), the United States students are "developing an understanding of and applying proportionality", which matches stage 8 advanced proportional on the New Zealand Number Framework, an expectation for the end of year 9. In New South Wales, the expectation at grade 6 (year 7) is that a student operates with fractions, decimals, percentages, ratios and rates (Board of Studies NSW, 2006). It is not clear whether or not students in other jurisdictions actually meet the expectations described in their curriculum documents. However, the implications of such documents are that New Zealand students at the end of year 8 should ideally be working at stage 8 advanced proportional part-whole thinking. Only about 12% of year 8 students were advanced proportional thinkers by the end of the school year (see 2003, 2005, 2007 in Appendix B). Hence, it is very clear that New Zealand's Mathematics Standards need to be at least as high as they currently are.

Data from cohorts of students whose teachers were participating in their first year of NDP training provide an insight into the initial impact of teachers' professional development on students' achievement after just one year of participation. The Longitudinal Study, which continued to gather data from schools after their initial involvement in the NDP, was designed to provide a longer-term view. From 2002 to 2006, the NDP Longitudinal Study reported on the performance of students in a group of schools who continued to collect numeracy data in the years following their initial implementation of the NDP (Thomas & Tagg, 2005, 2006, 2007). Since 2007, the Longitudinal Study has tracked the progress of a cohort of students from these schools for whom numeracy data is available for every year of their primary schooling (Thomas & Tagg, 2008, 2009). The data from the Longitudinal Study in 2006 was used to inform the setting of expectations for students at particular year levels. In 2006, the Longitudinal Study sample contained disproportionately more students from high- and middle-decile schools and fewer from low-decile schools, raising some questions regarding the generalisability of these results to students from low-decile schools.

Summarising the evidence presented in this paper thus far, it is clear that teachers have made considerable progress in their understanding of how numeracy develops and what they can do to support their students' numeracy learning. However, it is also clear that there is still a long way to go. Teachers are up against a systemic problem that dates back to their own schooling and to that of their teachers (and probably also to that of their teachers before them). Researchers have noted that teachers tend to teach the way they themselves were taught when they were at school (see Grootenboer, 2001, 2002). Teachers need to realise that low levels of mathematics achievement are a consequence of problems with the entire education system rather than solely the responsibility of a few individual teachers. The problems go back generations and are the legacy of decades of mathematics teaching being procedural rather than conceptual, a pattern that has become deeply ingrained over time (see Skemp, 2006). Moreover, the beliefs, attitudes, values, and feelings that teachers have about mathematics that were formed when they were at school have probably had a substantial impact on their teaching. Although there are international mathematics education reforms calling for a focus on conceptual understanding instead of procedural skills, there is little, if any, writing about what an enormous challenge that represents for the teacher education process, both pre-service and in-service. It would be naive to think that one or two years of professional development could miraculously change teachers' attitudes, feelings, beliefs, and values as well as their understanding of mathematics. It seems likely that many teachers may need considerably more professional development if they are to acquire the deep and connected understanding that they need to move a greater majority of their students closer to the expected levels of achievement. It is important that this process of ongoing professional learning (including having higher expectations for students) continues throughout the primary years of schooling. Becoming an advanced multiplicative thinker (stage 7) by the end of year 8 is only possible if all the teachers in the years leading up to year 8 have had appropriate expectations of their students. This needs to be seen as an issue for all primary teachers to address, not just those working at the year 7 and 8 level. Given the amount of time needed to develop an understanding of the complexities of additive and multiplicative thinking, a considerable lead-up time is necessary in order to provide a strong foundation on which those ideas can be built. We must not underestimate the challenge of bringing about deep and lasting change in teachers' understanding of mathematics learning.

