Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author’s right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.
http://researchspace.auckland.ac.nz/feedback

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.

Note : Masters Theses

The digital copy of a masters thesis is as submitted for examination and contains no corrections. The print copy, usually available in the University Library, may contain alterations requested by the supervisor.
Is sustainability of schooling improvement an article of faith, or can it be deliberately crafted?

Pamela June O’Connell

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Education

The University of Auckland, 2010
Sustainability of reform has been characterised in various ways in the literature, but most often as some sort of endurance test for educators. School leaders and teachers are left to weather the “tempests” of reduced funding, shifting priorities, and constant turnover of teachers and leaders (Wood, 2007). The review of research literature exposes sustainability as a problematic notion, since a plethora of ideas about school-based and system conditions has been reported and very few studies connect these conditions to ongoing achievement of student outcomes (Timperley, Wilson, Barrar & Fung, 2007). Rarely have researchers regarded sustainability in terms of how one project melds into the next initiative in a school. For the most part, schools have to beat their own path towards a nebulous destination, and they do.

This research reframes sustainability as two interdependent and dynamic continua that form the conceptual framework, namely, coherence of effective instructional practices and co- and self-regulated inquiry. The purpose was to investigate the significance of these dimensions in two interconnected studies using a mixed methods case study approach. Study 1 examined 16 schools, tracking students’ literacy achievement and participants’ actions in the two years following their participation in a national literacy programme in New Zealand. Ten of the thirteen schools that presented their data maintained or improved their gains in student achievement. Study 2 investigated four of these schools in more depth over 2006–08, their selection contingent on their having varied combinations of the framework dimensions and distinctively different mediating and outcome variables.

Analysis of findings from both studies suggested that participants’ perspectives of sustainability were limited to a single domain and to maintenance of newly developed practices, rather than being focused on an improvement paradigm. Inquiry as a model for schooling improvement was not fully embedded in these schools. However, when evidence-based practices were more systematically applied, this condition appeared to be a sufficient threshold to support sustained outcomes for new cohorts of students. Engaging with persistent issues of underachievement and establishing coherence of instructional practices were more strongly evidenced in those schools that improved on their achievement gains over time.
ACKNOWLEDGMENTS

My undertaking a doctoral study was partly serendipitous and partly typical of my enthusiasm for trying to seek out better solutions in my career as a teacher-educator. I manage the national Literacy Professional Development Project (LPDP) in New Zealand on behalf of the Ministry of Education as part of my role at Learning Media Limited. Our project directors design and implement the intervention with eight team leaders and 20 facilitators. The first cohort of 91 schools in this project was demonstrably successful but, as a project team, we were unsure whether the shifts made in teachers’ knowledge and pedagogy would be sufficient to support achievement gains for future cohorts of students: we simply needed to know. Would the significant learning that we had undertaken with our research team (Professor Helen Timperley and Associate-Professor Judy Parr from the University of Auckland) and the extraordinary efforts made by our facilitators, school leaders, and teachers be enough to ensure ongoing impact? We had paid particular attention to the research literature about sustainability, but we were operating in an “act of faith” paradigm, so aptly described in Teacher Professional Learning and Development: Best Evidence Synthesis Iteration (Timperley, Wilson, Barrar & Fung, 2007). The opportunity to undertake this research came at the end of 2005, when the research team did not have immediate capacity to add sustainability to their agenda for the following year. So I volunteered to take a rather large learning step – it certainly had not been on any career plan, but it seemed critical and central to the work ahead. It was also an opportunity to take a risk, to be a learner again under more formal conditions.

I had taken up many opportunities to write and publish texts for students as a secondary teacher of history, but this experience has been a much longer journey. This thesis represents the thinking of many minds. I have been supported enormously by my colleagues at Learning Media who have stopped to talk at the photocopier, shared gems on the subject, or given me feedback on research papers. The facilitator and leadership forums associated with the LPDP were another great source of support and critique as I continued to ask questions and grapple with being a researcher. In particular, I want to thank Carolyn English and Lyn Bareta, who co-direct the LPDP. They offered insights when I was perplexed and encouragement and humour when I needed it. Ann Anaru was my research assistant, and we learned much together as we coded and rated endless transcripts. Ian Reid and Marguerite Chadwick designed many of my diagrams as they
developed from a quick sketch. Marguerite also superbly supported administrative tasks related to the research and Jeannie Beauchamp taught me much about grammar and punctuation. I am also grateful to the Board of Learning Media who enabled me to have sabbatical leave to write this thesis. The journey was made much easier as a result.

The New Zealand Ministry of Education has funded this research as part of their literacy strategy. In particular, I want to thank two individuals who have responded to my work in ways that have helped me think more about system responses and the complexity of finding coherence across research, policy, and practice. Denise Arnerich and Dr Brian Annan are both passionate about what systems can achieve. They are making a difference for students with their drive to align their work programmes and to consider evidence of what works, often in the face of difficult bureaucratic boundaries.

My supervisors, Helen and Judy, have been exceptional teachers. Their scholarly critique directed me to new ideas and perspectives. I met them both when they evaluated a previous initiative that we had also developed and implemented at Learning Media. Their evaluation of the Literacy Leadership in New Zealand Schools Programme was, I believe, a watershed for in-service educators. I knew then that we would need both of them to support our learning as we went forward with our next project, especially when we needed to make a more lasting impact on students. I also need to thank the series of data analysts that worked with the research team at the University of Auckland – Paul Reddish, Cole Armstrong, and Kane Meissel. Kane, in particular, took a keen interest in the results and contributed much to our thinking about how to best design benchmarks for sustainability.

The schools, their leaders, teachers, and students were the highlight of this research. This work took me back to the classroom, and that was a great privilege. I want to thank those schools who participated for their willingness to have the disruption of a researcher in their midst at different times. They shared inspirational ideas and focused me on the balance between theory and practice.

Finally, my husband Craig has kept me in touch with the world by inviting our friends and family to share a meal with us most weekends. He is a wonderful cook, and this meant I could just keep writing. He works in the area of organisational change as well, so he provided me with a valuable perspective, other than education, on my work.
My experiences of those who work in education are that they are mostly driven people. It is important for them to know that they do make a difference to their students’ lives. Certainly this research was determined by that endeavour.
# TABLE OF CONTENTS

ABSTRACT .............................................................................................................................. II  
ACKNOWLEDGMENTS ......................................................................................................... III  
LIST OF FIGURES ............................................................................................................... VIII  
LIST OF TABLES ................................................................................................................... IX  
CHAPTER 1 INTRODUCTION ................................................................................................. 1  
  Key themes ................................................................. 1  
  Purpose of the research ................................................................. 3  
  The educational problem of sustainability ................................................................. 3  
  Research questions ........................................................................................................... 5  
  The context of the research study .................................................................................... 7  
  Organisation of the thesis .............................................................................................. 11  
CHAPTER 2 LITERATURE REVIEW: REFRAMING SUSTAINABILITY ......................... 13  
  What is the unit for sustainability? ................................................................................ 13  
  Time and context bound ................................................................................................. 15  
  Towards a definition of sustainability .......................................................................... 16  
  School-based conditions that promote sustainability .................................................. 28  
  A conceptual model for sustainability ............................................................................ 45  
CHAPTER 3 STUDY 1 SUSTAINABILITY: A PROBLEM OF DEFINITION AND PERSPECTIVE 49  
  Introduction: An investigation into 16 schools 2006–07 ............................................... 49  
  Methodology .................................................................................................................. 49  
  Ethical considerations ................................................................................................. 55  
  Selection of cases ......................................................................................................... 56  
  Data gathering instruments and procedures .................................................................. 58  
  Validity and reliability of student assessment tools ...................................................... 63  
  Analysis of qualitative data ......................................................................................... 65  
  Analysis of student achievement data ........................................................................ 69  
  Limitations and key assumptions ............................................................................... 72  
CHAPTER 4 RESULTS AND DISCUSSION: SCHOOLS’ PERSPECTIVES ON SUSTAINABILITY 73  
  Theories of action for sustainability ............................................................................ 73  
  Coherence as a theory of action for sustainability ......................................................... 88  
  Self-limited understandings of sustainability ............................................................. 91  
CHAPTER 5 SUSTAINABILITY OF STUDENT ACHIEVEMENT GAINS ......................... 93  
  Problems of measurement ............................................................................................ 94  
  Patterns for schooling improvement post implementation ....................................... 97  
  Benchmarks for sustainability of achievement gains ................................................... 99  
  Results for 16 case schools 2006–07 .............................................................................. 102  
  Comparative study ....................................................................................................... 105  
  Significance of findings for sustainability ................................................................. 106
LIST OF FIGURES

Figure 1. Teacher inquiry and knowledge-building cycle to promote valued student outcomes...38
Figure 2. Conceptual model for sustainability.................................................................47
Figure 3. School L achievement data in reading with successive cohorts of students........104
Figure 4. School P achievement data in reading with successive cohorts of students........105
Figure 5. Coherence and inquiry scores in 16 schools.......................................................129
Figure 6. School L reading comprehension results 2004–08............................................149
Figure 7. School L student achievement gains for at-risk students 2004–08.................150
Figure 8. School H reading comprehension results 2004–08............................................164
Figure 9. School H student achievement gains for at risk students 2004–08..................165
Figure 10. School A asTTle writing results 2004–08......................................................173
Figure 11. School A asTTle writing results for at risk students 2004–08.........................174
Figure 12. School C asTTle writing results 2004–08.......................................................181
Figure 13. School C asTTle writing results for at risk students 2006–08.........................182
Figure 14. Sustainability of educational improvement as two dimensions.......................204
LIST OF TABLES

Table 1. Comparison of Study 1 case schools to the LPDP Cohort 1 2004–05 ......................... 57
Table 2. Number and types of participants in Study 1 schools ................................................ 58
Table 3. Coding of participant responses to question on use of achievement data .............. 68
Table 4. Continuing or shifting professional learning foci for 16 case schools ...................... 75
Table 5. Principals’ descriptors of sustainability ................................................................. 77
Table 6. Leaders’ views of the key messages of the LPDP ................................................... 81
Table 7. Leaders’ indicators for sustainability of the LPDP ................................................ 85
Table 8. Teachers’ views of key messages of the LPDP ....................................................... 87
Table 9. Leaders’ views on connections between professional learning projects .................. 91
Table 10. Summary of student achievement results for 16 case schools 2006–07 ................. 103
Table 11. Case school contexts: decile, size, staff turnover, and sustainability ..................... 110
Table 12. Leaders’ responses about purpose and use of student achievement data .............. 112
Table 13. Analysis of school documents in 14 schools related to inquiry processes .............. 115
Table 14. Teachers’ responses about their purpose for analysing student achievement data .... 117
Table 15. Teachers’ responses about strategies for underachievement ............................... 118
Table 16. Inquiry and coherence aspects of literacy lessons in 16 case schools ..................... 120
Table 17. Self-regulation practices by students in 16 schools ............................................. 122
Table 18. Professional learning foci for teachers in literacy in 16 schools ............................. 123
Table 19. Analysis of school documents in 14 schools ....................................................... 126
Table 20. Teachers’ views on connections between professional learning projects .............. 127
Table 21. Numbers and types of participants in Study 2 schools ......................................... 140
Table 22. Study 2 school context variables 2006-08 ............................................................. 140
Table 23. School L STAR mean stanine gains 2004–08 ...................................................... 150
Table 24. School H STAR mean stanine gains 2004–08 ...................................................... 164
Table 25. School A net gains in asTTle writing points above mean gains calculated from asTTle norms ................................................................. 173
Table 26. School C net gains in asTTle writing points above mean gains calculated from asTTle norms ................................................................. 181
CHAPTER 1
INTRODUCTION

Key themes

The title of this thesis embraces the essence of the educational problem that confronts both New Zealand and international educators, namely the increasing pressure for lasting improvement to result from investment in large-scale professional learning projects in schools (Elmore, 2005; Fullan, 2006; Levin, 2008). The precise nature of this lasting improvement is not yet agreed among researchers and descriptions range from sustained changes in teacher practices (Century & Levy, 2002; Fullan & Hargreaves, 1996; St John, 2002), to a requirement for longer-term and accelerated impact on student outcomes (Lai, McNaughton, Timperley & Hsiao, 2009). The stance taken in this research is that improvement in teacher capability is a critical factor in raising student achievement (Alton-Lee, 2003) and that sustainability of reform must be measured for its lasting impact on student achievement (Timperley, Wilson, Barrar & Fung, 2007).

The question itself is derived from the following statement in Teacher Professional Learning and Development: Best Evidence Synthesis Iteration (Timperley et al., 2007):

Sustainability was not neglected in the literature, but it was treated as an article of faith rather than a condition subject to empirical verification. (p. 219)

These researchers located 97 New Zealand and international studies of professional learning that led to substantive and improved outcomes for students, and described the key characteristics of such endeavours that were associated with medium and high impact on learners. Only seven studies offered any insights into the impact on students after the input of expert support, resources, and/or funding had ceased.

Indeed, a plethora of ideas about district- and school-based conditions for sustainability is advanced in the research literature but, for the most part, it is left to the designers of professional learning and school leaders to make sense of these data and to forge their own path towards a nebulous destination, and they do.

To illustrate the key themes of this research, I have drawn on two statements from school leaders. The first comment was made by the literacy leader of a school in the study – a school where, three years from when they had participated in New Zealand’s
national Literacy Professional Development Project (LPDP), students’ achievement gains in literacy had almost doubled compared to when the school was in the project. The comment illustrates a key dichotomy, highlighted by this research, between maintenance and fidelity to project practices versus ongoing and systematic inquiry for improvement:

We are going along pathways, so we are not doing the same as we were doing a year ago [2007] . . . so we are not still in the same place we were at the end of the project [2005] (Literacy leader, School L, 2008).

This school leader viewed sustainability as being well beyond maintenance of the literacy practices that participants had learned in their two years of involvement in the LPDP. She understood the need to employ an iterative stance to improvement. When shifts in students’ literacy achievement were achieved each year, this school systematically investigated the remaining group of underachieving students, responding deliberately and in different ways when current practices were not supporting their needs. Other schools in this study took a more recursive stance to sustainability, maintaining fidelity to the new practices learned during the project with a new cohort of students, without explicitly refocusing on those students who were still not making adequate progress.

The second key theme of the research is the notion of coherence between professional learning projects. This is exemplified by a principal’s sense-making about the sustainability of her school’s involvement in the second cohort of the LPDP in 2006–07:

If it wasn’t for the timely arrival of our facilitator I suspect a lot of our learning [from another professional development project] would have remained as theory and not have been implemented in our classrooms. Even now the sustainability of what we learnt is dependent on vigilant monitoring and reinforcing . . . . We find that we are naturally linking our knowledge from both PD programmes and making practical connections between the two in the classroom. It’s like a Venn diagram – there is a big chunk of common knowledge and practice in the intersections of the Numeracy to Literacy “circles.” I am also using LPDP-gained knowledge to work with teachers in their numeracy lessons and accompanying appraisal observations. (Email, June 18, 2006)
While external providers can influence the outcome, sustainability is inevitably about what the school does in the wake of reform. This principal was crafting her own pathway to sustainability only six months into the project. Her approach illustrates the idea that sustainability does not begin at the end of a project, nor is it about a discrete schooling improvement project or programme of materials. Instead, the argument presented in this thesis is that sustainability of professional learning must be planned for and be explicit within the school, and that student outcomes will continue to be improved when successive professional learning initiatives are seen to contribute to and amplify coherent and effective practices across the curriculum. The current study findings confirm that there is not only a problem of definition for sustainability for educators, but also that there are self-limiting perceptions of the notion of sustainability at work in their schools.

**Purpose of the research**

The purpose of this research was firstly to examine the nature of sustainability in educational improvement initiatives, in particular when teachers and schools had moved well beyond an initial, more intensive phase of their professional learning with an external facilitator. Secondly, the aim was to define more precisely the pathways and destinations for sustainability, the relationship between, and relative importance of, particular components that research literature postulates as necessary conditions for sustainable professional learning programmes, particularly for those programmes that aim to make significant changes to teachers’ beliefs, knowledge and practices in order to raise students’ learning and achievement.

**The educational problem of sustainability**

Issues of sustainability have moved to the forefront of discussion amongst those concerned with designing, resourcing, and implementing effective professional development. For example, large-scale educational reform is expensive, and so those involved are increasingly accountable for improving student achievement outcomes (Konstantopoulos & Hedges, 2008). The imperative that New Zealand and other countries currently face, is the urgency of improving the learning achievement trajectories for those students who are at risk of never catching up over time to their higher performing peers (Elley, 2005). While New Zealand students achieve well overall internationally in literacy, there is a wide gap between those at the highest and lowest
bands of achievement (OECD, 2001, 2004; Ministry of Education, 2008), and the data also reveal a disparity of achievement for particular ethnic groups, such as Māori and Pasifika students. Therefore, any professional learning that demonstrates gains for these groups must be sustainable over time and replicable in a large number of schools.

Sustainability is problematic both from the point of view of a shared definition of the educational context and from lack of empirical evidence of the ongoing outcomes of initiatives. The problem arising from the discussions about the urgent need for improved and sustainable outcomes for students is that so few studies investigate the sustainability of educational professional learning initiatives beyond the years of implementation; fewer still link changes in teacher knowledge and pedagogy made at the time to any ongoing improvements in student outcomes (Coburn, 2003; Elmore, 2005; Knight, 2005; Timperley et al., 2007). Indeed, the focus of research on this topic is often on the difficulties that school leaders and teachers face in maintaining any attention on improvement in the face of teacher turnover, changing priorities, and the dissipation of funding supports (Cuban, 1984; Datnow, 2005; Datnow, Hubbard & Mehan, 2002; Honig & Hatch, 2004; Wood, 2007). Moreover, a common measure used in research to judge the sustainability of professional development is self-report and teacher satisfaction with their own learning; but improved teacher confidence alone does not always guarantee improved student outcomes (Timperley et al., 2007).

