SUSTAINING CHANGED TEACHER PRACTICE IN NUMERACY

Nicola Knight
The University of Auckland
New Zealand
nickyk@xtra.co.nz

Presented at the ICSEI Conference Auckland, New Zealand 6-9 January 2008

This paper provides an examination of the concept of sustainability, situating the analysis within the context of recent developments in mathematics education in New Zealand. The phrase "sustaining change" is widely used without a clear definition, shared meaning or an understanding of factors associated with long term generative mathematics improvement. The research (in progress) sought to identify factors associated with sustaining improved student outcomes and improved teacher pedagogical knowledge. The study explored three primary schools and teachers of students aged 7 to 11 and their senior managers over an eighteen month period after the completion of their formal numeracy training. As there is a need for current research in mathematics education to examine small samples in depth, a descriptive case study offered an appropriate methodological stance for this research. Constructivist and socio-cultural theories of learning provided a theoretical lens through which teacher practice was examined. Initial results show significant differences in student outcomes among teachers which may be attributable to the individual capacity of the teacher to manage the mathematical content and pedagogical practices and their attitude to being a learner. The research will help school leaders and educational agencies gain a better understanding of the explicit and implicit standards used to determine the sustainability of a major numeracy reform and determining which ones carry the most weight and why.

INTRODUCTION

Aidan Davison, in his 2001 book 'Technology and the Contested Meanings of Sustainability' stated that one of the key modern day arguments in environmental and ecological discourses was the question of what sustains us. He stated "the nub of my argument is that the ideal of sustainability - the focus of so much thinking about our…predicament – beckons us toward the now unfamiliar yet still resonant question of what sustains us" (p. ix). He suggested that questions concerning sustainability are essentially normative and therefore "any answers they prompt essentially contestable" (p. ix). The same debates including a question of definition, measurement, contributing factors to sustainability at differing systemic levels within and between organizations can be applied to a curriculum reform context. This paper, representing research in progress, seeks to determine the factors that are associated with the sustainability of changed teacher practice in the context of numeracy reforms in New Zealand. It seeks to answer the questions "what does it mean to sustain something and how it is sustained?"

Over the last two decades, central government agencies in a number of countries including England, Holland, the United States, Australia and New Zealand have committed considerable financial resources to developing new ways of teaching mathematics, specifically number. At an international level, mathematics reforms have called for substantial changes to classroom instructional practice and assessment (Amit & Hillman, 1999). Fraivillig, Murphy and Fuson (1999, p. 1) indicated that a major focus of the mathematics reform movement was to "change from traditional classrooms that focus on students acquiring proficiency in reproducing existing solution methods to classrooms that support instructional goals of helping students construct personally meaningful conceptions of mathematical topics." Under the current reform programme, procedural mathematical knowledge, which focuses on

using correct computational procedures, has been usurped by principled knowledge which involves students using key mathematical ideas and concepts to construct procedures for solving mathematical problems (Spillane, 1999).

When involved in New Zealand's Numeracy Development Project (NDP), teachers undertake intensive professional development over a one to two year period. The NDP aims to increase teachers' ability to improve students' competence in addition, subtraction, multiplication, division, fractions, decimals and proportion problems focusing on the development of part-whole relationships among numbers. In addition to developing students' numerical knowledge and strategies, teachers are guided and supported "in establishing more effective teaching and learning models informed by sound mathematics content knowledge" (Higgins, 2003, p. 7). Following the conclusion of the professional development, schools are left to fend for themselves to a lesser or greater extent.

Sustaining changed teacher practice is of particular importance to the developers and educators associated with the NDP in New Zealand. With the NDP now moving beyond the initial conception and implementation phases, the focus has shifted to developing ways of ensuring that the intense work done to support the initial adoption and implementation is not in vain. The current focus on sustainability signals a desire for the funding agency and schools to have feedback on the effectiveness of their substantial investments and highlights the complexity of the layers of curriculum reform from their initial implementation to the ongoing development required to keep it going in schools (Ministry of Education, 2004). The new challenge for teachers is to embed and sustain the changed practices, otherwise reversion back to previous pedagogical practices and processes can be easy and painless. Hand in hand with discussions about how to sustain practice, educators also should engage in a careful and critical look at what is being sustained, what criteria are being used of make judgments, whose criteria are being used and how schools change reforms as they are implemented.