If we accept that further efforts are required to support primary teachers to the point where they are all confident and skilled teachers of mathematics, the question then is what does research evidence indicate about possible directions that might help to achieve that goal. One area that might be fruitful is the place of counting on the Number Framework. It has been suggested by Sophian (2007) that in many education systems there is an overemphasis on counting. Sophian argues that the focus on discrete quantity (i.e., "how many?") draws attention away from continuous quantity (i.e., "how much?") and could help to explain the difficulties that students worldwide experience with fractional quantity, which requires co-ordination of both "how many?" (e.g., the numerator) and "how much?" (e.g., the denominator). The NDP data analysis shows that students in the early years of school are using counting strategies for longer than is desirable, according to the expectations conveyed in the curriculum and the national standards (Ministry of Education, 2007, 2009). For example, the highest counting stage on the Number Framework (stage 4, advanced counting) is, on average, attained by students at the end of year 3 (see Table 6). Likewise, students do not reach stage 5 (early additive part-whole thinking), on average, until the end of year 5. The strong emphasis on counting in the NDP may send students the wrong message, resulting in some students being reluctant to move away from counting at any stage because of its proven reliability for solving problems with small numbers. It is of great concern that even at years 8 and 9, approximately 5% of students are still using counting to solve addition and subtraction problems (see Appendix A).

Such a prolonged period using counting strategies may occur because teachers do not begin early enough helping students to develop an understanding of part-whole relationships among numbers and of the use of partitioning and recombining strategies as a means for determining the answers to simple problems. Teachers could begin introducing ideas about the composition of numbers as wholes made of different parts right from the early days of school. For example, there is no reason why five-year-olds could not investigate different ways to "make" six and other small numbers. They could explore the way that six can be made of five and one, or four and two, or three and three. They might also be introduced to the idea of six as double three or as three pairs of two, the beginnings of ideas important for work with the multiplicative domain. They need to become familiar with many different representations of six, such as the stylised pattern of six dots on a dice arranged in a rectangular shape, a row of five dots plus an extra dot as represented on a tens frame, or one hand of five fingers plus the thumb of the other hand. They should be estimating and making predictions about how many would result from adding another one on or taking one away from that collection of six objects.

Another reason for prolonged counting may be that subitising (the instant naming of quantities, such as the stylised patterns presented on dot dice and dominoes) is not encouraged as an alternative way (apart from counting) of determining quantity for small collections of objects. There is considerable research to suggest that subitising for small sets precedes counting and can provide a valuable introduction to concepts of numerical quantity. Some researchers have suggested that subitising might suit certain students better than counting (e.g., Hannula, Rasanen, & Lehtinen, 2007; Willis, 2000). A distinction has been made in the literature between perceptual and conceptual subitising (Clements, 1999). The latter uses the initial patterns that are learned as a basis for building more complex patterns. Although quantities represented on ten-frames may come to be subitised eventually, using other representations such as double two (four) or double three (six) displayed on dot dice and dominoes might be easier to learn than a linear display of four dots/counters or a line of five plus an additional dot (see Clements, 1999). Support for part-whole strategy instruction could also come from building knowledge of basic facts for small numbers (for example, "one and one", "two and one", "two and two", "three and one") right from the beginning of schooling, instead of (or as well as) counting. (Note: the smallest basic fact assessed in NumPA is "2 + 3".)

The existence of so many micro-stages at the lower end of the Number Framework may suggest to some teachers that the steps between these stages are of the same magnitude as those higher up the Framework. Hence they do not appreciate the importance of moving through those early stages at a reasonably brisk pace. Another possibility is that the alignment between strategy and knowledge at corresponding stages may suggest to some teachers that students need to be able to meet all of the stage 4 knowledge requirements before being given instruction in stage 5 (early part-whole thinking) strategy ideas. Teachers holding that view would require students to show that they could read 3-digit numerals, order numbers in the sequence up to 100 both forwards and backwards, and know all combinations of single-digit addends for sums and differences up to 20 before they could be considered ready for instruction in stage 5 strategy (using part-whole strategies to join or separate two small collections). Yet it is possible to work with concepts of parts and wholes with far less knowledge than is required at stage 4. Could that knowledge be developed alongside instruction focused on early additive part-whole thinking?

There is now considerable evidence to support the idea that in order to succeed in mathematics, students need to develop an appreciation of the pattern and structure of number (Mulligan & Mitchelmore, 1997, 2009; Mulligan, Prescott, & Mitchelmore, 2004). Mulligan and colleagues found that students who were low achievers in mathematics did not appear to notice the regularities and structure in the mathematics that was presented to them. NDP support materials are consistent with the idea of helping all students to appreciate pattern and structure within the number system. However, the suggestions above may help to strengthen that process.