The other problem that emerges from the current research literature is the plethora of conditions that researchers say contribute to the sustainability of professional learning at school level. These include: leadership for improvement (Fullan, 2006; Robinson, 2007; Spillane, 2006); inquiry skills so that leaders and teachers can assess if their changed practices are having the required impact on student outcomes (Copland, 2003; Earl & Katz, 2006; Reid, 2004; Robinson & Lai, 2006; Sutherland, 2004); sufficient depth of content and pedagogical knowledge for teachers (Bishop & O’Sullivan, 2006; Elmore, 2005; Parr & Timperley, 2006); communities of practice that enable ongoing and shared ownership for improvement in student outcomes (Elmore, 2002a; Lima, 2001; Little, 1993) and learning processes for teachers that penetrate teacher beliefs about learning and the efficacy of the practices they engage in so that principled knowledge can be transferred to different contexts and curricula (Timperley et al., 2007). Substantive research lies behind each of these conditions. Therefore, the issue for educators is how to balance and apply these to professional development design, alongside the knowledge
that context matters in all engagements with teachers (MacBeath & Dempster, 2009). The prior knowledge about teaching and learning that teachers bring to their learning, the structures and processes that already exist for teacher learning, the teacher practices in the school that endure and are untested for their efficacy, the students’ strengths and gaps in achievement and, finally, the expectations of the teachers and wider school community all provide a rich context that must be considered in the learning design for professional development and the expectations set for improvement.

For educators, evaluating whether professional learning has been sustained is as problematic as the lack of its precise definition. What lifts in student achievement can schools realistically expect over successive cohorts of students? The lack of longitudinal data to compare achievement patterns over time and the use of widely different, and relatively new, assessment tools by schools in the New Zealand context compound the problem.

Adding to this complexity for educators is the increasing importance of the wider educational setting in supporting sustainability of professional development, the policy and practices at system level that enable schools’ access to resources that support them to improve student outcomes. Such resources may include policy implementation, teacher materials, professional development programmes, and/or curriculum design (Cohen & Ball, 1999; Levin, Glaze & Fullan, 2008).

**Research questions**

The research questions sought to address this complexity alongside the critical need for further empirical evidence on how schools might sustain the early gains made during implementation. The research aimed to develop a conceptual framework for sustainability that could be used to promote more effective design and implementation of intervention programmes at system and school levels. The research questions included:

- What is the nature of sustainability?
- What perspectives do school leaders and teachers have about what leads to sustainability of professional learning initiatives?
- How do these perspectives impact on their actions to sustain the professional learning initiatives in their school?
- What is the relative importance of the conditions that are necessary to sustain reforms well after their initial implementation?
- Do school leaders and teachers isolate or transfer the learning that they appropriate in professional learning?
• Does the development of coherent, effective instruction in schools across the curriculum advance sustainable outcomes for students?

The hypothesis being examined in this research was that sustainability of professional learning that supports ongoing improved outcomes for students is determined by two continua: the programme’s contribution to the coherence of effective instruction across the curricula in a school and the capability of that school and its teachers to engage in co- and self-regulated inquiry for improvement. The content of professional development initiatives and the context in which the professional learning is implemented were implicit in this theoretical model.

To date only a few researchers, in New Zealand at least, have returned to the schools to examine the nature of the achievement trends after the initial implementation period for the professional development. Timperley and Phillips (2003) added much to the New Zealand discourse around sustainability with their follow-up in seven schools involved in a literacy initiative within the Strengthening Education in Mangere and Otara (SEMO) project. Their research pointed to teacher discussion about student achievement data in relation to new practices being implemented as central to sustaining gains over time. Lai et al. (2009) more recently found that schools in their study were able to continue to sustain acceleration in student achievement, at a rate similar to the intervention, when they engaged in ongoing inquiry and established professional learning communities that promoted organisational learning. In a much wider international search of reforms, Timperley et al. (2007) found only seven core studies that met methodological criteria with demonstrated gains in student outcomes after external support had been largely or completely withdrawn. They warned that any conclusions they came to must be viewed as “conjectures” rather than “solid findings” (p. xxxiv).

This research will, therefore, have important implications for practice for those policymakers, in-service teacher educators, school leaders, and teachers who are concerned with the need to advance the pace of progress and then sustain these gains for those individual students and groups of students who are currently underachieving within schools. Understanding the complexity of sustainability, and the conditions that promote it, is therefore critical if school districts, state departments, or education ministries want to scale up those programmes that research has already shown to support improvement to many more sites. Wood (2007) contends that researchers have not yet adequately addressed the mechanisms and responses that could enable sustainability of shifts in
practices in the face of those challenges that present to schools, such as leadership attrition, funding changes, and shifts in the external policy environment. The literature base for sustainability is indeed embryonic.

The context of the research study

The research study is situated in the context of New Zealand’s Literacy Professional Development Project (LPDP) delivered to almost 300 schools in several cohorts over 2004–09. Approximately 20 national facilitators act as external leaders in these schools and they, in turn, are supported in their learning by a development team of eight educational consultants and two researchers. The LPDP’s theory of improvement was based on the hypothesis that effective classroom teaching will lead to improved student achievement and, in turn, that effective facilitation of reform will enable schools to sustain their gains over time (Bareta, English & O’Connell, 2006). Schools participated in the project for a period of two years, focusing on several improvement outcomes: in students’ literacy learning and achievement, in teachers’ content and pedagogical knowledge, and in transfer of this knowledge into classroom literacy practices. Project schools were also supported to focus on improving the ways that teachers learn together as professionals and on the effective leadership of these communities of practice. In their study of the scaling up of seven different programmes in five United States education departments, Datnow, Hubbard, and Mehan (2002) differentiate highly specified programmes from programme designs that operate according to “concept dissemination”. The LPDP would fall into the latter group, as it is guided by a set of principles around inquiry rather than by directive curriculum, professional learning, lesson plans, and/or school structures.

The LPDP has particularly focused its efforts on promoting sustainable change in schools. The intervention design drew on and is explicit about a number of conditions that research evidence indicates will prepare schools to sustain improvements well beyond initial implementation. These conditions were grouped in themes that included:

- leadership of effective professional learning and improved student outcomes (Robinson, 2007);
- deepened content and pedagogical knowledge (Elmore, 2005; Parr & Timperley, 2006);
- learning processes for teachers that penetrate teacher beliefs about learning (Spillane, Reiser & Reimer, 2002; Timperley et al., 2007);
• knowledge about inquiry and data literacy in order to test the efficacy of teacher practices (Copland, 2003; Earl & Katz, 2002, 2006; Reid, 2004; Robinson & Lai, 2006).

The LPDP theory of action for co- and self-regulated inquiry involved systematic reflection, including supporting leaders to build their knowledge of the nature of teaching practices that were being deployed across individual classrooms and the effectiveness of those practices. The LPDP facilitators initially developed professional learning communities in the school. Over time, school leaders were to take over this role. Evidence of this hand-over of ownership was expected well before the schools exited the project. At the beginning of the project, an extensive needs analysis was undertaken collaboratively with school leaders in order to align the learning needs of the students and the teachers to determine the inquiry that would be undertaken, the professional learning to be delivered, and how outcomes would be measured. The project’s focus on an inquiry “habit of mind” has required extensive support in schools for evaluative capability – describing the challenging issues, devising rigorous questions for the inquiry, and analysing and applying the results in a way that would impact on student outcomes (Earl & Katz, 2006).

The LPDP also focused attention on the conversations that teachers had about changing practice so that beliefs about teaching and learning could be surfaced and investigated for their efficacy in improving learning. Both leaders and teachers were targeted as learners in this project so that classroom changes could be supported and enhanced by organisational changes. There was an expectation that the changes in both classroom instruction and school-wide practices would take time and that schools would move through the improvement process at different paces (Bareta & English, 2007).

A key feature of the LPDP was the considerable resource allocated to facilitator learning. Since 2006, the project has had an additional contracted outcome: to provide evidence of effective facilitation (Bareta, English & O’Connell, 2006). This outcome was the result of previous research that indicated facilitators would also need to make significant shifts in their knowledge and understandings and to change their ways of working in schools if sustainable outcomes were to be achieved (Timperley, Parr & Higginson, 2003). Most facilitators have participated in empirical research on the effectiveness of their professional learning conversations with school leaders and teachers (Parr, Timperley, O’Connell, Armstrong & Nelson, 2007). This came to be
known as “practice analysis” and the term was extended to school leaders who took over observation and feedback conversations with their teachers (Bareta, English & Winthrop, 2008). The guiding principle for facilitators was that the ongoing needs analysis in schools must be integrally linked to the professional learning that they facilitated and that their ways of working must lead to sustainable practices in schools.

The LPDP did not develop a formal written definition of sustainability until 2007, but described indicators for sustainability in a project tool that was used by facilitators in collaboration with school leaders to evaluate the school’s progress over time. The tool was called the “Progress through the Phases” and described progressions in relation to the school-based outcomes of the project. It refers to sustainability as a “habitual inquiry into the effectiveness of literacy practices” and stated that these indicators needed to be present for schools to have “embedded” (or have made “habitual”) their practices across the outcomes of the project. In relation to the current definitions of sustainability sourced so far, these indicators offered a comprehensive list of school-based conditions for sustainability that embraced leadership for improvement, organisational routines that supported ongoing exploration of teacher content and pedagogical knowledge in relation to patterns of progress in students’ achievement in literacy, and the monitoring of any changes to teacher and leadership practices for their impact on student achievement (Bareta & English, 2006). Reportedly, facilitators and leaders used this document at regular intervals with school leaders to assess progress over the two-year period of the LDPD and to prompt discussion about what would need attention when the school exited the project (Bareta & English, 2006).

The embedded research programme that supports the LPDP has focused on leadership of professional learning, on content and pedagogical knowledge, on learning processes for teachers, and on knowledge about inquiry, and the results have been used to inform ongoing improvements. However, as in most other projects, sustainability was still an “article of faith”. This research was also designed to provide empirical evidence for the LPDP team by means of returning to schools that had exited the LPDP to track their progress. One objective of the research was to test the project’s theory of improvement and the impact it may or may not have had on continued improvement in teacher practices, on deepened content knowledge, on leadership of professional learning communities and, more critically, on continued gains for students’ literacy achievement.
The first cohort of schools started in February 2004 and finished in December 2005, providing an opportunity for this research study to track schools over time to assess the impact of the LPDP in achieving long-term improvement. New cohorts of schools joined the LPDP in 2006 and again in 2008. The gains reported for the first cohort of schools in the LPDP were substantial. The first cohort of schools, providing two data points ($N = 91$), completed two years (2004 and 2005). The average effect size gain (that is, relative to where they started) for schools that chose to focus on writing (data from a moderated sample of 1,064 students) was 1.28 and for reading (data from 3,787 students) was 0.87. Moreover, for those students in the lowest twenty percent at Time 1, the gain has been more marked. In writing, the effect size gain was 2.05, and in reading, 1.97 (Bareta & English, 2006). An independent review of these data confirmed these results (McDowall et al., 2007). Similar large gains are reported in the second cohort of the LPDP (Bareta et al., 2008). When assessing the impact of interventions on student achievement it is important that any identified gains are compared to expected gains related to progress of normative samples. The expected effect size gain for the particular tool used to assess writing is 0.4 over two years (The University of Auckland, 2004). In terms of reading, the expected effect size gain for the reading assessment tool is 0. This is because the assessment tool uses stanines to describe students’ achievement levels. The expectation is that students remain at the same stanine level over time if they continue at the same rate of progress during the year (NZCER, 2001).

The two interdependent studies within the thesis research design aimed to provide empirical evidence related to the problematic notion of sustainability and add a conceptual framework to the debate at research and policy level. This research investigated understandings of sustainability promoted by the LPDP, its meaning, and its enactment in the schools over time. It aimed to challenge current thinking about sustainability (both within and beyond the school) as being self-limiting in nature, reinforcing sustainability as linked to a maintenance model rather than to an ongoing improvement model.

A sampling method was applied to select 16 schools from the first cohort of the LPDP, and data gathered from them at various points after the formalised external supports ended. The mixed methods research design examined if and how these schools maintained or transformed the key shifts that they made while in the programme, tracking shifts in students’ literacy achievement and teacher and leadership practices. A
purposive sample of four schools tracked over three years was then analysed to inquire into the relationship between school-based conditions for sustainability and evidence of sustained improvement in achievement data.

*Organisation of the thesis*

This chapter introduced the key themes related to sustainability of educational reform, its problematic position within current educational research, the consequential importance of this thesis topic, and the questions that the thesis addresses. The chapter also described the general context and parameters for the research.

Chapter 2 reviews the current research on the sustainability of educational professional learning programmes and from this offers a conceptual framework for sustainability that includes two school-based dimensions developed from this literature. These dimensions are co- and self-regulated inquiry practices and coherence of effective instructional practices. These dimensions were used to shape the design of the research instruments.

The thesis is then reported as two interdependent studies. Study 1 investigated the achievement data, perspectives, and actions in relation to sustainability within 16 schools in the two years following their exit from the LPDP. Chapter 3 describes the methodology for Study 1, the research tools, and the limitations of these tools.

Chapter 4 begins the analysis of the findings related to how schools in the study perceived sustainability of the LPDP. What did schools understand by the term sustainability, and what did they expect to see in the classrooms and school-wide if the project were being sustained? Was their view of sustainability a broad one, beyond that of a single project? Did school leaders and their teachers expect that separate professional learning projects had to be sustained in different ways, or did they recognise how different projects are bound by similar key ideas about effectiveness?

Chapter 5 then grapples with the problem of defining sustainability in relation to student achievement gains. In the research literature, there is a lack of longitudinal achievement data for successive cohorts of students, other than analysis of examination results over time for schools and/or districts. When accelerated achievement has been made for students during a professional learning intervention, there are few analyses to call on that compare subsequent student performance with those earlier gains. Therefore, the 16 schools in this study have two data sets over two-year periods for two consecutive
cohorts of students. Despite the relatively short-term nature of these data, they do provide a New Zealand context for comparison of students’ literacy performance between the implementation phase and subsequent years. The chapter compares and contrasts these data to results from one other New Zealand study that also focused on sustainability of a literacy intervention.

Chapter 6 then discusses actions that the 16 schools took over 2006–07 and relates these to their student achievement results in literacy. The qualitative data from interviews, classroom observations, document and meeting analyses are mapped onto the two dimensions of the conceptual framework offered in chapter 2.

Study 2 of the research begins in chapters 7 and 8. Chapter 7 outlines the methodology for investigating four cases from the original sample that had various outcomes for student achievement after one year and exhibited different characteristics of the dimensions discussed in Study 1. Chapter 8 examines each of the four schools in some depth, in relation to co- and self-regulated inquiry practices at school-wide and classroom levels and to ways of developing coherence within the school, in particular, whether schools explicitly make connections between learning from different projects in order to amplify and sustain effective practices. Specific attention is paid to mediating and outcome variables, to any disconfirming evidence found in each case school and to the implications these data may have on the specific elements of the two dimensions offered in the theoretical framework.

Finally, chapter 9 combines the results of both studies and offers theoretical conclusions and implications for the field and areas for further research.
CHAPTER 2
LITERATURE REVIEW: REFRAMING SUSTAINABILITY

The research literature on sustainability is neither abundant nor convergent in its conclusions. This chapter begins by examining findings of the relatively few studies that have specifically focused on sustainability of professional learning in order to describe how this notion has been defined and measured to date. It then explores a broader range of research themes to determine the school-based and systems-level conditions that are important in achieving sustained improvements in schools. The final section then reframes these conditions for sustainability into a conceptual model that forms the hypothesis to be examined in this study. The resulting model challenges the notion of sustainability as being described in the context of a single intervention only; instead, the model expands the scope of the notion to include other layers in the system and establishes a definition and theory of practice for achieving sustainable schooling improvement.

What is the unit for sustainability?

Levin (2008) contends that there is a “disjuncture” in the literature between the discussion of sustainability as applying to programmes or interventions, and sustainability as applying to the school or school system. Indeed, researchers have focused on the school as the unit for inquiry into sustainability, whether that is in terms of sustained shifts in teacher practices and/or student outcomes (Century & Levy, 2002; Lai et al., 2009; Timperley et al., 2003). There are gaps in this context as well around how sustainability might relate to individual teachers or, more importantly, their students. Typically, researchers undertake an investigation into one intervention across a number of schools, using their findings to describe features of sustainability (Century & Levy, 2002; Huberman & Miles, 1984; Knight, 2005; Lai et al., 2009). Later research into sustainability broadens the lens of inquiry to include the role of school district administrators in supporting schools to sustain their practices and/or student outcomes over time (Datnow, 2005; Levin et al., 2008; Wood, 2007).

A new construct is emerging in the research literature around schooling improvement. Schools in New Zealand and other jurisdictions are increasingly clustering together to help solve educational problems. (This is different from schools implementing a predesigned initiative in a school district.) The clustering may be
geographical entities, online collaborations, or communities of interest around a particular problem of learning and achievement. Clusters of schools analyse their achievement data collaboratively, design their initiatives together, and share strengths, resources, and ownership of the improvement agenda (Annan, 2007; Howes & Ainscow, 2006). The jury is still out on the effectiveness of such endeavours in terms of long-term sustainability, but there are promising indicators. Studies in the United Kingdom suggest that such collaboration is a necessary, but not sufficient, condition for successful large-scale intervention (Earl, Watson & Katz, 2004; Bell, Cordingley & Mitchell, 2006; Howes & Ainscow, 2006; NCSL, 2006). Networks may indeed support sustained ongoing inquiries (Kaser & Halbert, 2006), but not address large-scale and ongoing improvement. These researchers add, though, that the effectiveness of clustering or networking lies not within the structural collaboration itself, but is achieved through the presence of incentives, skilful leadership, shared improvement agendas, and credible external advisers who can provide sufficient challenge and critique to existing or new practices and their effectiveness. Certainly the potential for increasing the likelihood of sustainability lies within the ability of a network of schools to offset micro-level distractions as well as to absorb and align macro-level policy changes. In New Zealand, a particular feature of these clusters is their partnership with the Schools Monitoring and Support division of the Ministry of Education and, less frequently, with researchers who can monitor and evaluate their collaborative efforts (Annan, 2007).