THE CONTESTED NOTION OF SUSTAINABILITY: A QUESTION OF DEFINITION

Currently, there are multiple contested meanings of the term "sustaining" changed teacher practice. In their work on large scale reform projects in the United States, Century and Levy (2002, p. x) defined sustainability as "the ability of a programme to maintain its core beliefs and values and use them to guide programme adaptations to changes and pressures over time." Sustainability refers to the ability of teachers to not only maintain those practices mobilised through their participation in NDP but to adapt to local conditions and improve in response to internal and external changes (Century & Levy, 2002). A programme is maintained if its basic elements are well established and part of standard practice, while sustainability stresses the importance of adapting to local conditions. The maintenance phase is seen as an essential precursor to programme sustainability.

Bobis et al. (2005, p.35) described sustainability as the "long term success" and "consolidation and maintenance of gains made within the project since its initial implementation" (p. 48). In addition, the authors argued that sustainability involved "providing on-going access to support from outside the school, for example, facilitators' visits and access to on line resources" (p. 48). Neither long term success nor maintenance of gains was explained in any detail. In defining these common everyday words, hidden complexities emerge that should be considered prior to beginning any effort to continue a reform project. For example, "consolidation and maintenance of gains" is worth examining further. Are the authors referring to student gains? If so, are they academic, social, performance or cultural gains or a combination of some or all four? Are the academic gains referring to student scores on a standardised test? If so, what are the expected gains and outcomes? Arguably, what counts as sustainability should be empirically based. Young-Loveridge (2005) has attempted to make "long term success" more explicit. She stated that the NDP was having a continued, positive impact on student achievement with a statistically significant average effect size in multiplication and division (0.40) and in proportion and ratios (0.43). The article suggested that all students on the NDP have made improvements greater than would have been expected naturally over time.

Measuring the sustainability of teacher conceptual change or enacted pedagogical processes is arguably more challenging than improvements in student learning. If Bobis et al. (2005) are referring to maintenance of teacher gains, for example, gains in confidence, mathematical content knowledge and pedagogical knowledge or use of mathematical language, does just knowing about these mean sustaining them or is some level of adaptation required?

The longevity standard becomes difficult to measure because curriculum innovations invariably depart from their original designs. How much change is acceptable? Which is more important in the quest for sustainability; longevity of the project's vision and goals, or working to have the project survive in any fashion? Hamilton et al. (2003, p. 1) stated "The obstacles researchers face when evaluating such programmes include variations in the implementation of the reform combined with a lack of information about actual practice; variation in outcome measures available across sites and a lack of a straightforward set of analytic methods applicable to replicated cases." Their work on large scale reforms in mathematics and science in the United States outlined the difficulties in determining whether improved outcomes were associated with the desired changes in practice. In order to do this, implementation needed to be measured directly and associated with changes in student performance. These variations appear essential if the project is to be adapted to the unique conditions of each school.

A number of recent discussions and projects on sustainability are repeating a traditional fixation of keeping something going over time with continued support from external providers. They reduce it to maintainability, to the question of how to make the improvement last, and add little to the analysis of adaptation and institutionalization. Cuban (2002) suggested that the practitioner derived standard of adaptiveness became essential prior to applying any other criteria. Coburn (2001, p. 145) argued that teacher interpretation and adaptation occurred at the outset of the reform although there was little research into the "processes by which such interpretations and adaptations occur." Snyder, Bolin and Zumwalt (1992) and Berman and McLaughlin (1978) maintained that some level of local development or adaptation seemed necessary to sustain curriculum reforms particularly when teachers were implementing a highly structured model as in the case of the NDP.

In summary, finding out whose standards are being used to judge the worth of an innovation and the exact content of those standards including what constitutes acceptable evidence become critical information in deciding whether and how to continue a project. For the purposes of this research, sustainability is measured by improved student outcomes, the development of teacher pedagogical content knowledge and mathematical knowledge to improve practice for the benefit of students, evidence of teachers inquiring into their own practice, teachers being responsive to students and evidence of a strong numeracy learning community.