Evidence from research on the classroom practices of teachers in countries that scored high on international comparisons of mathematics achievement has shown that they spent considerable time focusing on making connections among mathematical ideas (Stigler & Hiebert, 2004). This contrasts markedly with the practices of teachers in the United States, who tended to turn most problems into procedural exercises. Students in the US scored significantly lower in mathematics than these countries (e.g., Czech Republic, Hong Kong, Japan, Netherlands). Components of the NDP such as the Number Framework, the diagnostic interview, and the books providing support materials are consistent with the idea of making connections among mathematical concepts, but such practices would require a good understanding of mathematics as well as a high level of familiarity with the complexities of the framework and the support materials. However, it is not clear just how much time is needed for teachers to develop a deep understanding of the NDP so that they can then help their students to make connections among mathematical ideas.

An important consideration with respect to students' progress in mathematics is the length of the NDP professional learning and development programme. The numeracy strategy policy determined that teachers would get support over two years in the NDP approach to teaching mathematics. However, research indicates that it takes considerable time to effect major change in approaches to teaching mathematics (Anthony & Walshaw, 2005; Lamon, 2007; Reid & Zack, 2010). In retrospect, it may be that teachers needed to receive support by facilitators for considerably longer in order to effect a more permanent change in their classroom mathematics practice. A research project with teachers in classrooms designed to explore the impact of the NDP approach on the teaching of addition and subtraction (see Young-Loveridge & Mills, this volume) and multiplication and division (Young-Loveridge & Mills, 2009a, 2009b) showed that many teachers stuck very closely to the printed NDP resources (the "pink" books), which describe examples of possible lessons that they could use (Ministry of Education, 2008c, 2008d). This could reflect the low levels of confidence that many teachers still have about their mathematics teaching, despite their involvement in NDP. Decisions about the ideal duration of NDP professional learning and development programmes were influenced by the need to balance the depth of impact with the provision of opportunity for as many teachers as possible. Over the decade in which the initiative was introduced to teachers, teachers in virtually all primary and intermediate schools were given the opportunity to participate. It would have been difficult to justify denying half of these teachers the opportunity for professional learning in order to give the other half twice as much support.

The issue of teachers' content (subject matter) knowledge (SMK) and pedagogical content knowledge (PCK) in mathematics has been the focus of considerable writing (Ball et al., 2005, 2008; Ward & Thomas, 2006, 2007, 2008, 2009). Evidence has clearly established the need for teachers working at the primary/elementary level to have greater understanding of mathematics and ways to support their students' mathematical development. International reforms in mathematics education have moved the emphasis away from drilling students in rules and procedures (instrumental learning) towards a deeper and more connected conceptual (relational) learning and understanding (Lambdin & Walcott, 2008; Skemp, 2006). Connected to this issue of teacher knowledge is the topic of classroom norms for mathematics instruction and the importance of getting students not just to explain their strategies, but also to critique them and justify them to each other. This is particularly important for students from low-decile schools who may not be familiar with argumentation processes. Hunter (2010) has shown how teachers in a low-decile school with many Maori and Pasifika students, who thought initially that they had high expectations for their students, later realised through their work with a teacher educator that their students were in fact capable of engaging in far more sophisticated mathematical argumentation and went on to achieve substantially higher on the Number Framework as a result.

Finally, it is important to think about the implications of the NDP data analysis for the implementation of the mathematics standards. The results presented here indicate that, in the initial implementation of the mathematics standards, many students in the middle and senior years may not reach the expected standards for their year level. Teachers need to understand that, at this point in time, the mathematics standards are *aspirational* for many students rather than *typical* and quite large numbers of students may be initially "below expectations". This should be recognised as a temporary situation that is inevitable while teachers are continuing to upskill themselves in mathematics content knowledge and pedagogical content knowledge. As part of that process, teachers of junior classes have the challenge of moving their students beyond counting more quickly than in the past by putting a greater emphasis on partitioning small numbers, by including subitising as another important knowledge domain in their classroom programmes, and by encouraging students to become familiar with the "basic facts" for small number combinations.

No one has ever claimed that mathematics education reform was going to be easy. Bringing about substantial change in the way mathematics is taught throughout the primary school (and beyond) is a huge challenge. This researcher believes that we have underestimated just how difficult that process is. As writers have recently pointed out, "elementary mathematics is anything but elementary" (Bahr & de Garcia, 2010, title page).