Others have viewed sustainability at the system level, arguing that sustainability is a shared responsibility (Elmore, 2005; Fullan, 2006; Fullan & Sharratt, 2007; Levin, 2008), though often with emphasis on actions to avoid such as educational change being bound in compliance rather than promoting capability for improvement in the system (Hargreaves, 2002; 2003). Levin (2008) views sustainability as the combination of three equally important elements – improvement of student outcomes, support for educators and strategically targeting public acceptance. His paper tracks the results of literacy and numeracy innovations since 2005 in schools across Ontario province and the principles by which they were implemented. The number of schools with very low performance had dropped by seventy-five percent, and overall ten percent more students were reaching the provincial standards. By paying attention to all elements of schooling simultaneously – building morale and capability for teachers and leaders, involving parents, changing policies, and adding resources – this track to improved outcomes, he argues, “augurs well” for sustainability, but at this time remains untested.
In New Zealand, Timperley and Parr (2007) also widen the lens to examine the relationship between various levels of the system involved in sustaining a schooling improvement initiative, specifically the LPDP. They describe how evidence-informed inquiry operates at all levels of this intervention and that these layers become “permeable” via deliberately created structures to overlap personnel involved in the national initiative. In this way, reflection and feedback have a two-way flow that, more importantly, allows for a direct pathway from the classroom focus on students and teachers to facilitators’ learning, to project leaders, and to policy makers. This view of sustainable change matches Fullan’s (2006) “tri-level reform” in which governments, school districts, and schools work together on common approaches and strategies.

Rarely have researchers looked closely at sustainability in terms of more than one intervention, at the intersection points as one project melds with the previous one, or when new initiatives take over. This perspective is mostly used to explain how multiple interventions and ever-changing reform agendas inhibit the likelihood of sustainability (Datnow, 2005; Hargreaves, 2003; Honig & Hatch, 2004). A case, therefore, emerges from these findings to broaden the theory of practice for sustainability to include networks of schools, district administration, and policy levels, as well as successive interventions.

*Time and context bound*

Sustainability is a comparatively recent theme in educational research literature in both international studies and those studies undertaken in the New Zealand context. Early findings (pre-1995) about sustainability were situated in and, in a sense, bounded by, the context of how the initial professional development programmes were implemented at that time (Gersten, Chard, & Baker, 2000). Most often such programmes were one-day workshops delivered by external experts that rarely led to long-term use of new practices (Berman & McLaughlin, 1977). However, these early findings were influential in bringing about significant improvements to later interventions and are worthy of consideration before examining how sustainability has been defined, described, and measured since 2000.

Studies by Berman and McLaughlin (1977), as part of the RAND Change Study in the United States, form some of earliest analyses of factors that lead to sustainability resulting from professional development programmes. In the first instance, they found
that the amount of resource used to initiate a programme bore little resemblance to the success of the programme long-term. Perhaps the most salient of their findings was that practices promoted in professional development that helped teachers succeed with the low-achieving students did, in fact, last over time. They also reported that professional development programmes, that took into account the complex local contexts, sustained their changes more than any externally imposed policy changes were able to (Berman & McLaughlin, 1977). It also became evident early in the research literature about sustainability that no one “solution” would last in large numbers of diverse schools and communities, and that programme developers needed to consider the uniqueness of schools and their students in their design of professional learning.

Active learning for teachers within the process for change was also explored as a factor that would deepen thinking about changes and so impact on sustainability. McLaughlin found that “outside expertise rarely led to long-term use of innovative practices” (in Gersten, Chard & Baker, 2000, p. 4). This finding must be viewed in the context of professional development programmes for this period that were typically a “one size fits all” approach. Such conclusions then led researchers to consider the roles and influence of leaders, in-school coaches and teacher membership of professional communities in leading and sustaining change (Huberman & Miles, 1984; Little, 1993).

Towards a definition of sustainability

Since 2000, researchers have more consistently attended to defining and describing sustainability and drawing conclusions about the links backward to the professional learning design. The term evolved out of economic theory and is now used widely in the literature about educational change, though it still lacks a precise definition or shared meaning. For example, Coburn (2003) argues that scaling up and sustainability of professional development are “undertheorised.”

[Sustainability] rarely appears in theoretical and empirical pieces. . . Of the 44 publications on efforts to scale up external reforms reviewed for this article, only 18 publications involved investigations of schools that had been involved in the reform for four or more years. And only one explicitly looked at schools for which an implementation period with additional resources and attention had officially ended. Instead, most studies focus on schools in their first few years
implementing a new external reform, failing, in our view, to capture sustainability. (Coburn, 2003, p. 6)

Wood (2007) also suggests that the sustainability is yet to acquire a shared meaning within educational contexts.

There are still relatively few research studies that offer a clear definition of sustainability, or trace how or why schools discontinue, adapt, or even transform the changes they make over time. Typically, sustainability is defined either in terms of shifts in school and teacher practices or, less commonly, in terms of ongoing improvement in student achievement outcomes. The unit of measurement is most often at system, school-based, or professional learning programme level (Coburn, 2003; Lai et al., 2009; Levin, 2008). Only a few commentators have established how an individual teacher might approach or secure sustainability of their own learning (Franke, Carpenter, Fennema, Ansell & Behrend, 1998; Spillane, 1998; Spillane, Reiser & Reimer, 2002). Therefore, measuring whether these practices have been sustained or not remains a complex issue. A focus on school-based interventions (linked to a larger programme) and individual teachers at least allows the context to be foregrounded in the mix of factors that may impact on sustainability. The documenting of school-based examples of how sustainability can be organised, developed, and pursued beyond the intervention are rare, at least in New Zealand. There are other layers to consider in terms of sustainability of outcomes. For instance, an evaluation of the LPDP found that individual external facilitators working with schools leaders and teachers in a school had an effect on outcomes for student achievement (McDowall, Cameron, Dingle, Gilmore & MacGibbon, 2007) although the link between the impact of coaching and student achievement is difficult to ascertain (Parr & Hawe, 2009). Therefore, the design of professional learning and the leadership of those who promote improvement in schools are important factors in supporting the capability of schools to sustain the changes they make while part of a particular programme.

The research traditions of schooling improvement and school effectiveness have undoubtedly led to similar debates and differences about the weighting of particular features of schooling and their impact on sustainability. Whether focusing on the school as the unit of sustainability, or integrating this with a wider view of the system and its input in defining the term, there is some agreement that a rich variety of pathways have been followed (West, Ainscow & Stanford, 2006).
Shifts in school and teacher practices

Sustainability has typically been examined with a lens on changes in teacher practices. More recent research has argued that the impact of shifts in teacher practices on student learning and achievement is critical to research on sustainability (Timperley, 2003). However, St John (2002) is critical of applying a “narrow focus” on just student data to define sustainability, arguing that this lessens the impact of any investigation into the valuable content knowledge and practice gains made by teachers. Fullan and Hargreaves (1996) support this view, arguing that enacted teacher practices, school processes for collaborative learning, and leadership practices must be included in any definition of sustainability. Evaluating if these practices have been sustained is also multifaceted. For instance, how much change to these practices and processes, either in the classroom or school-wide, is acceptable? How much “fidelity” should there be to the original reform?

Researchers working in the context of numeracy professional learning in New Zealand and Australia describe sustainability as the consolidation and maintenance of gains made within the project since its initial implementation (Knight, 2005). Knight challenges this definition as needing more precision, arguing that the gains are not clearly enough defined. These numeracy projects have addressed the emerging issues of sustainability by continuing the support from external agents and by targeting new teachers and leaders, presumably to fend off the challenge that turnover of teachers brings to schools. Knight refers to this continued support as a “traditional fixation” that also needs to be re-examined for its impact. She re-conceptualises sustainability as not being about a project but “about building on change and developing internal capacity for continued professional learning” (Knight, 2005, p. 473).

Adaptation versus fidelity

There appears to be no agreement about whether professional learning programmes that specify particular practices for teachers are more successful in terms of lasting improvements than those that are designed around a set of broadly defined principles. Some educational interventions have attempted to eliminate adaptations of their programme model by codifying and standardising practices with planning tools, processes and curricula materials that were designed by experts for teachers to implement. Elmore (1996) reports that some of these efforts merely “shoe-horned” new practices into old practices and that any changes that resulted were diluted and hybrid in
nature. However, Annan (2007) argues that standardised practices in programmes such as Success for All in the United States were successful where the instructional core of teaching needed considerable improvement.

Adaptation of newly learned practices can occur from the outset as teachers take up new ideas or try to decode policy (Spillane, 1998, 2004; Spillane et al., 2002). For instance, Huberman and Miles (1984) were interested in the modifications that teachers made in 12 school sites to various school professional learning initiatives. After three years, as many as 10 sites had dropped or modified the key components of their reform programmes to produce “unacceptable” versions of the original. Spillane et al. (2002) and Coburn (2001) argue that the complexity of the sense-making process between policy and practices is often ignored as researchers and policy makers try to explain the failure of many teachers to implement new practices as intended. These researchers contend that teachers’ active comprehension of reform combines with the richness of their individual experience, knowledge, and beliefs and that typically teachers encode new information by adapting it to fit what is known, or they encode it without exploring the implications of the new ideas in relation to what they already know, resulting in pockets of inconsistent knowledge. The usual approach to processing knowledge is a “conserving approach,” preserving existing frameworks rather than radically transforming them (Spillane, 2004; Spillane et al., 2002). Timperley et al. (2007) also indicate that this approach leads to important differences in the adaptations made by teachers. Teachers see the professional development through the lens of their current practices, and the understandings they construct may fail to reflect the fundamental shifts intended in the programme. That is, their tacit models are not challenged and may even conflict with the intentions of the professional learning programme (Timperley et al., 2007).

Furthermore, where change in teachers’ practices is largely self-reported in research on the impact of professional learning programmes, it is not clear if teachers have actually changed their beliefs or, in fact, held these beliefs prior to the programme (Spillane et al., 2002; Timperley et al., 2007). They can be misled by superficial similarities that they perceive between their own ideas and the programme ideas, which is sometimes referred to as over-assimilation in the literature (Spillane, 2004; Darling-Hammond & Bransford, 2005). Values and identity also play an important part in the sense-making that teachers engage in during their professional learning. The substantive
changes required in many programmes often involve core behaviours that are central to each teacher's self-image, so influencing personal motivation (Spillane et al. 2002). Geijsel and Meijers (2005) contend that professional learning programmes can threaten a teacher’s self-esteem, further reducing the likelihood of any lasting change in practice. The problem, therefore, with a single focus on practices is how much fidelity to the practice must occur if the programmes are to be judged as having been sustained and who should be the judge of that level of fidelity, other than teachers.

Century and Levy (2002) offer an often-cited definition of sustainability that addresses the notion of adaptation, at the same time embracing the institutionalisation of key changes, and resistance to any ideas that might derail the reform. They explain 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” (p. x).

Their definition resulted from a three-year research study into nine school districts in the United States that received National Science Foundation funding to change science programmes from being driven by textbooks to being centred on using hands-on materials with students. They examined factors that contributed to or inhibited the sustainability of district-wide, hands-on inquiry science programmes, with a specific focus on system contexts and external conditions that had an impact on lasting change. Century and Levy further distinguish their definition of sustainability from that of “programme maintenance,” where what is embedded into everyday practice is just a replica of what was introduced in the programme and has not been adapted or improved. They contend:

[that a] programme must be maintained before it can be sustained, that it cannot be stalled at maintenance, but must develop an ability to evolve and adapt. The adaptations, however, must still be guided by the core beliefs of the programme intent. (p. xi)

Reform models with flexibility to adapt to local circumstances were found to be more sustainable in the face of changing district and state constraints than those that required greater resources and could be adapted to multilingual and multicultural contexts (Datnow et al., 2002). Dede and Rockman (2007) comment further that “leverage” of successful initiatives may be the paradigm to consider when designing systemic change that is sustainable, rather than being “trapped” by the idea of scaling up
such programmes by replication. Cordingley and Bell (2007) also argue that the greater the scale, the greater the need for flexibility and adaptation to the school context. They support prescription in the first instance so that new practices are genuinely applied and understood before adaptation takes place. They believe teachers need to see what the practice looks like and what impact it has on students’ responses so that they can understand what can be adapted and what might be critical to be retained.

Expanding on this view of scaling up and adaptation of school reforms, Thompson and Wiliam (2007) make a case for rigour without rigidity. They develop a theory of scalability around their work on Keeping Learning on Track (KLT), a set of formative assessment interventions undertaken in 28 school districts across six US states. The impact of formative assessment teaching strategies on student achievement has been described by Thompson and Wiliam as yielding improvements of between 0.3 and 0.7 standard deviations, but these authors also argue that the cost–benefit ratio is several times greater than for other interventions. Their view is that there are “tight” components of KLT that require “obsessive adherence,” while the “loose” components are the contextual elements such as structures, communities, and political policies. While sustainability is not their primary argument or evidence base, their theory of action for scalability of reform integrates sustainability with a set of core design principles (predicated on formative assessment strategies) that can be used adaptively in schools to meet the needs of their students. They argue that teachers do need to understand the why as well as the how so that they can be adaptive in their responses. School leaders also need to understand the theoretical design of the intervention so that they can accommodate the core principles into their diverse contexts (Spillane, 2004). Thompson and Wiliam (2007) warn that the loose elements require alignment and the systemic influences need to be “swept in” and be adjusted to the school context. Rather than starting with whole system reform, they contend that focused interventions can more usefully “shine a light” on the environmental and contextual problems that need to be attended to before the original intervention can work. They also emphasise the critical nature of effective professional learning communities and the role of external expertise to support sustainable outcomes. As the KLT interventions developed, they identified a problem of “bootstrapping expertise” (the finite numbers of qualified professional developers), which was offset by developing more specificity in the written content provided for the schools and more support for school leaders of practice communities. They acknowledge that these actions only complemented the learning rather than
replacing the roles of external experts but, at least, the actions supported the scalability of the intervention design. Finally, Thompson and William background content pedagogical knowledge in their description of this programme and concede that they do not have a definitive answer to the issues of deficiencies in teachers’ subject matter knowledge being addressed with their learning of more generalised strategies. They recognise that deepened content knowledge leads to improved teacher questioning, better understanding of students’ progressions in learning, and more specific feedback, the latter among the most valued of interventions in terms of impact on student achievement (Hattie, 1999). However, they argue that their results imply that “simply improving teachers’ pedagogy works to boost student learning, even in the absence of deep content knowledge on the part of the teacher” (p. 18). Interestingly, for the particular context of sustainability, they add that further research is needed to determine whether or not this effect might plateau and whether further gains might well be achieved by subsequently focusing on improving teachers’ content knowledge.

So, adaptation is a given in the context of sustainability of professional learning, but the constraints of that adaptation are less defined in this literature. Adaptation of new practices may be superficial and not represent the core beliefs and values of a reform and, therefore, not survive or improve student achievement in the long term (Spillane, 1998). Timperley and Phillips (2003) studied seven suburban, low-decile\(^1\) schools in South Auckland, New Zealand, that participated in a professional development programme focused on literacy. They identified that those teachers who most adapted the literacy teaching approaches promoted in the programme had lower student achievement outcomes than those who implemented the approach as intended. The teachers who made these adaptations believed that the changes better met the needs of their students, but they had not tested them for their effectiveness in improving outcomes. Brown and Campione (cited in Wood, 2007) also discovered “lethal mutations” of practices after a programme had ended, that worked to negate the original objectives of the professional learning.

Some interventions try to limit teacher adaptation of practices by tightly prescribing teaching sequences and lesson structure. Stallings & Krasavage (1986)

---

\(^1\) A school’s decile in New Zealand indicates the extent to which the school draws its students from low socioeconomic communities. Decile 1 schools are the 10 percent of schools with the highest proportion of students from low socioeconomic communities. A school’s decile does not indicate the overall socioeconomic mix of the school.
investigated the Madeline Hunter model of schooling improvement that had a “relatively prescriptive” approach to implementation of the content and practices for teachers. There were substantive gains in student achievement in reading and mathematics while teachers were supported by external expertise, but when the responsibility for maintaining teacher practices shifted to school leaders then the classroom observations were reduced and implementation was highly variable and student achievement scores were reduced.

Therefore, two school-based conditions appear to support ongoing adaptation of practices that lead to sustained student achievement outcomes. Firstly, any adaptations of new practices need to be guided by the principles that underpin the professional learning. That is, teachers need to be able to transfer their learning to new contexts (Coburn, 2001; 2003). Secondly, teachers need to be equipped with the skills to notice and evaluate the impact of their teaching so that any adaptations that they make still support improved student achievement (Lai et al., 2009; Timperley et al., 2007).

System influences

There is substantial research literature about the primacy of the educational context that these schools operate within as a key factor in determining long-lasting change. Sustainability is described as a “two-way or multi-way street” where future directions are continually challenged and tested by leaders at both school and system levels (Fullan & Sharratt, 2007). Hargreaves (2002) supports the importance of system responsibility for sustainability, outlining interrelated characteristics that should be promoted by sector agencies. These characteristics include: improvements that sustain learning for students and not merely changes to school structures; improvements that can be supported by available and achievable resources; improvements that do not impact negatively on the surrounding environment for schools, and finally those improvements that promote ecological diversity and capacity in that environment. He writes of the social “geographies” in the educational landscape that together must be considered in all attempts to sustain new reforms, more particularly the limitations to sustainability that result from different approaches being promoted by various governments around the world.