METHODOLOGICAL APPROACH

Constructivist, socio-cultural and self-regulated (Butler & Vinnie, 1995) theories of learning provided a theoretical lens through which the sustainability of the NDP was examined. Bauersfled (1992) argued strongly for social constructivism in which meaning is constructed through discourse and interaction. A descriptive case study offered a methodological approach for this research because it could be formulated to suit the complexities of educational organisations and focused on gaining an in-depth understanding of phenomena in a real life setting. This enabled the use of multiple data collection techniques, multiple sources of data and was an ideal design for understanding and interpreting teachers' actions within their classroom, allowing for intensive study of specific instances and detailed contextual analysis of events or conditions. Quantitative data (from the questionnaires) enabled an examination of the impact of given variables on the curriculum reform outcomes. A quasi experimental (repeated measure) design was used for the student achievement data in the first year of the research project.

This research involved three higher socio-economic schools, their senior leaders and nine classroom teachers teaching students in years three to seven (grades two to six). All teachers had been involved

in the NDP for the previous two years and were identified as demonstrating changed practice, commitment and understanding to the principles and practices espoused through the NDP. Focusing on teachers whose current practice resonates with the reforms enabled a better understanding to be gained of the challenges involved in sustaining the practices. Teachers at these levels were specifically chosen because research data suggested that these teachers found the implementation of NDP more problematic with suggested pedagogical practices appearing less closely aligned to current practices. The research consisted of the following components repeated for six phases throughout an eighteen month period; semi-structured interviews for teachers, senior managers and numeracy leaders, classroom observations with follow-up discussions, student group interviews after the lesson, document analysis, collation of student achievement data- asTTle[1] and NumPA[2] (phases one and four only), teacher mathematical content knowledge and attitude data (phases one and four only) and observation of professional learning opportunities.

In order to analyse classroom lessons, an observation sheet covering pedagogical practice and inquiry practices concerning mathematical thinking was adapted from Fraivillig, Murphy and Fuson's (1999) pedagogical framework that supports children's development of conceptual understanding of mathematics (available on request). Data from the interviews were analysed using a process of content analysis. To determine teachers' content knowledge (TCK), teachers completed an asTTle curriculum levels three to five test and the Britt Algebraic Thinking Test (TCKALG) Sections A and B (Irwin and Britt, 2007).

FINDINGS

For the purposes of this paper, findings are reported in the following areas:

1. Student outcomes
2. Analysis of teacher pedagogical and mathematical inquiry practices
3. Teacher content knowledge
4. School systems and processes

1. Sustainability of Student Outcomes

Effect sizes were calculated for the nine classroom teachers using beginning and end of year asTTle and NumPa data. Effect sizes for asTTle ranged from 0.513 to 1.677 which are considered medium to large. Effect sizes for NumPA ranged from 0.073 to 1.545 showed more variation with a greater number of lower effect sizes.

Table 1

Means and Standard Deviations for asTTle and NumPA

Teacher	N	asTTle Mean 3/06	asTTle SD 3/06	asTTle Mean 11/06	asTTle SD 11/06	asTTle Effect Size	NumPA Mean 3/06	NumPA SD 3/06	NumPA Mean 11/06	NumPA SD 11/06	NumPA Effect size
A1	20	518.00	114.48	570.4	87.97	0.513	5.75	0.91	5.88	0.75	0.156
A2	22	343.73	71.36	420.14	87.33	0.958	4.45	0.60	5.14	0.64	1.112
A3	24	501.92	79.87	558.50	56.69	0.817	5.42	0.72	6.58	0.78	1.545
B4	18	469.17	82.49	541.50	79.98	0.890	5.06	0.80	5.17	0.86	0.133
B5	26	371.12	92.72	475.77	75.36	1.238	4.73	0.66	5.35	0.49	1.043
B6	22	496.77	101.88	577.86	92.52	0.833	5.50	0.74	6.09	0.64	0.925
C7	21	410.81	120.32	518.86	108.36	0.944	5.29	1.31	5.38	1.15	0.073

[1] asTTle – Assessment Tools for Teaching and Learning is a nationally referenced assessment tool which produces an overall score and a curriculum level and sub level ranging from 2 Basic to 6 Advanced.
[2] NumPA – The Numeracy Project Assessment tool administered as a diagnostic interview.