References

Anthony, G., & Hunter, R. (2005). A window into mathematics classrooms: Traditional to reform. *New Zealand Journal of Educational Studies, 40,* 25–43.

Bahr, D. L., & de Garcia, L. A. (2010). *Elementary mathematics is anything but elementary: Content and methods from a developmental perspective.* Belmont, CA: Wadsworth Cengage Learning.

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, Fall, 14–22, 43–46.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Berry, A., Loughran, J., & van Driel, J. H. (2008). Revisiting the roots of pedagogical content knowledge. *International Journal of Science Education, 30,* 1271–1279.

Board of Studies NSW (2006). *Mathematics K–6 Syllabus 2002.* Retrieved on 14 May 2010 from: http://k6.boardofstudies.nsw.edu.au/go/mathematics

Brown, G., & Quinn, R. J. (2006). Algebra students' difficulty with fractions: An error analysis. *Australian Mathematics Teacher, 62*(4), 28–40.

Clements, D. H. (1999). Subitizing: What is it? Why teach it? *Teaching Children Mathematics, 5*(7), 400–405.

Crooks, T., &. Flockton, L. (2002a). *Mathematics assessment results 2001:* National Education Monitoring report 23. Wellington: Ministry of Education.

Crooks, T., & Flockton, L. (2002b). *Assessment results for Māori students 2001: Information skills, social studies, mathematics:* National Education Monitoring report 24. Wellington: Ministry of Education.

Fan, X. (2001). Statistical significance and effect size in educational research: Two sides of a coin. *Journal of Educational Research, 94*(5), 275–282.

Flockton, L., & Crooks, T. (1998). *Mathematics assessment results 1997:* National Education Monitoring report 9. Wellington: Ministry of Education.

Flockton, L., Crooks, T., Smith, J., & Smith, L. (2006). *Mathematics assessment results 2005:* National Education Monitoring report 37. Wellington: Ministry of Education.

Garden, R. A. (1996). *Mathematics performance of New Zealand form 2 and form 3 students: National results from New Zealand's participation in the Third International Mathematics and Science Study.* Wellington: Research and International Section, Ministry of Education.

Garden, R. A. (1997). *Mathematics and science performance in 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.

Garden, R. A. (1998). *Mathematics and science literacy in the final year of schooling: Results from New Zealand's participation in the Third International Mathematics and Science Study.* Wellington: Research Division, Ministry of Education.

Geary, D. C. (1999). Sex differences in mathematical abilities: Commentary on the math-fact retrieval hypothesis. *Contemporary Educational Psychology, 24,* 267–274.

Goya, S. (2006). Math education: Teaching for understanding: The critical need for skilled math teachers, *Phi Delta Kappan, 87*(5), 370–372.

Grootenboer, P. (2001). How students remember their mathematics teachers. *Australian Mathematics Teacher, 57*(4), 14–16.

Grootenboer, P. (2002). Affective development in mathematics: A case study of two preservice primary school teachers. In B. Barton, K. C. Irwin, M. Pfannkuch, & M. O. J. Thomas (Eds), *Mathematics Education in the South Pacific* (Proceedings of the 25th annual conference of the Mathematics Education Research Group of Australasia, pp. 318–325). Sydney: MERGA.

Hannula, M. M., Rasanen, P., & Lehtinen, E. (2007). Development of counting skills: Role of spontaneous focusing on numerosity and subitizing-based enumeration. *Mathematical Thinking & Learning, 9*(1), 51–57.

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Hunter, R. (2010). Changing roles and identities in the construction of a community of mathematical inquiry. *Journal of Mathematics Teacher Education, 13,* 397–409.

Lambdin, D.V., & Walcott, C. (2007). Changes through the years: Connections between psychological learning theories and the school mathematics curriculum. In W. G. Martin, M. E. Struchens, & P. C. Elliott (Eds), *The learning of mathematics: Sixty-ninth yearbook* (pp. 3–25). Reston, VA: National Council of Teachers of Mathematics.

Lamon, S. (2007). Rational numbers and proportional reasoning: Towards a theoretical framework for research. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629–667). Charlotte, NC: Information Age Publishing.