In slightly later work than that of Hargreaves, Datnow (2005) also widens the notion of a single school-based context when framing a definition for sustainability of
interventions. She examined six comprehensive school reform\(^2\) (CSR) models in 13 schools in one US school district in the period three years after their initiation. Only five of the 13 schools were still continuing to implement their reform designs with moderate to high levels of intensity. Datnow argues that sustainability of the CSR models was affected by the changing district and state contexts and that the schools’ ability to be resilient was dependent on their strategy for dealing with the changes, their experience with the reforms, and their own capacity. Her definition of sustainability, like those before it, encompasses the notion of institutionalisation, a multilevel process of embedding an innovation within the structure and norms of the school. Datnow describes structural, procedural and cultural norms as being part of this institutionalisation. However, she cites another study that found that schools sustained reforms over eight years or more when there was overarching political support and an alignment of the “cultural logic” to the innovation (Yonezawa & Stringfield, 2000, cited in Datnow, 2005).

Datnow’s (2005) research aimed to uncover these external conditions impacting on sustainability, more particularly schools’ interactions with district- and state-level policy makers and how schools operating in the same policy contexts make strategic decisions that enable them to be more effective at sustaining their shifts in teaching and learning than others. She found that reform models lasted when they helped meet policy and accountability demands or, at least, when they did not come into conflict with them. Resilience in the face of competing priorities was difficult to achieve for most of the schools in her study. Therefore, the notion of system- and school-wide coherence is taken up later in the chapter as a condition that promotes sustainability.

**Time frames**

There appears to be no agreement in the literature about time frames for sustainability. What are the expected progressions as schools move on from externally supported professional learning? Century and Levy (2002) describe three phases that advance a programme from implementation to sustainability – establishment, maturation, and evolution. The establishment phase focuses on the concrete elements of the programme, making sure they are well in place, accepted, efficient and found both across

---

\(^2\) CSR models refer to whole-school reform designs in the United States where schools and districts purchase programmes from design teams often associated with universities, non-profit organisations, or companies. Federal funding for the use of these models was first made available in 1998, and then changes to Title I regulations for school funding accelerated their use in schools all over the country.
the school and district-wide. At maturation the changes are habitual, both for school-wide processes and for classroom teaching and student learning. The hallmarks of the evolution phase are that there is growth and improvement. Sustainability, in their view, requires multiple strategies – at the classroom level, school level, and district level.

With respect to time frames, this review of literature on sustainability did not establish how much time should elapse before it is reasonable to observe if a programme has been sustained. Century and Levy’s (2002) study examined schools that had programmes in place for between 10 and 30 years. In contrast, New Zealand’s Numeracy Development Project focus on sustainability examined student achievement data only one to three years after the implementation phase (Thomas & Tagg, 2005). Another potential tension arises in the New Zealand context, where schools undertake a one to two year “contract” with professional development providers and then must turn to the next opportunity in a new curriculum area. This provision model reinforces the sense that sustainability means “going it alone,” with no sense that each new professional development contract may be able to reinforce and strengthen key ideas about effective practice already learned in previous projects. Even if an ongoing inquiry process is promoted in the intervention, schools may have experienced only one cycle of the inquiry (Timperley et al., 2007). The danger here is that leaders and teachers may not continue to deepen their understandings of how to support those students who are still not progressing in one area of the curriculum before they move onto the next professional learning project. Where the context for the professional learning is in schools that are serving linguistically diverse communities in low socioeconomic areas, this time frame may not be enough to ensure accelerated rates of progress required for students to overcome disparities in achievement. Teachers in these schools face educational issues that are complex and they may still require ongoing external expertise to support then in successive inquiries (Lai et al., 2009). In New Zealand, early schooling improvement initiatives set out to support schools for a period of three to four years, but this time frame proved too short a period to achieve sustained outcomes. Newer initiatives are framed around a model of support for a period of three years and then an additional three years if complex achievement problems persist.

---

3 Information supplied by Early Childhood and Regional Education Division, New Zealand Ministry of Education, 6 October, 2009.
Not surprisingly, Century and Levy (2002) found that none of the programmes that they examined was an exact replica of the earliest years and that the longer the time span the more clearly the trends to evolution emerged. It is worth noticing that implicit within Century and Levy’s definition, together with those of most other commentators reviewed here, is the view that sustainability can only be described in the context of each programme and its core beliefs and values. However, the argument proposed in this research is that this definition limits educators’ views of sustainability and, in fact, may create an additional barrier to sustainability. Century and Levy’s definition continues the notion that each professional learning project is an entity in itself, bounded and separated by the content or curriculum that it focuses on, rather than emphasising, in some part, its alignment to effective teaching and learning practices and to system-wide policy.

**Student outcomes**

The critical issue related to sustainability and any adaptation of practices over time is whether these adaptations lead to improvement in student outcomes. Timperley and Phillips (2003) found that those schools that had sustained improved student outcomes used student achievement data as a “touchstone” for measuring the effectiveness of their teaching methods as they moved on from their professional learning programme. This study was a watershed in the research about sustainability in the New Zealand context and its messages have been further endorsed in the findings of *Teacher Professional Development: Best Evidence Synthesis Iteration* (Timperley et al., 2007). The fundamental criteria used to determine the inclusion of studies in this synthesis was evidence of “substantive student outcomes associated with teacher professional learning” (p. xi). Ninety-six studies were located in their search of national and international databases, but only seven studies demonstrated ongoing improvement of student outcomes after the programme funding and support had been reduced. The evidence from these seven studies was limited but consistent. Where sustainability was achieved, the providers of the professional learning had focused on developing pedagogical content knowledge in enough depth for teachers to make principled decisions about practice, as well as supporting teachers to develop evidence-based inquiry skills that allowed them to test the impact of new practices. What is interesting is that both prescriptive and more flexible approaches for professional learning were represented in the studies that documented ongoing improvement. Two interventions that were both sustained and scaled up were more prescriptive about lesson structure and some elements of content.
The LPDP would not be considered prescriptive in its delivery, so this adds yet another element of complexity to this research.

In recent years, there is a more clearly defined position and agreement among researchers on judging sustainability in terms of continuous improvement of student achievement, whether in particular school-based inventions or in large-scale system contexts (Fullan & Sharratt, 2007; Lai et al., 2009; Levin, 2008; Timperley et al., 2007). More critically, though, the research literature yields few analyses of patterns that might be expected for student achievement data in the years following an intervention. If students do make gains in the initial period, will these same improved trajectories be maintained, or even improved, for new groups of students once the schools exit the training programme? For example, one study of an English school district revealed that only one in 16 schools managed to improve examination results continuously for more than four years over the 10 years of tracking examination data sets (Thomas, Peng & Gray, 2007). Lai et al. (2009) found that after a three-year intervention in reading comprehension in seven Decile 1 schools in New Zealand, students continued to accelerate their learning at a similar rate as they had during the intervention (four months in addition to expected progress) in the following three years. These researchers had identified significant plateau and slumps in achievement over the holiday period whilst schools participated in the intervention and these continued to occur in the follow up study. There were differences among the schools in the cluster, with some making higher gains during the year. This study is more fully described in chapter 5, where sustainability is investigated in relation to student achievement gains.

Longevity is implicit within the notion of sustainability. If one accepts that student achievement data are fundamental to evaluating whether an intervention has been sustained or not, then a longitudinal lens over these data is a key feature to include in a school-based study such as this. Will schools follow an iterative inquiry model over time, where they ask more searching questions about their own practices and knowledge to help solve the issues at each cycle of data analysis or will they continue to analyse each data set as they did the year before? One can extrapolate from Teacher Professional Learning and Development: Best Evidence Synthesis Iteration (Timperley et al., 2007) that privileging student achievement as a measure of sustainability may sharpen the urgency and improve the responses that schools make to ongoing professional learning opportunities and the stance they take to adapting practices.
Indeed, identifying the evaluative capability being applied by schools to ongoing sets of achievement data is a key feature of this research. The term evaluative capability is derived from the literature that focuses on inquiry as a mechanism for ongoing improvement and is expanded on in the next section.

School-based conditions that promote sustainability

The research literature offers considerably more guidance on the challenges that schools face in trying to retain the shifts that they make. These constraints include the impact of changing personnel, shifting priorities, and internal competition for funding. They have most often been studied in the context of the early years of school participation in professional development (Century & Levy, 2002; Coburn, 2003; Datnow, 2005; Fink & Brayman, 2006; Gersten, Chard, & Baker, 2000; Hargreaves, 2003; Jasinski, 2007).

Researchers promote a plethora of conditions that they argue contribute to the sustainability of reform at school level. Substantive research lies behind the promotion of each of these conditions. The issue for educators is to work out how to balance and apply these conditions to professional learning design, to identify which is more influential in supporting sustainable changes in schools, while acknowledging that context matters in all engagements with teachers. Therefore, these school-based conditions for sustainability have been grouped into five themes in the following section.

Leadership

Hargreaves and Fink (2006) link their findings about sustainability to leadership in a study of change in eight US and Canadian high schools over three decades. They describe sustainability in terms of being long lasting and having “meaningful improvements in learning”, suggesting that a key force in this process is leadership. They outline several principles of sustainable leadership that include: planning for succession, with the grooming of new leaders in schools; distributed leadership for improvement within the school; sharing of the responsibility of improvement beyond the school, with leaders supporting other schools and the district as a whole to change. Leaders, they argue, enable people to adapt and prosper in their increasingly complex environments by learning from one another’s diverse practices. Their research found standardisation to be “the enemy of sustainability”: schools with leaders that brought about long-term change adapted programmes from the beginning. They refer to leader “resilience” as critical to
fending off actively those programmes or policies that work against what is already supporting improvement in their schools.

Spillane’s investigations (2006) into distributed leadership as a mechanism or condition for sustainability are also relevant here. Leadership for reform that can be “stretched over” a school can support the “critical mass” indicator, which Century and Levy (2002) refer to in their study. Critical mass is meaningful for sustainability when adequate numbers of teachers can access leadership and support for change and can continue their professional conversations in both formal and informal learning situations. In a New Zealand study of the sustainability of an ICT initiative in a large urban secondary school, the lead teachers (roles that were established in the implementation phase) were found to be the key to sustaining the changes in the school as long as they were supported by senior management and had allocated time to work with their colleagues (Okey, 2006).

Leadership for sustainability is also about knowledge of change processes (Copland, 2003). Waters, Marzano, and McNulty (2003) discuss the concept of the “order” or magnitude of change in schools, as a result of their meta-analysis of 70 studies of leadership that had demonstrated improved student achievement. These authors describe second order change as occurring when it is not obvious how the change makes things better, when it requires individuals or groups to learn new approaches, or when it conflicts with prevailing values and norms. Second order changes require leaders to work far more intensively with staff and the community. These researchers argue that second order change may disrupt co-operation, a sense of wellbeing, and cohesion in a school. It may confront and challenge expertise and competencies and throw people into states of “conscious incompetence.” They argue that this knowledge, coupled with skills and strategies for leading second order change, is a prerequisite for effectiveness and presumably sustainability.

Finally, Robinson (2007) identifies five dimensions of leadership that are associated with impact on student achievement in her analysis of 11 studies that sufficiently quantified the relationship between types of school leadership and a range of social and academic student outcomes. These dimensions include: establishing goals and expectations; strategic resourcing; planning, co-ordinating, and evaluating teaching and the curriculum; promoting and participating in teacher learning and development; and ensuring an orderly and supportive environment. Robinson’s major message is that the
closer leaders get to the core business of teaching and learning, the more impact they have on their students.

Leaders are, therefore, central to promoting particular school-based conditions that link to sustainability by aligning messages about change, ensuring the coherence of multiple professional development initiatives being undertaken in a school and accessing external agencies that promote and support improvement.

Content and pedagogical knowledge

Coburn (2003) promotes depth of professional learning as a key dimension to achieving scale and sustainability. By this she means that teachers need to have a particular type of content knowledge. They need to know their subject in a way that supports others to learn. This dimension, termed content pedagogical knowledge, encompasses knowledge of the subject, knowledge of how to teach, and knowledge of the learner. Professional learning programmes have to support teachers to build principled knowledge, that is, knowledge that is trustworthy for its effectiveness and transparent enough to be translated to new teaching contexts to improve student learning (Cohen & Ball, 1999; Davies, 2007; Hiebert, Gallimore, & Stigler, 2002). Once school leaders and teachers can recognise the principles behind the knowledge, they can filter and test new ideas and innovate practices in relation to these principles (Timperley et al., 2007).

Increasingly, federal and state government policies are setting ambitious goals for student achievement that require transformative changes to core-of-teaching practices in order to meet the challenge of large numbers of underperforming ethnic minority groups (Borko, 2004; Elmore, 2002a; Stein & Coburn, 2007). Parr and Timperley (2006) refer to the depth of knowledge required to shift achievement. They argue that teachers require deepened content knowledge to understand the progressions that learners take and to analyse the strengths and gaps that learners present in their work. Teachers also need to develop strategic pedagogical responses to these needs, well beyond knowledge of general strategies and activities, if they are to quicken the pace of progression and sustain improvements. Teacher Professional Learning and Development Best Evidence Synthesis Iteration (Timperley et al., 2007) reinforces this point. These authors found that interventions demonstrating ongoing improvement in student outcomes had emphasised equipping teachers with strong theoretical knowledge as the foundation of changes to their practices. Franke et al. (1998) elaborate on this notion of principled knowledge in
their studies of the Cognitively Guided Instruction approach to mathematics in the United States. They argue that the learning for teachers needs to be guided by the idea that teachers are professionals rather than technicians and that teachers require the theoretical frameworks that lie beneath how children develop their thinking so that they can make instructional decisions based on that theory. They report that teachers learning how to apply this theoretical knowledge more flexibly in their own classroom context resulted in consistent gains in student outcomes.

Learning processes for teachers

This theme combines several features of professional learning that research has identified as leading to sustainability. These features are mostly situated in the key construct of professional learning communities, within and between schools. While sustainability may begin with the individual, it also requires school organisational structures to place the changes as a central component of all routines (Datnow, 2005; Fullan & Sharratt, 2007; Jasinski, 2007; Wood, 2007).

Teaching (and leadership) practices are influenced by an individual’s personal theories about how to be effective (Spillane et al., 2002). Teachers may have to “unlearn” to be able to change these beliefs so as to reposition their views and their power relationships with students. For instance, Bishop and O’Sullivan (2006) argue that profound changes in teacher–learner interactions are required to lever up and sustain outcomes for those groups of students who have a long history of underachievement.

Hiebert et al. (2002) consider teacher knowledge as being characterised by its concreteness and contextual richness. They argue that research knowledge is, by contrast, often general and context independent and that professional learning should, therefore, uncover both sets of knowledge and promote a public knowledge-base for the profession, which is open for discussion, verification and modification. These authors contend that teachers respond more effectively when they can translate the theory they are working with into the context of their own classroom practice, or when they recognise how their current practice fits a coherent theory.

Structures and processes that enable teachers’ theories to be unpacked and analysed for their efficacy are promoted as critical to improving the likelihood of sustainability (Borko, 2004; Timperley et al., 2007). These structures may be formal or informal, but require explicitness in their purposes and practices. Five interrelated
variables, distinct and critical to effective professional communities, have been commonly identified (Fullan & Sharratt, 2007; Timperley et al., 2003). These include:

- shared values and expectations about students, learning, teaching, and teachers’ roles;
- a clear specific focus on student achievement;
- sharing expertise in a collaborative culture;
- deprivatised practice through peer coaching and observations;
- engagement in reflective dialogue about teaching that examines assumptions and beliefs.

On their own, professional learning communities that work in collaborative ways may not be a guarantee of continued improvement (Coburn, 2001; Timperley et al., 2007). In New Zealand, Timperley and Phillips (2003) delved more deeply into the nature of professional communities and what specific factors would ensure more traction with lasting improvement. This work sharpens and deepens the knowledge needed to lead and participate in evidence-based communities. Timperley and Philips argue for the need to “accentuate the goal of promoting student achievement and learning” in these communities. They examined conversations within these communities about student achievement, finding that those schools that closely linked evidence of student achievement to the impact of their teaching practice (and not to external factors), that took ownership of the educational problem, and that had a sense of urgency about improvement had sustained the gains made in earlier professional development. Professional development has to impact on a teacher’s sense of self-efficacy. Otherwise, if teachers believe that there is not much they can do to improve student achievement, there is little incentive for them to engage in alternative instructional practices or sustain any new practices (Timperley & Phillips, 2003). These authors found that before the professional development in their study, over 80 percent of teachers nominated external factors as reasons for underachievement, while the post-course questionnaire recorded over sixty percent of teachers nominating school-based factors as critical to this issue.

More recent literature on particular conditions of professional learning communities that led to improved outcomes for students has been identified in *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration* (Timperley et al., 2007). These conditions include where teachers were challenged about the efficacy of their beliefs about teaching, where teachers had high expectations for all students and
where they used artefacts representing student learning to ground their discussions about student achievement.

Leaders and those supporting the improvement effort in these professional communities of practice need to be equipped with strategies to promote cognitive conflict and to balance challenge with support, and they need to be able to ask incisive, informed questions that challenge teachers to re-examine their practice in a way that is respectful, basing their dialogue on data and evidence of effectiveness (Annan, Lai, & Robinson, 2003; Lima, 2001; Timperley, 2001). Interestingly, all the studies of professional learning communities cited in the synthesis by Timperley and her colleagues (2007) that did not lead to advanced achievement outcomes for students lacked external input. Presumably external experts can recognise and mediate competing theories at work in a practice community, offer new insights, and better support teachers to test their theories and assumptions.