C8	21	490.33	89.30	565.48	81.63	0.878	5.57	0.87	6.047	0.86	0.623
C9	21	331.38	111.26	494.48	80.83	1.677	5.19	1.17	5.62	0.97	0.451

Year level asTTle gains in the research classes were considerably more than expected gains based on national samples of students. A Scheffe test of multiple comparisons on asTTle gain scores established that the teachers fell into two groups; Group 1- C9, B5, C7 and Group 2- B5, C7, B6, A2, C8, B4, A3 and A1. C9 is indistinguishable statistically from C7 and B5 in group 1 but is distinguishable from all other members of group 2 except C7 and B5. Initial results using Scheffe in February indicate that C9 is indistinguishable from four teachers with the exception of A1, A3, B6 and C8. Therefore we can conclude that the gains made are a function of the teacher not group membership. There was little difference in the gains made between schools on both tests.

Characteristics of the teacher (C9) whose class performed most highly on asTTle included:

- Highly motivated to improve classroom practices
- Individual student conferences about progress and achievement
- Thorough understanding of the NDP
- Very structured classroom
- Reads professional material including the NDP research reports
- Accesses support material from websites
- Focused on numeracy knowledge initially to address gaps

Student outcomes from the NumPA data produced very different results. NUMPA effect sizes ranged from 0.073 to 1.545. Again a Scheffe test of multiple comparisons on NumPA gain scores established that the teachers fell into two groups with four teachers represented in both groups. A3 is statistically indistinguishable from A2, B5, B6 and C8 but distinguishable from C9, A1, B4 and C7. The NumPA gain scores for School A were higher than those of Schools B and C.

Characteristics of the teacher whose class performed most highly on NumPA included effective teacher prompts during the group teaching sessions, a high level of responsiveness to students' thinking, the recording and analysis of incorrect as well as correct answers, the consistent eliciting of multiple strategies for problems and an approach based on "What is going to make it easy for them [students]?" Students in this class were very aware of their next steps for learning and could clearly articulate their individual knowledge gaps in number identification and sequencing, place value, basic facts and fractional numbers.

2. Sustainability of Teacher Pedagogical and Mathematical Inquiry Practices

Structural, pedagogical and practices that promoted mathematical inquiry and mathematical thinking were key elements of the NDP training. Scores for each individual indicator were aggregated for each teacher for the first group teaching session of the lesson.
Table 2

Teacher Pedagogical and Mathematical Inquiry Practices for Teaching Group One (phases 1-3)

Teacher	Phase 1 Feb 06	Phase 2 May 06	Phase 3 Aug 06	Total for Phases 1-3
A1	18	12	12	42
A2	8	8	8	24
A3	24	21	19	64
B4	20	22	10	52
B5	13	14	33	60
B6	13	11	13	37
C7	20	15	9	44
C8	9	12	11	32
C9	12	19	17	48

Teacher A3 had the highest NumPA effect size and scored amongst the highest for pedagogical and mathematical inquiry practices. Teachers B5 and C9 were also high scorers and achieved high asTTle effect sizes. In contrast, teacher A2 scored very low on pedagogical and mathematical inquiry practices yet her students made impressive gains on both NumPA and asTTle. She is considered and slow in her delivery with her practices remaining consistent but limited over the first three classroom observations (Phases 1-3).