Ministry of Education (1997). *School entry assessment.* Wellington: Ministry of Education.

Ministry of Education (2001). *Curriculum Update 45: The numeracy story.* Wellington: Ministry of Education.

Ministry of Education (2007). *The New Zealand Curriculum.* Wellington: Ministry of Education.

Ministry of Education (2008a). *Book 1: The Number Framework: Revised edition 2007.* Wellington: Ministry of Education.

Ministry of Education (2008b). *Book 2: The diagnostic interview.* Wellington: Ministry of Education.

Ministry of Education (2008c). *Book 5: Teaching addition, subtraction, and place value.* Wellington: Ministry of Education.

Ministry of Education (2008d). *Book 6: Teaching multiplication and division.* Wellington: Ministry of Education.

Ministry of Education (2009). *Mathematics standards for years 1–8.* Wellington: Ministry of Education.

Ministry of Education (2010). *Mathematics standards for years 1–8* (Poster: Item number 11475). Wellington: Ministry of Education.

Ministry of Education (n.d.). *Curriculum expectations.* Retrieved on 1 March 2010 from http://new.nzmaths.co.nz/using-expectations.

Moch, P. L. (2004). Demonstrating knowledge of mathematics: Necessary but *not* sufficient. *Curriculum and Teaching Dialogue*, *6*(2), 125–130.

Mulligan, J., Prescott, A., & Mitchelmore, M. (2004). Children's development of structure in early mathematics. In M. J. Hoines & A. B. Fugelstad (Eds), *Proceedings of the 28th PME international conference* (Vol. 3, pp. 393–400). Bergen: PME.

Mulligan, J., & Mitchelmore, M. (1997). Young children's intuitive models of multiplication and division. *Journal for Research in Mathematics Education, 28*(3), 309–330.

Mulligan, J., & Mitchelmore, M. (2009). Awareness of pattern and structure in early mathematical development. *Mathematics Education Research Journal*, *21*(2), 33–49.

National Council of Teachers of Mathematics (n.d.). *Curriculum focal points.* Retrieved on 14 May 2010 from: www.nctm.org/standards/default.aspx?id=280

Reid, D. & Zack, V. (2010). Observing the process of mathematics teacher change – part 1. *Journal of Mathematics Teacher Education*, *13*, 371–374.

Ritchie, G. (2004). Teacher movement between schools in New Zealand: To high SES schools that have, more shall be given. *Journal of Educational Policy*, *19*(1), 57–59.

Royer, J. M., Tronsky, L. N., Jackson, S. J., & Marchant, H. (1999a). Math-fact retrieval as the cognitive mechanism underlying gender differences in math test performance. *Contemporary Educational Psychology, 24,* 181–266.

Royer, J. M., Tronsky, L. N., Jackson, S. J., & Marchant, H. (1999b). Reply to the commentaries on the math-fact retrieval hypothesis. *Contemporary Educational Psychology*, *24,* 286–300.

Schagen, I. & Hodgen, E. (2009). *How much difference does it make? Notes on understanding, using, and calculating effect sizes for schools.* Retrieved on 14 May 2010 from http://www.educationcounts.govt.nz/publications/schooling/36097/36098

Skemp, R. (2006). Relational understanding and instrumental understanding. *Mathematics Teaching in the Middle School*, *12,* 88–95.

Sophian, C. (2007). *The origins of mathematical knowledge in childhood.* New York: Erlbaum.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership*, *61*(5), 12–17.

Thomas, G., & Tagg, A. (2005). Evidence for expectations: Findings from the Numeracy Project Longitudinal Study. In *Findings from the New Zealand Numeracy Development Projects 2005* (pp. 21–34). Wellington: Learning Media.

Thomas, G., & Tagg, A. (2006). Numeracy Development Project Longitudinal Study: Patterns of achievement. In *Findings from the New Zealand Numeracy Development Projects 2005* (pp. 22–33). Wellington: Learning Media.

Thomas, G., & Tagg, A. (2007). Do they continue to improve? Tracking the progress of a cohort of longitudinal students. In *Findings from the New Zealand Numeracy Development Projects 2006* (pp. 8–15). Wellington: Learning Media.