*Evidence-based inquiry practices*

Inquiry into the efficacy of teaching practices that impact on student outcomes appears to be a fundamental shift from traditional ways of working for leaders and teachers in schools. Inquiry, as a means to discover more evidence, to build teacher knowledge, or to decide next steps for learners is a relatively new term in educators’ language. As a process for teacher learning and decision making, evidence-based inquiry is the most recent of school-based conditions for sustainability to be found in the research literature. The term embraces a number of evaluative capabilities, is variously described by researchers, and can operate at several different levels in an education system simultaneously (Reid, 2004; Timperley & Parr, 2007). The literature on student self-regulation has been applied to the field of teacher learning in a number of studies (Butler, Lauscher, Jarvis-Selinger, & Beckingham, 2004; Donovan, Bransford, & Pellegrino, 1999; Timperley & Robinson, 2001) adding to the repertoire of inquiry skills required by teachers.

As discussed in the previous section, leaders may have fostered new practices but not necessarily worked with the teacher beliefs that underpin those practices, or examined the evidence of the impact of these beliefs on student achievement (Robinson & Lai, 2006; Timperley et al., 2007). Processes within professional learning communities that gather and analyse evidence to find the impact of current theories of practice on student achievement, and that also examine the impact of new practices, are at the heart
of an inquiry model. Robinson (1993) views changing teaching practice as a process of engaging in problem-solving and this requires change agents to support communities of practice to examine the tacit and untested theories that they hold about teaching and their learners and the validity of the reasons they have for students who are not progressing adequately. Robinson and Lai (2006) argue that using this problem-based methodology enables current and new practices to be tested for their efficacy in supporting valued student outcomes.

Indeed, student achievement data support teachers to problem-solve issues of underachievement and to question, challenge and articulate their practices rather than simply adopt some new ideas or strategy (Timperley, 2003). Teachers’ professional dialogue about classroom-based problems and issues of practice is then grounded in evidence, not anecdotes. Earl and Katz (2006) provide a further list of capabilities that collectively ensure ongoing improvement. They report that teachers need to:

- know what evidence to gather to investigate problems;
- be data literate with all of the assessment tools available in their curriculum area of expertise;
- be able to identify trends and patterns;
- know how to compare results over time;
- value data.

These capabilities underpin the goal of improving student learning and achievement in schools by supporting analytical professional conversations (Annan, Lai, & Robinson, 2005). They are, by inference, essential to promoting sustainability of gains in student outcomes. Beyond working with data per se, is a wider notion of co- and self-regulated improvement practices that operate at several levels of the school, including students’ capability to use inquiry processes to self-regulate their learning.

The term self-regulated learning (SRL) became popular in the 1980s because it emphasised the emerging autonomy of students to take responsibility for their own learning (Paris & Winograd, 2001). Timperley and Parr (2009a) audiotaped 17 writing lessons with teachers who were participating in the first cohort of LPDP schools. They were looking for evidence of core practices of the LPDP: whether lesson aims were clearly articulated by teachers and whether students were able to identify deeper features of writing as their lesson aims. They found that even when the aims were made explicit, the mastery criteria and lesson activities were often misaligned and, consequently,
students could not identify the learning intentions of the lesson. The related research literature argues that, without these understandings, it is unlikely that students will be able to self-regulate their learning (Black & Wiliam, 1998; Paris & Winograd, 2001). Lefstein (2008) emphasised the challenges that teachers face in changing the nature of learning in their classrooms. He was investigating whether particular classroom strategies, exemplified in core teacher materials for the English National Literacy Strategy, were actually in use, but also what impact the micro-interactions with students had on teachers’ practices. The enacted curriculum in this study of two classrooms fell far short of prescribed aims for the lesson from either the materials or the teacher. This was because the students “colluded in construction of the lesson, constraining possible teacher actions” (p. 731), and so lessons persisted with little variation from the normal set of teacher–student interactions. Butler and Cartier (2004) found that if task interpretation was absent or faulty, then learning was derailed (p. 1735) suggesting that clarity of learning for students is critical to their ability to learn effectively.

This body of research on SRL subsumed the literature on metacognition, motivation, and cognitive strategies under one construct and, like many terms about student learning, has been applied to the field of teacher education. While the term refers to self, this does not conflict with the idea of cognition being situated in a sociocultural context for learning, where teachers set out to analyse and solve problems about learning and achievement. Franke et al. (1998) describe this capability as “self-sustaining, generative change” brought about by individual and collective problem solving, rather than just acting out a process or a set of procedures. These inquiries occur within a collaborative model that supports collective responsibility for rigorous investigation of evidence, theory engagement and selection of solutions. Reid (2004) describes inquiry in the school context as:

a process of systematic, rigorous and critical reflection about professional practice, and the contexts in which it occurs, in ways that question taken-for-granted assumptions. Its purpose is to inform decision-making for action. Inquiry can be undertaken individually, but it is most powerful when it is collaborative. It involves educators pursuing their “wonderings” (Hubbard & Power, 1993), seeking answers to questions or puzzles that come from real-world observations and dilemmas. (p. 3)
Thus, inquiry is a process where it is essential for people to develop the ability to regulate their own learning, both at the individual level and as a collective enterprise. Integral to co- and self-regulation is evaluating what you know and what you do not know. It is quite reasonable to expect that in the busy organisational milieu of a teacher’s work, the individual does not question their own knowledge for validity and efficacy on a daily basis. Reflection may occur, but this may not have the necessary rigour, or be as systematic in nature as Reid describes above.

*Teacher Professional Learning and Development: Best Evidence Synthesis Iteration* (Timperley et al., 2007) offers an inquiry model that enables systematic review in order to improve student outcomes. The construct denotes dynamic actions rather than a ritual of processes or teacher practices and promotes flexible and adaptive responses. The authors are clear in making the distinction between inquiry as using evidence about student needs to inform teaching decisions and using the same data, or data on current teaching practices, to also reveal teachers’ learning needs. They argue that inquiry without a determination to build new knowledge with teachers will not necessarily lead to improvement. In their model, new learning is undertaken as a result of looking at both student and teacher data, new or adjusted teaching actions are transferred to practice and then these practices are regularly monitored for their effectiveness.

Evidence-based practices are therefore a subset of what is described as inquiry as a self-regulated process for the building of new knowledge. For instance, the early literature in the field of assessment for learning focused on more generic teaching practices that were known to support improved student outcomes (Black & Wiliam, 1998; Black & Harrison, 2001). These practices were described as evidence-based practices and included use of effective feedback by teachers, the sharing of learning intentions and success criteria with students, and the increased involvement of students in their own learning using self and peer assessment. However, these practices were not necessarily linked to an inquiry-based approach where student achievement data were the starting point to determining teacher responses.

A distinction between this inquiry cycle and evidence-based decision making based on student achievement emerges in a study by Coburn, Toure and Yamashita (2008). They describe the “interpretative space” that follows being confronted by evidence of student underachievement. In their study, decision makers at district level did not return to the data to ask further questions about the data, often determining new
pathways without fully framing the problem that they were dealing with. If we apply this finding to the school setting, school leaders and teachers also need to acquire skills to support this interpretative space so that they can interrogate their data to gain a finer grained understanding of the issues of student achievement that they face and the relationship this has to teachers’ knowledge.

Figure 1 taken from *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration* (Timperley et al., 2007) graphically illustrates this idea. The cycle determines teacher knowledge-building as a pre-requisite for new responses and then systematic monitoring of the impact of those responses on issues of achievement. A similar model in the recently published *The New Zealand Curriculum* (Ministry of Education, 2007) reinforces teaching as an inquiry into individual effectiveness. Timperley and her colleagues noted that teacher self-regulatory skills were a key feature of those studies where teacher and student outcomes were sustained after the regular involvement of an external provider had been withdrawn. Implicit in these models then is the collaborative nature of the inquiry process, its need for leadership and its relationship to other communities, especially congruence with larger institutional policies and procedure. Similar versions of the model can be described for all levels of professional inquiry (Timperley & Parr, 2007). For instance, leaders or school district administrators can ask similar questions as they inquire into what new knowledge they need to build as a result of examining and interpreting data (Coburn et al., 2008), what new actions they need to take as a result, and how they might evaluate the impact of their new practices on student outcomes. Wenger (1998) argues that school-based communities of practice hold the key to real transformation and improvement. In themselves, professional learning communities may not be intrinsically beneficial or harmful but, with systematic inquiry as their focus, they can become producers of new knowledge that is debated and evaluated for its effectiveness. Hannon (2008) uses the term “next practice” to offset the idea of scaling up best practice in education that may just serve to support learning for a few. She argues for inquiry as the central paradigm for schools so they can innovate for themselves and keep checking for impact of new practices on their learners.
Copland (2003) also draws attention to different outcomes in relation to the development of inquiry in his study of a region-wide school renewal effort in the Bay Area School Reform Collaborative (BASRC). Schools reached differing levels of sophistication in their understandings and practices about using student achievement data to inform schooling improvement. Three “stages” were identified. By the end of their fourth year of implementation of BASRC’s cycle on inquiry, teachers in “novice” schools remained daunted by the process of collecting and analysing data. By contrast, schools in the “intermediate” stage were able to work through a cycle of inquiry and connected small improvements in their practices to this process. He noted that this group of schools tended to rush to a solution after they had gathered the data, choosing a strategy before defining the problem. Finally, the “advanced” schools moved beyond

---

carrying out inquiry as a procedure to more fundamental changes in the way that teachers worked. Teacher communities in these schools created knowledge within the school site, inventing ideas for reform after closely examining their own practices and the work of students who did not meet grade-level expectations in literacy. Classroom-based inquiries were linked to whole school inquiry in these schools and inquiry became a “stance” rather than a process.

This distinction between evidence-based and inquiry-based practices was again observed by Lai et al. (2009) in their investigation of two case study schools. The factors that differentiated one school as a lower and one as a higher gain school were the approaches that the schools had taken to solving specific problems and whether they had reframed the issues that they faced as problems that could be overcome by strategic action. For example, the higher gain school began an inquiry into the summer effect drop in scores that their achievement data were indicating and then integrated theories about reading mileage and partnerships with parents and the community to support new responses to this problem. By contrast, the lower gain school adopted routine responses to issues of achievement, looking for ideas that “worked well” in classrooms but were not targeted to the specific nature of the problem. Secondly, the higher gain school deliberately developed strategies to offset and reduce high teacher turnover, not leaving the issue as an explanation of low achievement that could not be solved. Lai et al.’s (2009) study further support the knowledge-building inquiry cycle described in Teacher Professional Learning and Development: Best Evidence Synthesis Iteration (Timperley et al., 2007) as being adaptive in nature, involving reflection on school-wide achievement data and teaching practices, as well as reflection on the research evidence of “what works,” in the decision-making process.

Another term being used interchangeably in the literature with evidence-based inquiry is “evaluative capability.” Essentially, definitions of this term nominate competencies that are closely interlinked with those identified in the inquiry literature, with more emphasis on the “interpretative space” as outlined by Coburn et al. (2008). In this literature, the educational problem to be explored needs to be described, theorised, and debated for its relevance and validity before solutions can be promoted. In turn, each solution needs to be evaluated for its validity and likely effectiveness. These are highly skilled capabilities that few schools have access to within their own professional community. Timperley, McNaughton, Lai, Hohepa, Parr & Dingle (in press) suggest that
Chapter 2 – Literature review: Reframing sustainability

high dependence on external agents to support these evaluative decisions may lead to
dependence and, indeed, limit schools’ ability to be adaptive and to sustain inquiry. This
idea returns to the arguments around the concept of sustainability as being about “going
it alone.” Where dependence on external agents may be the enemy of sustainability, so
too might independence work against sustainability. This emerging literature argues that
some form of interdependence should be maintained since it is unrealistic to assume that
schools could develop full autonomy in each curriculum domain. These authors argue
that schools need to develop collaborative partnerships with experts in order to share
responsibility for inquiry, innovation and improvement.

Coherence and strategic action

The final theme raised in the research literature in relation to sustainability of
professional learning is focused on coherence and alignment of the messages that schools
receive about effective instruction. These messages are carried by many means: policy,
curriculum materials, professional learning programmes, professional organisations,
assessment systems, and individuals (Stein, Hubbard, & Toure, 2008). Researchers have
examined the notion of coherence at both school and system levels to determine its
importance in schooling improvement. At both levels, coherence has been described and
examined in different ways. For instance, researchers may view coherence as the
alignment of effective teaching and learning practices across curriculum programmes or
as the development of consistent and effective practices within a single subject domain.

Coherence at the school level

Newmann, Smith, Allensworth, and Bryk (2001) define instructional coherence at
the school level as working across the curriculum, describing it as:

…a set of interrelated programmes for students and staff that are guided by a
common framework for curriculum, instruction, assessment, and learning climate
and that are pursued over a sustained period. Reform that strengthens
instructional programme coherence contrasts with efforts to improve schools
through the adoption of a wide variety of programmes that are often un-
coordinated or limited in scope or duration. (p. 297)

Newmann and his colleagues contend that too many unrelated improvement
programmes exist for schools and that their teachers divide themselves across various
initiatives because they want to access all available support and knowledge. Many school
districts now favour whole-school initiatives but, as these authors point out, there is little research yet on how instructional coherence might constitute an important schooling improvement strategy, nor a theoretical explanation about why coherence might advance student learning. “Although whole-school models may spur more extensive improvement or restructuring efforts, they may or may not address links between the reforms and their collective links to student achievement” (Newmann et. al., 2001, p. 299).

Sustainability is inevitably about what the school does in the wake of reform, how their community of practice interprets and holds fast to the core beliefs and values of each professional development initiative that they undertake. Newmann et al. (2001) explain further that while the responsibility for the alignment of key messages is located at the school level, it seems likely that little support is given to schools to continually craft the fit between external policy demands and the schools’ own goals and understandings. According to Honig and Hatch (2004), when multiple messages converge on schools via professional learning projects, this coherence has to be deliberately “crafted” by school leaders so that schools have internal mechanisms to “buffer and bridge” the complexity of messages that exist in an increasingly “demand rich” educational environment. Educators have to translate the messages into day-to-day operations and thinking and then, more importantly, test their efficacy in achieving improved outcomes for their students (Timperley, 2003).

Newmann et al. (2001) found a strong positive relationship between improving programme coherence and improved student achievement in their sample of over 200 Chicago schools. Field studies in 11 schools involved in a large-scale schooling improvement project indicated they were able to sustain their efforts and work collaboratively to define a common framework for curriculum, instruction, and assessment over a significant period of time – three years at the time of the study. These authors, however, do not view coherence as the overriding task of schooling improvement but as a strategy for “maximising” supports for improved teaching and learning.

Robinson (2007) found that leadership practices that supported the planning, co-ordination and evaluation of teaching were associated with improved student achievement. For instance, school leaders can co-ordinate and align effective instructional practices in schools by taking strategic actions that offset all those factors that inhibit ongoing uptake of new practices. These actions might include:
• deliberate and explicit links made between professional learning projects and between the macro-policy context and school practices so that the key messages about effective instruction can be recognised and transferred into new contexts by teachers (MacBeath & Dempster, 2009);

• engaging with parents and the community to support alignment between learning at home and learning in schools (McNaughton, 2002);

• induction programmes for new teachers to the school that amplify these key messages and set expectations for their application (Coburn, 2003).

Stein and Coburn (2005) and Stein et al. (2008) also argue that school leaders can spearhead and frame coherence of instructional changes to ensure that there are scaffolds to support cognitive levels of improvement at work in the school. They draw on a framework from the National Academy of Education (1999, cited in Stein et al., 2008) that identifies how learning is facilitated by ideas, people, and tools and the various mixes of these. These authors also describe a fourth dimension at work in the learning process: the “pedagogy of travel,” by which they mean the stance taken by those promoting the changes to instructional practice and their adaptation to a new context. The tools that are developed by the school (or by the programme developers) can embody the theory both behind the change in the practice and behind how people learn as well as guide individuals’ actions. The tools become a defacto representation of what the educational idea looks like in practice, and support new users of the idea to understand the design principles and how to enact them. Robinson, Hohepa, and Lloyd (2009) use the term “smart tools” in their research on leadership and its impact on improved student achievement. They describe a range of tools that have particular routines that enable users to understand the theories that are applied in their teaching. These tools can therefore promote coherence and alignment of practices across a school consistent with best evidence about effectiveness. They may be developed by central government, at district level, or within a single school. Coldren and Spillane (2007) also confirm the influential role of tools in instructional leadership in a case study of one principal and her routines around writing reviews to support 50 teachers in her school. These tools may prove to be a critical element to support sustainability. Timperley and Parr (2009) warn, though, that while a particular tool may represent the core learning of the project, it is how the tool is integrated into the routines of practice within each school that may affect its impact on student achievement.
New Zealand researchers also take up the idea of coherence in relation to sustainability of professional learning, focusing more on coherence as representing consistency of practices with a single domain across the school, than on ways to develop a “common framework” across curriculum, instruction, assessment and learning. For instance, Lai et al. (2009) posit that sustainability of student achievement gains were enhanced when schools embedded a professional learning community into their normal school routines so that they could build coherent literacy instructional programmes, calling on expertise when required.

**Coherence at system level**

Newmann et al. (2001) view coherence as a dynamic process, a process, in fact, that might provide a more productive organising construct for policy makers as well as schools. Fullan (2000) argues that the “outside in” model of professional learning promoted at system level needs to be coherent and aligned with the “inside” or “inside out” model of learning that occurs in schools. School and system leaders need to work on the same agenda for sustainability (Fullan & Sharratt, 2007). Datnow (2005) comes to similar conclusions. Supovitz and Taylor (2005) also focus on the theme of system coherence, arguing that powerful improvements in teaching and learning can come from developing coherence and alignment across the complex and different elements of an educational system. Indeed, Levin et al. (2008) are optimistic about recent school reforms in Ontario being sustained because they have been implemented with attention to all elements of schooling simultaneously, with fewer initiatives and more focus on deeper implementation.