Teachers were asked to describe and justify the adaptations they were making to the NDP from what they believed to be in the initial training. Teachers indicated a duplication rather than a development idea of sustainability stating that they intended to implement the strategies taught during the training and adhere closely to the teacher resource materials. However, teachers B5 and C9 recounted explicitly what they had changed and why. Adapting meant using different equipment than indicated in the teaching resources or missing lessons if teachers were unable to make sense of them after an initial read through. In some cases "adapting" meant reverting back to the rule they were taught at school despite a different approach suggested in the teaching sequence. For example: 2/3 of 21 = 3 divided into 21 is 7 and 7x2=14. Eight out of the nine teachers were not able to discuss what the impact of this might have on student achievement. The quality of the initial training had an impact on how teachers perceived the NDP. If teachers found an alignment of the NDP to their own beliefs about teaching and learning, the transition to new teaching practices appear less "painful." This was illustrated by teacher A2's comment:

> I think the whole programme is a bit hard for me. Because there's such a lot in it and sometimes it almost overwhelms me with the things that I'm supposed to be trying to impart… my maths has never been an easy subject for me as a person so I think that, learning the new way of maths after all these other ways of maths that I've had to learn, and having to now relearn a new way… that you really need to learn it to be able to teach it proficiently, that's the hard bit. That's why I keep on doing it because it's hard and I want to get better at it. (Teacher A2 Phase 1 Feb 06)

3. Teacher Content Knowledge

Teachers completed an asTTle test with the results ranging from level 4 advanced to level 6 proficient indicating a wide range of mathematical ability equating to year seven to year eleven student groups.

Table 3

Results of teacher content knowledge

Teacher	TCK Feb 2006	Ranking	TCKALG % Nov 2006	Ranking
A1	918-5A	4=	40	6=
A2	766-4A	9	40	6=
A3	842-5B	7=	40	6=
B4	918-5A	4=	90	3=
B5	857-5B	6	100	1=
B6	976-6P	1	90	3=
C7	935-5A	3	80	5
C8	842-5B	7=	40	6=
C9	954-6B	2	100	1=

Teachers B6 and C9 performed well on both tests. Teacher A2 performed poorly on both tests.

Table four outlines a correlation between teacher content knowledge and student achievement effect sizes.

Table 4

Correlation between Teacher Content Knowledge and Student Achievement Effect Sizes

		TCK	**TCKALG**	**asTTle effect size**	**NumPA Effect size**
TCK	Pearson Correlation	1	.378	.124	-.560
	Sig. (2-tailed)		.316	.750	.117
	N	9	9	9	9
TCKALG	Pearson Correlation	.378	1	.599	-.066
	Sig. (2-tailed)	.316		.089	.866
	N	9	9	9	9
asTTle effect size	Pearson Correlation	.124	.599	1	.052
	Sig. (2-tailed)	.750	.089		.895
	N	9	9	9	9
NumPA effect size	Pearson Correlation	-.560	-.066	.052	1
	Sig. (2-tailed)	.117	.866	.895	
	N	9	9	9	9

A correlation of 0.66 is needed for significance at the 5% level therefore none of the results are significant. The better teachers did on the TCK test, the higher the gains students made on the asTTle test. This relationship is more pronounced with the TCKALG test. There is a correlation of .599 between the TCKALG and asTTle results. Gains on asTTle have no relationship to gains on NumPA. NumPA appears to be a tool that helps lower content knowledge mathematics teachers. Teacher content knowledge was not necessarily a good indicator of student outcomes. TCK as measured by asTTle is not a strong predictor of ability to cause learning. The teacher (A2) with the lowest TCK and lowest equal TCKALG achieved high effect sizes in both asTTle and NumPA. On her TCKALG test paper she wrote "Haven't got a clue. I've never been taught this stuff and never needed it so of course I don't know!!" A little further on in the test "Can't be bothered trying to work this out. I've got no clue at all …at all!" (Teacher A2)

4. School systems and processes

Professional learning communities

Teachers appreciated the informal discussions they had with their colleagues (often a teacher working at the same year level) and found this an invaluable support mechanism for further learning. Schools averaged one to two whole staff meetings a year with a focus on the more structural elements of the NDP such as NumPA testing. Over the eighteen months, there were few differentiated learning opportunities for staff with seven teachers receiving no feedback on their classroom practice. This left them wondering if they were doing the right thing. The Education Review Office provided positive feedback to two teachers in school A despite this not being part of their usual practice.

Use of student achievement data

At the beginning of their third year teaching the NDP, all teachers had access to the previous end of year student achievement NDP data was used to group students for instruction using strategy stages

for addition and subtraction. Most teachers used student participation and attitude to gauge the success of their teaching rather than student outcomes. School A's principal was highly skilled at data analysis and interpretation and staff relied on her for this expertise.