Thomas, G., & Tagg, A. (2008). What do the 2002 school entrants know now? In *Findings from the New Zealand Numeracy Development Projects 2007* (pp. 5–15). Wellington: Learning Media.

Thomas, G., & Tagg, A. (2009). The Numeracy Development Projects' Longitudinal Study: How did the students perform in year 7? In *Findings from the New Zealand Numeracy Development Projects 2008* (pp. 3–11). Wellington: Learning Media.

Ward, J., & Thomas, G. (2007). What do teachers know about fractions? In *Findings from the New Zealand Numeracy Development Projects 2006* (pp. 128–138). Wellington: Learning Media.

Ward, J., & Thomas, G. (2008). Does teacher knowledge make a difference? In *Findings from the New Zealand Numeracy Development Projects 2007* (pp. 50–60). Wellington: Learning Media.

Ward, J., & Thomas, G. (2009). Linking teacher knowledge and student outcomes. In *Findings from the New Zealand Numeracy Development Projects 2008* (pp. 77–85). Wellington: Learning Media.

Wigfield, A., & Byrnes, J. P. (1999). Does math-fact retrieval explain sex differences in mathematical test performance? A commentary. *Contemporary Educational Psychology, 24,* 275–285.

Willis, S. (2000). Strengthening numeracy: Reducing risk. Proceedings of the ACER Research Conference: *Improving numeracy learning: What does the research tell us?* (pp. 31–33). Brisbane.

Wu, H. (2002). Chapter 2: Fractions (draft). Retrieved on 14 May 2010 from www.math.berkeley.edu/~wu/EM12a.pdf

Young-Loveridge, J. (1987). Learning mathematics. *British Journal of Developmental Psychology, 5,* 155–167.

Young-Loveridge, J. M. (1991). *The development of children's number concepts from ages five to nine: early mathematics learning project: Phase II. Volume I: Report of findings.* Hamilton: Education Department, University of Waikato.

Young-Loveridge, J. M. (1993). *The effects of early mathematics intervention: The EMI-5s study.* Hamilton: Department of Education Studies, University of Waikato.

Young-Loveridge, J., Carr, M. & Peters, S. (1995). *Enhancing the Mathematics of four-year-olds; Volume I: Report of findings.* Hamilton: University of Waikato.

Young-Loveridge, J. (2004a). *Patterns of performance and progress on the Numeracy Projects 2001–2003: Further analysis of the Numeracy Project data.* Wellington: Ministry of Education.

Young-Loveridge, J. (2004b). Effects on early numeracy of two programs using number books and games. *Early Childhood Research Quarterly, 19,* 82–98.

Young-Loveridge, J. (2005). Patterns of performance and progress: Analysis of 2004 data. In *Findings from the New Zealand Numeracy Development Project 2004* (pp. 5–20, 115–129). Wellington: Ministry of Education.

Young-Loveridge, J. (2006). Patterns of performance and progress on the Numeracy Development Project: Looking back from 2005. In *Findings from the New Zealand Numeracy Development Projects 2005* (pp. 6–21, 137–155). Wellington: Learning Media.

Young-Loveridge, J. (2007). Patterns of performance and progress on the Numeracy Development Projects: Findings from 2006 for years 5–9 students. In *Findings from the New Zealand Numeracy Development Projects 2006* (pp. 16–32, 154–177). Wellington: Learning Media.

Young-Loveridge, J. (2008). Analysis of 2007 data from the Numeracy Development Projects: What does the picture show? In *Findings from the New Zealand Numeracy Development Projects 2007* (pp. 18–28, 191–211). Wellington: Learning Media.

Young-Loveridge, J. (2009). Patterns of performance and progress of NDP students in 2008. In *Findings from the New Zealand Numeracy Development Projects 2008* (pp. 12–26, 169–184). Wellington: Learning Media.

Young-Loveridge, J. & Mills, J. (2009a). Teaching multi-digit multiplication using array-based materials. In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides* (Proceedings of the annual conference of the Mathematics Research Group of Australasia, 5–9 July, Wellington pp. 635–642).

Young-Loveridge, J. & Mills, J. (2009b). Supporting multiplicative thinking: Multi-digit multiplication using array-based materials. In *Findings from the New Zealand Numeracy Development Projects 2008* (pp. 48–64). Wellington: Learning Media.