From the outside in, then, different reforms might work to amplify the key messages about effective practices as a way to scale up reforms that do produce effective outcomes for students. To achieve amplification and breadth as well as depth of professional learning, Coburn (2003) suggests building “normative coherence” between professional learning projects. By this she means that school districts can recast their role from supporting individual reforms to a more deliberate strategy to spread reform norms and beliefs and to creating knowledgeable leaders who can influence policy and procedures at this level. She sees this as an alternative to thinking solely in terms of expanding individual programmes to more schools as a means of achieving breadth of outcomes across a school system. Madda, Halverson, and Gomez (2007) also comment on how coherence can be deliberately designed at a system level. The focus of their
research was on one school district’s efforts to develop an artefact to access student achievement data across their district that would fit coherently into existing initiatives. They concluded that while district leaders did design an artefact that incorporated coherent messages about effective collation and analysis of data for informing instruction by involving a stakeholder-based process, they had not paid enough attention to local use of that tool for it to be effective in sustaining changes in data-gathering practices.

Stein and Coburn (2007) focus on another system dimension that can promote coherence of effective instruction. They investigate networks of practice, arguing, like Wenger (1998), that people (“brokers”) working across communities and the tools or “boundary objects” that they bring to each community each support learning that is coherent. Wenger, however, warns that if there is too much reliance on either people or tools, then teachers’ uptake will be literal and procedural. Stein and Coburn (2007) contend that alignment is a key ingredient of the large-scale, co-ordinated forms of learning required for successful and long-lasting initiatives:

New forms of engagement within an individual, micro-community of practice will spur learning within that particular community, but they will not – on their own – connect that learning to the broader enterprise of whole district reform. It is through consciously designed processes of alignment that individuals become invested in tasks that are defined beyond the engagement of their own community. (p. 7)

Annan (2007) promotes another form of system coherence in his model of effective schooling improvement initiatives. His research mapped professional learning models in a range of curriculum reforms in New Zealand, the United States and the United Kingdom. He found that vertical transfer of standardised effective practices from national centres of expertise to classrooms can be balanced with horizontal transfer of practices promoted via learning communities within and across schools to achieve accelerated and sustainable outcomes for students.

The potential for developing coherence of instructional programmes to support sustainability at school and system level across the curriculum and within subject domains is certainly clear and powerful but not yet fully established as a necessary condition in the research literature.
A conceptual model for sustainability

Based on what is currently argued in the research literature on sustainability, it seems reasonable to collapse these divergent themes into two dimensions that are more closely interwoven and can potentially impact on learning at two or more levels of the system, because learning at one level needs to be supported by learning at another (Timperley & Parr, 2007; Reid, 2004). These levels include learning for individual teachers, for school leaders, for school communities, for local or district leaders, and systemically. The two-dimensional model represents the fundamental hypothesis being examined and serves as a framework for the structure of this thesis.

Co- and self-regulated inquiry practices for improvement

The first dimension of the conceptual framework is an orientation for teacher and system learning that enables co- and self-regulation for improvement. This dimension conceptualises teacher change in terms of teachers carrying out authentic inquiry into problems of practice and student achievement that requires deeper analysis with each new cycle of inquiry. There is a distinction emerging in the literature between use of evidence-based practices, such as using student achievement data to inform next steps in learning (Black & Wiliam, 1998; Sutherland, 2004) and the use of an inquiry model that links teacher knowledge and practices explicitly to student outcomes (Copland, 2003; Earl and Katz, 2006; Robinson & Lai, 2006; Timperley et al., 2007). The inquiry dimension in this thesis differentiates between those schools that engage in a recursive inquiry, where they ask the same questions of their student achievement data as they did the year before, and those that work iteratively, connecting each ongoing inquiry and asking different questions as they begin to pinpoint groups of students that still need attention, while always testing their assumptions and linking their teacher practices to the impact on achievement. Emerging literature defines sustainability as transcending maintenance of recently learned practices to building capability to be successful in new and demanding contexts (Century & Levy, 2002; Davies, 2007; Wood, 2007).

Coherence of effective instructional practices

The second dimension is the degree to which the professional development programme is integrated into coherent, effective instruction across the curricula in a school. Elmore (2002a) describes coherence as the connection between the “big ideas” about improved practice and the “micro-world of teaching practice” (p. 29). These big ideas may compete or conflict and so they must be processed or resolved in some way so
that teachers can lift their thinking beyond the practices to the underlying principles that they represent. In this way, a school can develop a coherent set of principles and practices, both shared and explicit, which are regularly tested for their efficacy in improving outcomes for all students. Elmore (2002a) also argues that professional development is a collective good rather than a private or individual good. Its value should be judged by what it contributes to the quality of instruction in the school and school system. In other words, each professional learning initiative promotes new practices and ideas about effective teaching and individual teachers build their knowledge of effective practices over time and meld them to their own contexts (Spillane et al., 2002). Transfer of strategies across curriculum areas occurs where teachers find them effective for students. For instance, the assessment for learning research literature actively promotes transfer of formative assessment strategies across the curriculum to support improved student outcomes (Black & Wiliam, 1998; Wiliam, 2007). Thus, this research argues that the lens for sustainability should be broader than a single professional learning project and focused across projects within the system or within a school.

Figure 2 incorporates these dimensions on two axes. Each axis represents a continuum of understandings and practices that may be present in the school. The horizontal axis describes the development of effective and coherent classroom instruction. One end of this continuum might begin with a perception of professional development projects as disparate, each concerned with its own set of changes to knowledge and pedagogy. The other end may represent a stance that can allow teachers to make connections across projects and to explore and understand the beliefs that exist about the nature of effective pedagogies. The vertical axis is about inquiry into issues of learning and achievement. It encompasses the development of a professional learning community in a school, starting with learning activities and structures that support teachers to learn from and with each other, and with experts in an ongoing way. This continuum works towards an inquiry culture where teachers can self-regulate their practices, both as individuals and as a community, to support ongoing improvement in student outcomes (Timperley et al., 2007). In other words, inquiry is ongoing and leads to more focused investigations and better questions (Copland, 2003; Earl & Katz, 2006). The curriculum content knowledge of the professional development project serves as the backdrop to these axes, as does the rich context of each school.
Figure 2. Conceptual model for sustainability.

This model provides a framework to evaluate whether professional learning has been sustained, at the same time expanding the notion of sustainability beyond the context of a single professional learning project. Its purpose is to focus closer attention on whether schools “craft” their new learning and practices onto a coherent set of principles about effective instruction, within and across the curriculum. The model supports a definition and theory of practice for achieving sustainable school improvement that can be used by all levels in the system. The model is underpinned by four key assumptions derived from this literature on sustainability. These assumptions are:

- Sustainability is measured by continuous improvement in student achievement, particularly for those students whose patterns of achievement place them at risk of not meeting expected levels over time;
• Any adaptation of professional learning programme practices over time, and any new programmes and practices, are explicitly reviewed against the principles of effective instruction that are supported by research as improving learning and achievement and that are tested for their efficacy using school-based inquiry and knowledge-building improvement processes;

• A professional learning project contributes to the overall coherence of effective instruction in a school (and is not “sustained” as a single entity) and to the knowledge and practices of individuals in that school that relate to effective schooling improvement;

• Sustainability of effective practices and continuous improvement in student achievement requires strategic action in schools and by policy makers to make coherence explicit in order to offset the impact of teacher and leader attrition, any reduction of funding, and changing educational priorities.

This framework of sustainable improvement also assumes that the critical focus for each curriculum professional learning project needs to be on developing teachers’ pedagogical and content knowledge. The theoretical framework promotes recognition and transfer of pedagogical knowledge across professional learning projects so that new practices and the beliefs that underpin them are aligned and further strengthened. Understanding this coherence can avoid the creation of separate processes for each project in the wake of their implementation period, instead promoting commonality of purpose in improving existing infrastructures.

This model shaped the design of the research instruments used in this study and supported the analysis of the data gathered in 16 schools that had participated in the first cohort of the LPDP as they set out to sustain the gains they had made.
CHAPTER 3
STUDY 1

SUSTAINABILITY: A PROBLEM OF DEFINITION AND PERSPECTIVE

Introduction: An investigation into 16 schools 2006–07

Karl Hostetler (2005) reminds us that “the ‘answers’ to research questions do not end things but offer new circumstances” (p. 21). Study 1 set out to investigate the nature of sustainability of educational change in schools and to raise possibilities. Its purpose was not to reach definitive answers, but to raise new questions and to offer new insights into the problem. Study 1 examined a purposively selected sample of 16 schools from the first LPDP cohort of 2004–05 (N=91) as they began their drive towards sustaining the shifts in teacher and leader practices that they had made during the implementation phase and, more importantly, towards sustaining the literacy gains that their students had made with a new cohort of students. This chapter outlines how these schools were selected and what methods were used to analyse their responses to the research instruments.

Chapter 4 then analyses the theories of practice for sustainability held by teachers and school leaders in these schools in 2006, one year from their exit from the LPDP. These data raise the possibility that schools’ views of sustainability are self-limiting in nature, reinforcing sustainability as being only about single professional learning programmes and unintentionally restricting the longevity and institutionalisation of new and effective practices. Chapter 5 reflects on the student achievement data gathered from the case schools and also investigates the issues associated with measuring sustainability in terms of students’ literacy achievement data in this context. The achievement data from Study 1 are compared to the results from one other similar New Zealand study. Chapter 6 presents findings from Study 1 that compare the student achievement data patterns to qualitative data gathered from the 16 schools.

Methodology

Study 1 examined 16 schools in the two years after they had completed their participation in the LPDP. The research methodology described in this section aims to match the problem being investigated and the research questions to the chosen methods. It also describes the features of that methodology that support generalisable findings.
Chapter 2 has explained the need to reveal much more about the complex notion of sustainability and the substantive problem of deciphering which conditions might best support schools to sustain the gains they make when significant professional learning supports are withdrawn. The problem is currently undertheorised since so few studies have focused on schools in the period after implementation (Coburn, 2003). Yet questions about long-term impact of policies and practices on diverse student populations, particularly on those students who are not well served in education systems, have emerged as priorities in all educational jurisdictions (Konstantopoulos & Hedges, 2008; Raudenbush, 2005). In this environment, the research agenda is most often the pursuit of causal relationships that, in turn, impact on the choice of methodologies.

Causation is a hotly debated phenomenon in all areas of inquiry including education. Some argue that causation cannot be directly observed and that only the regularities in the relationships between factors can be counted as causal, while others determine that context is an essential component of cause and that the role of meaning-making by individuals is essential to describing how or why an event occurs (Maxwell, 2004). Huberman and Miles (1994) use the phrase “local causality” in their work because they argue that cause may be contingent on social and cultural contexts. Maxwell (2004) claims further that qualitative methods can uniquely strengthen causal investigations by identifying the influence of contextual factors that cannot be statistically or experimentally controlled, by describing the unique processes at work and by elucidating the role of participants’ beliefs and values in shaping outcomes. The context for this research into sustainability encompassed numerous variables and it was impossible to control for all competing explanations, or argue whether or not particular actions or approaches may have caused results. However, measures were taken to support generalisable findings about school-based conditions that led to sustainability and these measures included the use of case methodology and cross-case analysis within a mixed methods design.

**Mixed methods design**

Tashakkori and Teddlie (2003) describe mixed methods design as research in which both qualitative and quantitative methods are used in parallel or sequential phases of the study. Smith (2006) defines mixed methods design or “multiple methodology” as when “studies or projects employ at least one quantitative and one qualitative method to produce knowledge claims” (p. 459). Mixed methods design is therefore different from
merely applying a mix of methods for data gathering like surveys, interviews and observations. The difference lies in why the quantitative and qualitative methods are used and how this relates to the analysis of the research hypothesis. Mixed methods design supports investigations of whether findings from the quantitative and qualitative methods are associated in some way, whether they confirm or disconfirm research hypotheses, and whether their interaction can lead to new or unanticipated knowledge or more synergistic findings (Day, Sammons, & Qu, 2008).

Mixed methods afford just that, a mix of data sets that combine to unpack different aspects of particular educational problems, also providing checks and balances between sources. As explained above, the goal of mixing methods design is not only to seek corroboration of sources but also to expand one’s understanding (Johnson & Onwuegbuzie, 2004). For instance, semi-structured interviews can expose the sociocultural beliefs and values at work in particular contexts, while document analysis can track developments in thinking and actions over time. Yin (2006) explains that the boundary between phenomena (such as sustainability) and context may not be clearly discernible and requires different sources of evidence to create checks and balances to interpretation. Mixed methods design can provide stronger evidence for a conclusion through convergence of findings. Questionnaires, semi-structured interviews, lesson transcripts, teacher meetings, and document analyses made up the mix in this study to embrace not only the breadth of participants that matter in the conceptualisation of sustainability, but also the breadth of activities associated with sustainability in a school.

The goal of a mixed methods approach in this study was as much about expanding knowledge about the characteristics of sustainability as it was about a search for associations between school-based conditions and sustainability of achievement gains with new cohorts of students over time. Day, Sammons, and Gu (2008) illustrate this idea of combining qualitative and quantitative data in ways that produce “more synergistic findings.” Their study of the lives, work and effectiveness of 300 teachers over a four-year period involved “qualitizing quantitative evidence and quantitizing qualitative evidence and integrating the two” (p. 333).

As long as each method is acknowledged with its assumptions, strengths, and limitations, then results can be analysed and made accessible for further testing in other contexts (Smith, 2006). In this study, qualitative responses from questionnaires, interviews, observations and document analyses have been summarised and presented
quantitatively, while quantitative data (student achievement gains) have been categorised in qualitative ways in relation to benchmarks set for sustainability.

Case methodology

A case study is an exploration of a bounded system in which detailed data are collected. The reason for adopting an in-depth case methodology was to investigate whether school-based conditions that promote sustainability might generalise across all case sites or if other conditions may be site-specific or even transient (Chatterji, 2005). The choice of case methodology was also to support an in-depth perspective within a real-life context (Yin, 2006). This preference reflects the need to investigate fully the issues at the school and classroom level to evaluate whether specific conditions relating to sustainability might be operating and where leader and teacher theories of sustainability might impact on actions taken, or not taken, to support ongoing improvement. In this way, qualitative research is useful for understanding connections in the lived experiences of participants and their impact in that context. Chatterji (2005) contends that temporal factors related to the stage of an intervention should also be considered when deciding on method. She argues that employing appropriate descriptive methods in site-specific contexts in the early stages of an intervention yields valuable qualitative variability in progress and checks theory development. It is not likely that there is a regular chain of contingent events leading to sustainability. A case study design, with one feedback loop, that is, one study informing the second, accommodates this idea.

The theoretical framework for sustainability being proposed in this research has guided the choice of data collection instruments and enabled the conceptual framework described in chapter 2 to be explored in the cases. The methodology for this research utilised qualitative analyses combined with quantitative investigations into student outcomes in each study. The latter was used particularly to evaluate the eventual effectiveness of approaches being taken in each case school. Maxwell (2004) adds:

To develop adequate explanations of educational phenomena, and to understand the operation of educational interventions, we need to use methods that can investigate the involvement of particular contexts in the processes that generate this phenomena [sic] and outcomes. (p. 7)
In this way, the study could discover and warrant the most promising practices around sustainability in particular contexts for others to examine and test further in similar or different contexts. Detailed descriptions of practices that do or do not work often supply new ideas on how to intervene. This scale of research is useful as a prerequisite to supporting larger-scale inquiries, improving interventions, and/or impacting on policy. In other words, such studies can be built upon to provide increasingly nuanced understanding of complex notions such as sustainability (Shaffer & Serlin, 2004).

The challenge presented to all those who engage in case methodology is the generalisability of the findings. The context of each school and its particular problems, theories, and actions for sustainability may be highly individualised, negating any generalisations for other schools not involved in the study. The logic of sampling theory says that generalisable findings can be made if the sample is representative of the whole but, of course, variables at work in schools, such as class size, instruction, and teacher theory and beliefs are often highly contextual. Such variables are difficult to measure quantitatively to infer across contexts. However, Robinson (1993) argues that findings from such cases can be generalised where researchers argue from generalised theory and existing empirical evidence and investigate such theory in their cases. This generalisability depends on the prevalence of the problem being researched and the extent to which the researchers’ explanations draw on school conditions that are known to operate in similar situations. The theories about sustainability of reform are not idiosyncratic. The problem is shared across all educational sectors and the researcher has drawn on extensive sources of evidence to develop a principled conceptual framework about sustainability in order to explore this in an empirical study. The choice of case methodology, with its focus on the why and how in each context, in relation to a well-specified hypothesis, also supports this view that findings can be generalised.

Yin (2006) argues that the case study method has attained routine status as a viable method for undertaking research in the educational setting. Its key strength is its ability to examine a complex phenomenon in its real-life setting. The fundamental research question posed in this initial study was: What is the nature of sustainability and the school-based conditions that promote it? This descriptive question required inquiry into the theories and actions of key participants – principals, literacy leaders and teachers. The case study design enabled the in-depth exploration required for exploring and
building theory. The research framework was grouped around two dimensions: the development of coherent effective instructional practices in a school and the capability of school leaders and teachers to undertake co- and self-regulated inquiry for improvement.