The role of the numeracy leader

The role and activities of the numeracy leaders varied considerably between schools. School B released a numeracy leader one day a week. Her previous role had been as a school support services numeracy facilitator. Her health failed half way through the year and she relinquished the position which was not filled by anyone else in the school. School C appointed the current deputy principal to the fully released role of curriculum director in the second year of the research. The assistant principal fulfilled the role in the first year of the research although indicated that mathematics wasn't her strength and she lacked confidence in providing support for classroom practice as she had not been teaching the NDP during the initial training.

DISCUSSION

Factors contributing to sustained changed practice can be categorized into three capacities; personal, interpersonal and organizational/systemic. Each of these is discussed in turn.

Personal Capacity

The personal attributes of the teacher had a significant impact on their ability to cope with sustaining changed practice. The teachers (A3 and C9) whose classes performed most highly, embraced the changes with enthusiasm and actively continued their own learning often with little support from the schools' senior managers. These teachers reflected on and critiqued their own practice, accessed additional resources and information from research reports and websites and attended area wide professional learning sessions. Earl, Watson and Torrance (2002) confirmed this by claiming that ultimately any changes that occur in schools happen because of the motivation of individual teachers teaching children in classrooms. Teacher C9 likened teaching the NDP to being a skilful chef. "But if you understand cooking, you're a chef, you can just take ingredients and pull them together and the numeracy project gives the children the skills to be able to take any aspect of number and manipulate it however they want" (Teacher C9 Phase 1 Feb 06). Interestingly, the teacher (A2) most challenged by the changes advocated through the NDP, compensated by being very considered, didactic and methodical in her approach with students and to learning the new material. Learning mathematics was challenging for her at school so she spent many hours at home preparing the lessons and making sure the students understood the key mathematical concepts. However, she knew when she had done enough.

> I just can't think that out, that concept would take me days to get to think it out. And I've already thought about it for days. That's it. And I'm not really prepared to spend that time doing it again when I can teach them perfectly well as far as I'm concerned in a way that they understand. I don't understand that lesson at all, it was a total disaster. (Teacher A2 Phase 3 August 2006)

Teacher beliefs about teaching and learning mathematics had an impact on their approach to suggested pedagogical and mathematical inquiry strategies. In all three schools these were not accessed as part of the initial training in NDP. Timperley et al (2007, p. 222) noted "teachers need to have time and opportunity to engage with key ideas and integrate those ideas into a coherent theory of practice." Any strategies that were contrary to their philosophical beliefs about how students learn mathematics were eventually dropped from their repertoire. This is illustrated by teachers A2 and B4. Teacher A2, for example, used a limited number of pedagogical and mathematical inquiry practices which remained consistent over the entire research period.

Interpersonal Capacity

If sustainability of outcomes is associated with ongoing professional learning experiences, then the nature and type of these experiences needs to be examined further. All three schools provided limited (one to two) whole staff professional learning opportunities in numeracy over the eighteen month research period with all teachers stating that most support and learning was developed through informal discussions at team meetings or a discussion one on one with a colleague. Discussions centred on clarification of lesson structure, understanding of mathematical knowledge and pedagogical approaches and discussion of appropriate resources. Timperley et al. (2007, p.216) noted a number of factors in the professional learning context that may contribute to sustainability such as teachers having "ongoing opportunities to deepen relevant knowledge and skills and to work and learn collaboratively with colleagues as they tested the impact of their teaching on student outcomes." Interestingly, teachers didn't question the accuracy of the information they were receiving from colleagues and rarely discussed student outcomes. Higgins (2004, pp. 54-55) also claimed the importance of peer support structures. She noted:

> Peer support is an important component of the internalisation process for individual teachers. In broad terms this is about teachers talking together about the changes they are making to their classroom practice and their developing knowledge of the Framework and how best to develop children's understanding of number…Some schools encouraged teachers not only to talk together, but to also observe each other's practice as a means of developing it. One lead teacher described the characteristics of those likely to sustain the changes as being "open minded, willing to learn and to change their ways of maths teaching were applicable, keen to share and to be observed.

Organizational Capacity

The existence of change goals within a sustainability plan was not sufficient in itself to sustain change. It was more about the quality of the goals, the school's execution of the goals and the personnel involved. Only one of the three schools (School B) developed a comprehensive NDP sustainability plan for the third year of its implementation. However this appeared not to advantage the school. Once the numeracy leader in this school was unable to continue due to ill health, goals on the plan were sidelined until her return the following year. Such a reliance on one individual to manage staff learning proved to be problematic for the school with no-one feeling confident to take over the responsibility. All had a sense of being "new to this." Timperley et al. (2007, pp. 215-216) commented "Sometimes the conditions that make for sustainability are not considered until the end of a professional development programme—or even after it has finished. Our view is that the conditions for sustainability are set in place during the professional learning experience as much as after it. For example, if ongoing inquiry into the impact of practice on students is a condition associated with sustainability, then teachers must given the opportunity to learn the skills to engage in such inquiry." School A had a larger group managing the NDP within the school and this spread the risk. The principal coordinated the data analysis and interpretation and the numeracy leaders organised the resources and coordinated the programmes within the junior and senior teams. Numeracy leaders were created at a number of levels within the school. Fullan (2001, p. 18) points out "the main reason that change fails to occur in the first place on any scale, and does not get sustained when it does, is that the infrastructure is weak, unhelpful or working at cross purposes."

A common theme amongst the schools was a belief that the changes were having a positive impact on student outcomes and on their ability to teach mathematics more logically and coherently and more specifically to understand the student progressions in acquiring number strategies and knowledge. For some teachers, outcomes related to student attitude and enthusiasm of participation rather than student achievement despite Timperley and Wiseman (2003) argued that student achievement data and evidence are increasingly important dimensions for educational decision making. Although evidence is valued to inform decisions, many of the teachers were not sufficiently data literate to interpret and use it appropriately. A key step for the long term development of the numeracy projects in New Zealand is continued teacher professional development in data collection, aggregation,

interpretation and use. This should include factors that are associated with the highest levels of student achievement.

The simplicity of programme resources was a factor in teacher uptake and the length of time teachers continued to use the programme resources. If teachers didn't understand the lessons in the NDP booklets after reading them (once or twice), they were ignored or not included. In subsequent months, the more skilled teachers (A3 and C9) were less reliant on the booklets as a guide, using the underlying strategy as a guide to their approach and practice. Young-Loveridge (2004, p.39) indicated that the high levels of compliance required for the NDP in New Zealand initially did not necessarily bode well for long term sustained practice in the future. "In some places, there is evidence that attention to the strategies might be short-lived, superficial compliance." In future, schools will need to balance this prescriptiveness with the need for flexibility; teachers to exercise choices, to see purpose in their learning, and to have opportunities to develop their voices (Lieberman & Wood, 2003).

Coburn (2001) argued that teacher observation and feedback were more beneficial than just teachers reflecting and talking about their practice. Over the research period, seven of the nine teachers received no feedback on their classroom practice which left them feeling frustrated and not confident about the effectiveness of their practice.

CONCLUSION

This research has intended to capture the complex vision of sustainability with the initial findings providing some explanation and understanding of effective and equitable practice required for long term generative mathematics improvement. In particular, the research informs our knowledge about the relative contributions of disparate factors in the engagement and sustainability of large-scale educational reform. There is no doubt that the NDP has made significant changes to the way number is taught in New Zealand primary schools. Arguably, sustainability is about building on change and developing internal capacity for continued professional learning. From Higgins' (2004) perspective, sustainability increasingly presents itself as an ethical and political activity, as well as a learning, a teaching and an epistemological activity. The current emphasis should be focusing on investment in individual capacity-building, local creativity, reflection through goal setting and networking.

The NDP was conceived as a dynamic, evidenced based initiative. Therefore it is important for developers and researchers to gain a better understanding of the explicit and implicit standards used to determine the sustainability of a major curriculum innovations and determine which standards carry the most weight and why. If we are serious about changing schools, then we must be equally serious about the issues relating to sustainability. Clearly, without paying close attention to this issue, we are destined to continue down a path of fragmentation.