Cross-case analysis using mixed methods

Cross-case analyses, informed by theory, can highlight similarities and differences across contexts and allow for greater opportunity to generalise across several representations of a phenomena such as sustainability (Borman, Clarke, Cotner, & Lee, 2006). The choice of a mixed methods design that incorporated both quantitative and qualitative data within case methodology was deliberate, so that rich data about these complex practices around sustainability could be examined and compared from all angles and across each context. A mixed methods approach can draw from the strengths of quantitative and qualitative methods, at the same time minimising weaknesses of both in single research studies (Johnson & Onwuegbuzie, 2004). They work in combination. For example, semi-structured interviews and observation frameworks adds some degree of uniformity across sites, but still generate wide-ranging data within various specified contexts. Combined with a close-ended instrument such as student achievement assessments, the qualitative data can be more systematically compared and contrasted (Johnson & Onwuegbuzie, 2004).

Methods for making generalised claims from a set of observations in cases all depend on the idea of sampling. Schaffer and Serlin (2004) describe the conditions by which a selection of schools best renders a “fair” sample to build more generalisable findings and minimise any confirmation bias. This includes being able to triangulate from other sources and to provide participants with the opportunity to give feedback on findings. In this study, the mixed methods design helped to triangulate data and establish converging lines of evidence. For instance, what leaders in interviews stated as the key messages sustained in the school could be corroborated or challenged by the transcript analysis of the teacher meetings and classroom observations. Member checking also occurred in this study.

A sequential design for inquiry has involved the data from Study 1 informing the content and conduct of the second study of the research. Accuracy of findings from the qualitative research in the 16 case study schools was increased by ensuring that disconfirming evidence was noticed and coded as such, giving it appropriate weighting in the data analysis process. Cronbach (cited in Robinson, 1993) argues that testing a
hypothesis involves serious attempts to falsify it. Case study approaches often have difficulty evaluating causal pathways because any given case selected for study may fail to display such a path even when it exists in the larger population of potential cases. Numerous alternative theories may also be consistent with data gathered from a case study. To offset this issue, triangulation with student achievement data and systematic coding of qualitative data for both confirming and disconfirming evidence was used in Study 1. The qualitative data have been coded (after establishing appropriate inter-coder agreement) and used to rate responses so that schools could be mapped onto the theoretical model presented in chapter 2. The student achievement data provided a critical check as to the strengths and weaknesses of the theoretical model. Did the schools that sustained their achievement gains in 2006–07 in any way match the dimensions promoted in the conceptual framework as furthering sustainability?

**Ethical considerations**

The research project is a component of the embedded research programme for the LPDP being undertaken by The University of Auckland 2004–09. Ethics approval for this research was granted by The University of Auckland Human Participants Ethics Committee on 13 May, 2004, with a further extension granted for 2007–08 (reference no. 2004/059).

The researcher was a member of the research team while enrolled as a PhD candidate at The University of Auckland. Appropriate participation information for the Study 1 research activities was sent to all participants and consent forms gathered according to the guidelines provided by The University of Auckland Human Participants Ethics Committee. Appendix A contains the participant information and consent forms for this study.

The researcher is also the manager of the LPDP and is employed by Learning Media Limited which has the responsibility for the implementation of the project under contract to the New Zealand Ministry of Education. The potential for any bias from the researcher in this dual role has been mitigated by several factors. This includes the use of a research assistant to conduct the structured interviews (following training by the researcher), to record the classroom observations and to collect student achievement data and documents. All coding of qualitative data in Study 1 has been subjected to inter-coder reliability testing by the research assistant, who has no other interest in the project.
The selection of the schools for Study 1 was determined by criteria largely independent of any decisions made directly by the researcher. In this way, the researcher could avoid any ethical issues of power or control where participant responses might be tempered by a desire to be supportive or uncritical of the LPDP. In Study 2, the researcher selected four schools based on criteria and data gathered in Study 1 and carried out the interviews and observations with the assistant researcher. Any ethical issues around selection bias no longer applied since the schools came from the original sample.

There was a second ethical issue related to the researcher’s dual role: she had access to all the LPDP databases and often to informal comments made by project facilitators about specific schools. Any data used for the analyses were therefore only those for which school boards, principals, teachers, students, and facilitators have given consent.

Selection of cases

Ninety-one schools participated in the original cohort of the LPDP in 2004–05. Forty-nine schools (593.6 teachers) had a writing focus, and 42 schools had a reading focus (589.5 teachers; Bareta & English, 2006). In effect, there were two selection processes to determine the schools for Study 1. Firstly, the selection criteria for the sample included that they were a facilitator “case-study” school for the LPDP. Approximately six months into their work with schools, each facilitator was asked to select two schools (unless they had only one school in their caseload) that were at different stages of their development and that appeared fairly representative of their caseload of schools at that time. The purpose of this selection was so that facilitators could detail the progress of these schools in more depth in milestone reports for the project leaders during the implementation period. In total, there were 37 case-study schools in cohort 1. From this group of schools, a purposive sample was finally selected that aimed to include both reading and writing focus schools and a range of deciles, school types and facilitators.

Participants

These criteria were used to select 20 schools and to invite them into this study on sustainability. Sixteen schools (17.6% of cohort 1) agreed to participate in a one-day school visit and to provide collations of student achievement data. Table 1 indicates that the Study 1 schools differed in two ways from the original cohort of 91 schools.
Comparative percentages indicate that significantly more intermediate schools (Years 7 and 8) and fewer full primary schools (Years 1–8) were represented in the research sample and that fewer writing schools were represented. These differences were mostly related to the size of schools and the year levels being taught. Neither of these variables was particularly critical to the findings. Most research indicates that it is teacher and leader turnover that impacts on the likelihood of sustainability (Fink & Brayman, 2006; Wood, 2007). However, school size still ranged from 5 to 24 teachers in the Study 1 sample and small rural and large urban schools were both represented.

Table 1
Comparison of Study 1 case schools to the LPDP Cohort I 2004–05

<table>
<thead>
<tr>
<th>Type of school</th>
<th>No. Cohort 1 schools</th>
<th>%</th>
<th>No. facilitator case study schools</th>
<th>%</th>
<th>No. Study 1 case schools</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading focus</td>
<td>42</td>
<td>46.15</td>
<td>21</td>
<td>56.76</td>
<td>10</td>
<td>62.50</td>
</tr>
<tr>
<td>Writing focus</td>
<td>49</td>
<td>53.85</td>
<td>16</td>
<td>43.24</td>
<td>6</td>
<td>37.50</td>
</tr>
<tr>
<td>Total</td>
<td>91</td>
<td>100.00</td>
<td>37</td>
<td>100.00</td>
<td>16</td>
<td>100.00</td>
</tr>
<tr>
<td>Full primary</td>
<td>40</td>
<td>43.96</td>
<td>13</td>
<td>35.13</td>
<td>5</td>
<td>31.25</td>
</tr>
<tr>
<td>Contributing</td>
<td>33</td>
<td>36.26</td>
<td>11</td>
<td>29.73</td>
<td>4</td>
<td>25.00</td>
</tr>
<tr>
<td>Intermediate</td>
<td>13</td>
<td>14.29</td>
<td>11</td>
<td>29.73</td>
<td>7</td>
<td>43.75</td>
</tr>
<tr>
<td>Years 7–15</td>
<td>2</td>
<td>2.20</td>
<td>1</td>
<td>2.70</td>
<td>0</td>
<td>6.25</td>
</tr>
<tr>
<td>Special</td>
<td>2</td>
<td>2.20</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Restricted</td>
<td>1</td>
<td>1.10</td>
<td>1</td>
<td>2.70</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>91</td>
<td>100.00</td>
<td>37</td>
<td>100.00</td>
<td>16</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Note. Full primary refers to schools in New Zealand with Year 1–8 students. Contributing schools have only Year 1–6 students and intermediate schools have only Years 7–8 students.

Table 2 indicates the numbers and types of participants who offered their responses in either questionnaires or interviews. Sixteen principals, 28 literacy leaders, 126 teachers (including 21 new teachers who had participated in LPDP for only one year and 24 teachers who did not experience the project at all), and 94 students offered their responses in questionnaires, semi-structured audiotaped interviews, or classroom lessons. Fifty-eight percent of teachers in the selected case schools returned a questionnaire to the researcher, with an additional 14.5 percent of teachers being interviewed instead of completing the written questionnaire.
Table 2
Number and types of participants in Study 1 schools

<table>
<thead>
<tr>
<th>School</th>
<th>Principals</th>
<th>Literacy leaders</th>
<th>Teachers 04-05</th>
<th>New teachers 05</th>
<th>New teachers 06</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reading focus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>J</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>K</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>M</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>N</td>
<td>1</td>
<td>2</td>
<td>13</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>P</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Writing focus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>2</td>
<td>13</td>
<td>3</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>O</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>16</td>
<td>28</td>
<td>81</td>
<td>21</td>
<td>24</td>
</tr>
</tbody>
</table>

Note. Numbers of teachers noted in the principal questionnaire may differ from actual numbers of the teacher questionnaires collected as not all teachers completed the questionnaire or they were interviewed instead after their lesson. New teachers in this study have been identified in two groups: those teachers who were new to the school in 2006 or those who participated in the LPDP for the final year only.

Data gathering instruments and procedures

The literature review on sustainability of professional learning and the theoretical framework underpinned the design of data collection instruments for Study 1 site visits. The final set of research instruments included:

- a questionnaire for the principal;
- a semi-structured interview schedule for the principal and literacy leaders;
- a questionnaire for all teachers in the school, including those who may not have participated in the LPDP;
- a literacy lesson observation schedule for two teachers, one who was new to the schools and a literacy leader who had participated in the LPDP;
• a semi-structured post–observation interview schedule for both teachers involved in the classroom observations;
• a semi-structured post-observation interview schedule for three students from each classroom (the teachers were asked to select three students who represented the upper, middle, and lower bands of literacy achievement in their class but did not represent the extremes).

Each research instrument probed one or both dimensions in the theoretical framework. Responses to questions would be used to determine the depth of knowledge about inquiry and coherence that schools were operating with in relation to the core beliefs and values espoused by the LPDP. Most often the questions asked of participants required exemplification, providing a critical incident method to check authenticity of their responses to actual practices.

The pre-visit questionnaire asked principals to describe what sustainability of the LPDP meant for their school and how they might articulate that with other school leaders. The researcher wanted to determine whether principals recognised the schooling improvement processes inherent in the project as being transferable. A second, closely related question asked principals how they would know if the project was being sustained in their school and what evidence they might seek for this. The questionnaire also asked principals if they had continued or changed their professional learning focus since exiting the LPDP and their reasons for doing so. These questions aimed to reveal whether or not principals were following an inquiry model of decision-making by linking their decisions about professional learning to student achievement and teacher data.

The leaders’ interview involved the principal and those literacy leaders responsible for ongoing literacy learning for teachers. This research instrument aimed to establish practitioner understandings of sustainability and the results of their sense-making of the key messages of the LPDP in particular school contexts (Coburn, 2001; Spillane et al., 2002). Firstly, what key messages were important to sustain from the LPDP? How did they plan to continue their own literacy learning? Did school leaders value the same conditions for sustainability as project leaders? Did they value adaptation of the original reform, or fidelity to the structures and learning, as a sign of success?

A second focus for the leaders’ interview was the external challenges to ongoing improvement. What was the exact nature of staff and leadership turnover? What were the new or shifting learning priorities in the school? How were new staff recruited and
inducted in relation to the reform? What, if any, continuing external activities had been
promoted by regional and national agencies and undertaken by schools to support
ongoing learning and improvement? What resilience was there in the school to external
barriers to sustainability?

Thirdly, the leaders’ interview aimed to determine if the schooling improvement
processes implicit in the LPDP were being continued or strengthened, particularly those
processes that supported schools to be self-regulated in their inquiry around evidence
(Copland, 2003). Had schools deepened their inquiry into teaching and learning by
continuing to investigate the things that matter? For example, did schools look further
into learning for students who did not progress adequately after the first inquiry? Had
they used their data to investigate what teachers still needed to understand and learn?
How did leaders continue their action plans and leadership practices for ongoing
improvement?

The LPDP theory of action for co- and self-regulated inquiry involved systematic
reflection, including supporting leaders to build their knowledge of the nature of teaching
practices that were being deployed across individual classrooms and the effectiveness of
those practices. Therefore, the school visits in Study 1 focused on all school-wide
elements of inquiry, including whether theories of action articulated by the principal and
literacy leaders might be confirmed by school documents and teacher meetings and how
well these reflected inquiry as a core structural and knowledge-building element of the
LPDP. Had the professional development experience impacted on the way collaborative
professional learning was now being approached and how new learning was undertaken?
These questions probed understandings of evidence-based inquiry related to their
leadership of each step of the teacher inquiry and knowledge-building cycle as outlined
in chapter 2.

A further focus across all instruments was the notion of coherence of effective
instructional practices. Did leaders recognise, and explicitly make clear for their teachers,
links between the professional learning projects that they engaged on? Did teachers
understand the theoretical links about effective pedagogies across the curriculum? Did
students have any notion of these connections across subjects?

The teacher questionnaire asked about their understandings and practices in
relation to co- and self-regulated inquiry and coherence of effective instructional
practices. In particular, the questionnaire probed the perspectives that teachers held about their ongoing use of the knowledge and practices that they had engaged with while they participated in the LPDP. The questionnaire addressed how they gathered and used student achievement data, if and how they continued to learn about literacy practices and what similarities and differences they could discern between their learning in different professional learning projects that had participated in. New teachers to the schools were also included in the questionnaire so that any differences and similarities in their responses about sustainability could be analysed.

In each school, a literacy leader and a teacher new to the school were interviewed after a literacy lesson observation and asked about: their understandings of the core beliefs and messages of the LPDP and how these might have been evident in their lesson; the purpose of their assessment practices; and the links, if any, they were making to other areas of the curriculum. Each of the 32 teachers was asked about students they found really hard to progress and what they thought needed to happen to raise the achievement levels of these students. The purpose of this question was to establish whether the LPDP had impacted on teachers’ beliefs about their own influence on student achievement, or whether they looked to external factors, such as the support of parents and caregivers and additional funding for specialist literacy interventions, or felt that student motivation was at fault for low achievement patterns.

The lesson observation aimed to substantiate if teachers applied the effective literacy practices that were promoted in the LPDP, particularly those practices related to supporting students to use self-regulated learning and to establishing connections between their literacy learning and learning in curriculum areas. The LPDP classroom observation schedules required teachers to link students’ literacy learning to their needs, to make links to prior knowledge, and to provide effective feedback. For example, how did the teacher support students to understand the specific learning for the lesson, and was this linked to student needs and strengths? Did they support students to self-monitor by developing success criteria? Did teachers require students to reflect on where they were at and what they needed to do next (Paris & Winograd, 2001; Timperley & Parr, 2009a)?

5 Learning Media provided a manual for all facilitators in the LPDP and classroom observation schedules were included.
The three student interviews that followed each classroom observation asked students about what they understood to be the learning objectives, how well they thought they had learned these objectives, how they knew this, and what were their next learning steps. These questions for students linked to notion of self-regulation as outlined in the theoretical framework described earlier in this thesis.

The school’s literacy action plan, any written analysis of student achievement data used with the board, school community, or teachers and any school-developed resources to support teachers’ literacy knowledge and practices were requested to examine the schools’ understandings of ongoing inquiry into issues about achievement and any approaches to coherence of effective instructional practices that may be occurring in the school. Schools also taped a teacher meeting that focused on literacy and student achievement so that leaders’ and teachers’ understandings of inquiry as a means for improvement could be analysed. These were either staff or syndicate meetings that were a scheduled part of their professional collaboration.

Procedures

Mixed methods were used to collect data, including interviews and questionnaires, and audiotaped recordings of lessons and teacher meetings. Student achievement data, and other documents relating to their analysis of that achievement data, were gathered from each school about their specific literacy focus for the LPDP. The data collection instruments for interviews, observations, and questionnaires were developed and trialled in two schools by the principal researcher and were then refined. Both schools had participated in the first cohort of the LPDP but were not included in the sample invited to be part of Study 1. The research assistant was trained during the trial and observed in the first research school to ensure that the instruments would be used in a valid and reliable manner.

Instructions for introducing the data collection instruments to participants accompanied each tool to ensure consistency of data gathering across schools. The research assistant visited each school for one day in the period October 11 to November 30, 2006. All the interviews with leaders, teachers, and classroom lessons were subsequently transcribed, as were the teacher meetings.

---

6 This term is used to describe a smaller unit of teachers in a school, generally a group of teachers working with students who are of a similar age.
The teacher meeting that discussed some form of students’ literacy achievement data took place on the day of the researcher visit in three schools since they were a part of the school’s regular schedule of meetings. The other schools were asked to tape their next scheduled meeting and to send this to the researcher.

Each school involved in the study had the opportunity to receive their transcript and to feed back to the researcher anything that they wanted to make clearer or have corrected. The researcher also sent the initial findings to the four case schools selected for Study 2, as well as the graphed achievement data results. These findings and results were specifically discussed in the structured interviews with the school leaders as a means to check on their interpretation of the results in relation to school-based conditions for sustainability.

**Validity and reliability of student assessment tools**

The analysis of the schools’ 2006–07 student achievement data aimed to describe whether each school was able to maintain or improve on the shifts in students’ literacy achievement that occurred while participating in the LPDP or whether the patterns of improvement had reduced.

**Instruments to measure student achievement**

The LPDP had used Supplementary Tests of Achievement in Reading (STAR) to determine the shifts in reading comprehension. STAR is a standardised test that uses “the scores of a large representative sample of students to establish Stanine Norms for each class level” (NZCER, 2001, p. 18). Because the conversion is age adjusted, the expectation is that students remain at the same stanine level over time if they continue at the same rate of progress during the year. Any shift seen in the mean stanines for each year group therefore represents improvement over expected levels of progress.