REFERENCES

Amit, M., & Hillman, S. (1999). Changing mathematics instruction and assessment: Challenging teachers' conceptions. In B. Jaworski, T. Wood & S. Dawson (Eds.), *Mathematics teachers education: Critical international perspectives.* (pp. 17-25). London: Falmer

Bauersfled, H. (1992). The structuring of the structures: Development and functioning of mathematics as a social practice. In L. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 137-144). Hillsdale, N.J.: Erlbaum.

Berman, P., & McLaughlin, M. W. (1977). *Federal programs supporting educational change. Volume 3: Factors effecting implementation and continuation.* Santa Monica, California: The Rand Corporation.

Bobis, J., Clarke, B., Clarke, D., Thomas, G., Wright, R., Young-Loveridge, J., Gould, P. (2005). Supporting teachers in the development of young children's mathematical thinking: Three large scale cases. *Mathematics Education Research Journal*, 16(3), 27-57.

Butler, D. L., & Winnie, P. H. (1995). Feedback and self-regulated learning: A theoretical analysis. *Review of Educational Research, 65*(3), 245-281.

Century, J. R., & Levy, A. J. (2002, 3 April). *Sustaining change: A study of nine school districts with enduring programs*. Paper presented at the Annual Meeting of the American Educational Research Association.

Coburn, C. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis, 23*(2), 145-170.

Cuban, L. (2002). *Answering tough questions about sustainability*. Retrieved 16 May, 2004, from http://sustainability2002.terc.edu/invoke.cfm/lsc_project/214

Davison, A. (2001). *Technology and the contested meanings of sustainability.* Albany, New York: State University of New York Press.

Earl, L., Watson, N., & Torrance, N. (2002). Front Row Seats: What we've learned from the National Literacy and Numeracy Strategies in England. *Journal of Educational Change, 3*, 35-53.

Fraivillig, J. L., Murphy, L. A., & Fuson, K. C. (1999). Advancing Children's Mathematical Thinking in Everyday classrooms. *Journal for Research in Mathematics Education, 30*(2), 148-170.

Fullan, M. (2001). *Leading in a Culture of Change.* San Francisco: Jossey-Bass.

Hamilton, L. S., McCaffrey, D. F., Stecher, B. M., Klein, S. P., Robyn, A., & Bugliara, D. (2003). Studying large-scale reforms of instructional practice: An example from mathematics and science. *Educational Evaluation and Policy Analysis*, 25(1), 1-29.

Higgins, J. (2003). *An Evaluation of the Advanced Numeracy Project 2002:* Wellington: Ministry of Education.

Higgins, J. (2004). *An Evaluation of the Advanced Numeracy Project 2003*. Wellington.

Irwin, K., & Britt, M. (2007). The development of algebraic thinking: Results of a three year study. In Ministry of Education (Ed.), *Findings from the New Zealand numeracy development projects 2006.* (pp. 33-43). Wellington: Ministry of Education.

Lieberman, A. & Wood, D. (2003). Sustaining the professional development of teachers: Learning in networks. In B. Davies and J. West-Burnham (Eds.) *Handbook of educational leadership and management.* (pp. 478-490). Harlow: Pearson Education Limited.

Ministry of Education (2004). *Teacher professional learning 2004 effective numeracy teaching in years 1 to 8: Sustaining the numeracy project.* Wellington: Ministry of Education.

Snyder, J., Bolin, F., & Zumwalt, K. (1992). Curriculum implementation. In P. Jackson (Ed.), *Handbook of research on curriculum* (pp. 402-435). New York: Macmillan.

Spillane, J.P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: the mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, 31(2), 143-175.

Timperley, H., & Wiseman, J. (2003). *The sustainability of professional development in literacy Part 2: School-based factors associated with high student achievement.* Auckland

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration.* Wellington: Ministry of Education.

Young-Loveridge, J. (2004). *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. (2005). Patterns of performance and progress: Analysis of 2004 data. In J. Higgins, K.C. Irwin, G. Thomas, T. Trinick, & J. Young-Loveridge. *Findings from the New Zealand numeracy development project 2004* (pp. 5-20). Wellington:Ministry of Education.