The LPDP used asTTle writing (Assessment Tools for Teaching and Learning) to determine the shifts in writing achievement from 2004–05. asTTle tests students’ writing performance over seven “deep” and “surface” features and provides normed data to enable schools to compare their students’ performance with other and similar groups of students from New Zealand schools (see www.tki.org.nz/r/asttle/; Ministry of Education, nd).
Procedures

There are some limitations related to the choice of STAR and to the data sets collected by the LPDP facilitators in 2004–05 and, subsequently, by the researcher in 2006 and 2007. STAR data were collected by the facilitators on behalf of the project developers in 2004–05. They collected data from the same students at each year level in their schools near the beginning of the intervention in 2004 and again at the end of 2005. The following comments are sourced from the project milestone once the data were gathered and at the time of its analysis:

Once guidelines for administering STAR had been established with the teachers, the administration, marking, and recording of scores seemed relatively straightforward. Some facilitators spoke anecdotally of schools using the wrong set of national data when identifying students with critical scores or determining stanines. In addition, some schools appeared not to have systems for checking teachers’ copying of scores or the addition of scores to determine a total. We did check the addition of subtest scores for the total score and then for the correct stanine. There were a few errors in teachers’ addition of the subtest scores. These were rectified before conducting the analysis. There were no mistakes in the stanine allocations. (Bareta & English, 2006, p. 20)

The 2006 and 2007 achievement data collected by the researcher were largely from a different set of students than that of 2004–05 because of the fact that intermediate schools have students for only two years and there were seven intermediate schools in the sample of 16 schools. Mostly data were collected at two or more time points, with testing at the beginning and end of the year. The data sets were requested from the schools at the end of each year of the study. They were checked for accuracy of addition of raw scores and stanine score allocation, and any errors were rectified before analysis. Phone interviews were conducted and transcribed to ascertain the checking of scores and methods of moderation.

The STAR tool appears to have a ceiling effect for Year 6 students and those students at stanines 7–9 (Parr, Reddish, & Timperley, 2007). However, the year level data were not considered separately, rather they were used as part of the school mix. Only nine of the thirteen schools which supplied sufficient data for analysis had Year 6 students in their data and numbers in Year 6 in each of the schools were not large so any ceiling effect would be minimised.
The administration, marking, and scoring of asTTle writing is complex because it relies on the markers having a very sound knowledge of writing features as described in the asTTle indicators. However, the national norming data determined an average of 75 percent exact score marking and an average dependability of 0.77 (The University of Auckland, 2003). These levels of agreement provide sufficient confidence in the reliability of marking in the conditions used within the LPDP. The data presented from each LPDP school in 2004–05 were a random sample of approximately ten percent of the students in each school, moderated and/or marked by the facilitator. Five of the six schools that had assessed their students in writing in 2006 and 2007 used similar moderation processes as to those that were applied in 2004–05, which was identified when the researcher gathered their data for comparison with the previous school cohort.

Note that the asTTle tool was still under development in 2004–05. The project milestone reported stated that “all schools used asTTle version 3 at the beginning of 2004; most schools used asTTle version 4 at Time 2, but some used the earlier version at the end of 2005” (Bareta & English, 2006). The major difference between these two versions is the level 4 indicators in the marking schedule. It is not known whether any of the five sustainability study schools that submitted writing data post 2004 used different versions of the marking schedules so the analysis was unable to take this into account. The asTTle data collected in 2006–07 consistently applied version 4 indicators.

Analysis of qualitative data

The qualitative data gathered from mixed methods have been analysed to establish how schools viewed sustainability of the practices that they has learned in the LPDP, what actions were taken to promote sustainability and how these actions reflected understandings about co- and self-regulated improvement practices and coherence of effective instructional practices. Data were aggregated across schools where the research questions aimed to examine the leaders’ and teachers’ perspectives and actions associated with sustainability, removing individual contexts from the analysis. Study 1 also aimed to examine the relative importance of particular school-based conditions in supporting sustainability of student achievement outcomes so qualitative and quantitative data were then compared and contrasted across schools to support this analysis.

---

7 School F did not collect writing data using asTTle in 2006. At the time of their interview (October, 2006), the principal indicated that one asTTle school-wide writing analysis would be completed in November but this was apparently not actioned when the researcher later requested these data and subsequent data for 2007.
The method of analysing interviews followed standard qualitative analysis. All of the interview and observation data were transcribed, and the scripts were read to elicit the range of the responses. Responses to open-ended questions were grouped into similar ideas, and direct quotes were highlighted and listed on a spreadsheet. The researcher developed coding categories using an iterative process of considering both the actual quotes and the research questions being addressed, and also of considering how the theoretical framework about inquiry and coherence might frame decisions about categories. For example, the responses that leaders gave about how they would know whether the LPDP was being sustained in the school suggested that broad categories such as ongoing teacher learning structures and using student achievement data to inform teaching might be used, whereas the theory and empirical research suggested that further fine-tuning of categories was required.

The theoretical framework developed in this research requires analysis as to whether there was an understanding of the inquiry process in the case schools, and how teachers might have used data to evaluate the effectiveness of their teaching and to problem-solve where current practices were not effective. The researcher was also investigating if schools had understood the distinction between evidence-based decision making and inquiry approaches. For instance, Lai et al. (2009) found that the higher gain school in their investigation of two case study schools had evaluated their data about the drop in students’ reading achievement over the holiday break and integrated their theories of effective literacy practices in order to determine responses to the problem. However, the lower gain school had simply applied ideas that “worked well” in classrooms rather than targeting actions to the specific problem. The coding for the research instruments therefore differentiated between evidence-based practices and co- and self-regulated inquiry practices.

Similarly, the theory being used promoted the concept of transfer of skills and knowledge about schooling improvement to other curriculum areas, so categories were added to reflect a broader view of the sustainability of the LPDP.

For documents and lesson observations, an analysis framework was developed based on the coding categories built up for the interviews. These frameworks incorporated coding related to the continuums about inquiry practices and to coherence of effective instructional practices described in chapter 2 and how these might be played out at leadership, teacher and student levels of the school. The analysis frameworks were
used, in the first instance, to check information gained in the interviews and to test if espoused practices were, in fact, theory-in-action (Robinson, 1993). Where espoused actions were not verified in this way, this was noted in the final analysis.

To avoid confirmation bias, the coding procedures took account of disconfirming evidence. For example, Table 3 illustrates how all responses to one question were coded, including those that did not necessarily align with the two advanced ends of the dimensions of the theoretical framework. The codes were designed to distinguish between, and further refine, elements of the two dimensions of coherence and co- and self-regulated inquiry, indicating a nominal (but yet untested) set of progressions towards the end of each dimension’s continua. Each question in the interviews, in the observations, and in document analysis has been linked to one of the continua, coded, and scored.

In addition, the researcher and the assistant researcher independently coded all of the responses from Study 1, including from research interviews, questionnaires, observations, and documents. Reliability of coding of responses was established by calculating the percentage of codes that were coded the same between markers. Percentage of inter-rater agreement was calculated using this formula:

\[(\text{Total number of ratings that agree} / \text{Total number of ratings}) \times 100 = \% \text{ agreement}\]

The agreement ranged between 80.5 and 100 percent across all questions and observation and document analyses. Any disagreements were then discussed and resolved.
### Table 3
*Coding of participant responses to question on use of achievement data*

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Category: Question 6 Leaders Interview: 2006: What purpose do you have for data gathering?</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No answer</td>
</tr>
<tr>
<td>1</td>
<td><strong>Reporting</strong>: Involving parents or caregivers in learning, e.g., used for teachers to be able to give feedback to parents and/or community about where students were in relation to others or to national standards.</td>
</tr>
<tr>
<td>2</td>
<td><strong>Evidence-based decision making</strong>: Classroom focused evidence or data used for finding gaps in learning, planning for needs, e.g., grouping, next steps for teaching.</td>
</tr>
<tr>
<td>3</td>
<td><strong>Involving students in learning</strong>: Used for teachers to be able to give feedback to students, e.g., to set goals.</td>
</tr>
<tr>
<td>4</td>
<td><strong>Comparison</strong>: Longitudinal data needed for tracking progress of cohorts over time, comparisons, e.g., school-wide data over time, trends, for planning and reporting to MoE, comparisons with normed groups</td>
</tr>
<tr>
<td>5</td>
<td><strong>Co- and self-regulative improvement practices</strong>: Inquiry approach or design for school, specific statements about challenging practice or justifying practice, data being used to test the efficacy of the practices they engage in, e.g., checking individuals or testing their practices against the data, looking at the data for gaps in their knowledge, focus on the data to monitor their own impact, looking for teacher effectiveness, triangulation of data</td>
</tr>
<tr>
<td>6</td>
<td><strong>Transfer</strong>: Use in other contexts or comparing curriculum areas</td>
</tr>
<tr>
<td>7</td>
<td><strong>Other</strong></td>
</tr>
</tbody>
</table>

Each school in the study has been mapped to the theoretical framework by scoring their data using the following process: Scoring occurred after the initial coding of evidence from the questionnaires, interview responses, documents and transcripts. Each coding schedule had a category of co- or self-regulated improvement practices and/or transfer of practices that represented the advanced ends of the two continua of the theoretical framework. Where responses or transcripts had been scored in this way, they were allocated a score of 1. If responses or transcripts were coded in any other categories then they were allocated a score of 0. In this way, it was possible to clearly discriminate between schools in the study by relating the scoring of evidence to the definition and theoretical framework for sustainability argued in chapter 2. In Table 3 for example, if evidence from question 6 of the leaders’ interview had been coded in 5 then the school would have scored 1 for understanding of inquiry for improvement. Similarly, if evidence had been coded in 6 then the school would have scored 1 for understandings about ways to develop coherence of instructional practices within and across curriculum. If evidence from question 6 of the leaders’ interview had been coded in any category other than 5 and/or 6 then the school would have been coded as 0 since they had some or no evidence of coherence of effective instruction or some or no evidence of co- and self-regulated improvement processes. In the teacher questionnaire, where there were
different numbers of teachers in each school, the scores were averaged so that schools were equally weighted.

The results of the scoring were totalled across the research instruments for each school and then graphed as co-ordinates on axes labelled for coherence of effective instruction and for co- and self-regulated inquiry practices. To illustrate the process, Appendix B presents the collation of the scoring across all research instruments and schools in Study 1.

Analysis of student achievement data

There is evidence that changed practices do not necessarily lead to improvement in student achievement (Timperley & Phillips, 2003; Wood, 2007). From the outset, this study aimed to examine the relationship between leader and teacher practices and the sustainability of student gains in achievement, given this relationship is not always the focus of studies related to this topic (Timperley et al., 2007).

As noted in chapter 2, the research literature on sustainability offers little guidance on what improvements in student achievement might be expected for new cohorts of students after the school exits a professional learning initiative. Nor did the LPDP set expectations for schools about their ongoing data patterns, focusing instead on embedding the processes and structures for inquiry and knowledge-building in each school. Certainly it seems unreasonable to expect shifts in achievement to improve ad infinitum over time; statistically there are limits. Without national testing, or assessment against literacy or numeracy standards, schools in New Zealand can choose from a number of assessments to analyse their students’ progress over time. Increasingly, there has been government pressure for schools to use those tests available to them that have normed populations and/or are linked to the New Zealand Curriculum\(^8\) so that they can compare their students’ performance to national means. To that end, the researcher considered the individual assessment tools that the LPDP schools were using over the period of this study and their test characteristics in order to propose reasonable benchmarks for sustainability specific to each tool.

---

\(^8\) Within the time span of the research study, the New Zealand Ministry of Education published the Literacy Learning Progressions, supported the development of asTTle in writing, reading comprehension, and mathematics, and upgraded the Progress and Achievement Tests (PATs) to encourage schools to use nationally referenced assessments. National standards in literacy and numeracy are due to be in place by 2010.
The measure of gain for students was the key issue to consider in relation to this study, particularly comparing gains in achievement between different cohorts of students. In the LPDP, the measure of gain in students’ achievement in reading and writing over two years was illustrated in a numbers of ways using aggregated data from 91 schools, including effect size for asTTle writing and mean stanine gains where STAR data were used. The LPDP analysed gains in terms of particular year levels, gender and ethnic groups using the aggregated data. The project leaders also investigated the gains made by students in the lowest 20 percent in writing achievement and also Stanines 1–3 in reading comprehension using data aggregated across 91 schools. Gains in students’ literacy achievement were not analysed for individual schools.

However, this study was interested in individual schools data, comparing student gains in achievement in reading or writing in 2004–05 with a new cohort of students in 2006–07. There was a relatively small student sample size in each case school which meant that establishing effect size would not be a valid measure to compare gains recorded while in the LPDP and then in subsequent years. Nor was it possible in every school to compare the mean shifts for all students with those of the lowest 20 percent since the student numbers were often too small for valid analysis.

In order to offset small numbers of students in particular year groups in individual schools in the study, the 2004–05 and 2006–07 achievement data aggregated whole school data only. Data collated from each writing school recorded the mean difference of each student’s test score from the asTTle norms at each test time and averaged this for the whole cohort. These data were then compared with 2004–05 data collated in the same manner from these individual schools. Any slope on these graphs indicated progress beyond that of the normed populations, that is, beyond expected progress. The question remained as to how much slope in 2006–07 was considered enough to suggest that any school had sustained their student achievement gains.

The following rationales were applied to this question. For schools that used the asTTle writing assessment9, the researcher considered their gains as a proportion of the standard deviation for the normed groups (1 standard deviation = 100 points) and judged

---

9 In asTTle writing, the average growth per year based on multi-year cross-sectional sampling is generally 25 points to 50 points in primary schools and around 50 points in secondary schools (p. 25, The University of Auckland, 2004). The standard error of measurement for asTTle writing is calculated as 15 scale points.
that if students’ net gain in asTTle scores represented more than a 0.5 standard deviation increase, their achievement results had been sustained.

Similarly, for schools that had used the STAR tests for reading comprehension over 2004–07, any increase in mean stanine for each year level meant students had scored beyond current rates of progress. When considering stanine shifts, it appears easier to make gains if students enter at scores lower than Stanine 5; alternatively it is harder to make gains if students start at a higher mean stanine because of the ceiling effect on this test already noted (Parr et al., 2007). Taking these factors into account and for the purposes of this research, sustainability of achievement using STAR data has been defined as when students have a net gain greater than 0.33 mean stanines.

The rationales for choice of these benchmarks are discussed further in chapter 5. They are reasoned decisions that acknowledge that little is known about the constraints of accelerated growth within normal classroom conditions. The difference in benchmarks between the two tools resulted from the different characteristics of the tests themselves.

The researcher was also interested in whether individual schools had improved the gains that they were making for new groups of students over time. The conceptual framework proposed in this thesis describes a key dimension of sustainability as co- and self-regulated inquiry processes where schools work in iterative cycles to build their knowledge and support new responses for those students who are persistently underachieving (Earl & Katz, 2006; Timperley et al., 2007). Improvement is therefore a key lens for sustainability in this study. The benchmarks provided one lens for decision-making about sustainability that took into account different entry levels for students and the cumulative effects of any drops in achievement over the holiday period. The data from each school that had met the sustainability benchmark could then be analysed to evaluate if there was a pattern of increasing gains being made for each new cohort of students in that school over time. This second lens would differentiate between those schools that had maintained gains according to the benchmarks for sustainability and those schools that had improved on their gains for students over time.

Chapter 5 investigates the patterns of student gains that emerged for the case schools and compares these with one other research project working with one of the same assessment tools.
Limitations and key assumptions

Study 1 is limited to a single project focus, the sustainability of the LPDP in selected schools who exited the project at the end of 2005. While the theoretical framework proposes that coherence of effective instruction across the curriculum and transfer of the improvement processes teachers learned in the LPDP to other professional learning may enhance sustainability, the researcher did not observe teachers in learning situations in other curriculum areas. However, in Study 2, the lesson observations and interviews did move beyond literacy in their orientation so that the notion of coherence of instruction could be more fully explored.

A second limitation is that this research has not tracked schools beyond three years of their participation in the LPDP and focuses only on those schools that were originally named as case study schools while in the project. Limitations of the student assessment tools used in this study have been noted in the previous section. The following assumptions are integral to the research:

- that changed practice does not necessarily lead to improvement;
- that sustainability must be measured in terms of ongoing and improved student achievement in literacy, albeit for different cohorts of students;
- that the focus for the study remains on the schools that have exited LPDP, rather than ongoing improvements to the programme itself;
- that the focus of the research is on those factors that have helped schools sustain their gains.
CHAPTER 4
RESULTS AND DISCUSSION: SCHOOLS’ PERSPECTIVES ON SUSTAINABILITY

Sustainability of educational reform is not only a nebulous notion in the research literature, its parameters are sometimes unclear to those who are most responsible for its enactment. This chapter begins by examining the perspectives of principals, leaders, and teachers involved in the study and how this compares to both dimensions of the conceptual model that is proposed in chapter 2. Qualitative data from the 16 schools have been aggregated for this set of results to enable cross-school analyses of practitioner views about sustainability. The perspectives of practitioners are particularly valued here to reveal more about the tension between maintenance of programmes and ongoing improvement that is yet to be resolved in the research literature.

Theories of action for sustainability

In this thesis, a theory is taken to mean a set of connected ideas to explain particular circumstances (Agyris & Schön, 1974; Robinson & Lai, 2006), so a theory for sustainability is similar to the ideas described by educational researchers about particular school-based conditions that promote ongoing improvement (see chapter 2). These are formal theories but, as important as these are, the tacit theories held by practitioners more often determine the actions taken in particular situations (Argyris & Schön, 1974). Theories of action are situational in that they link consequence to particular actions in particular contexts. They depend on a set of assumptions that may not be obvious or fully described. Theories of action include espoused theories and theories-in-use. Espoused theories are those that are articulated as the reasons for particular actions or beliefs but they may not be directly observed in the context of the study. Furthermore, an individual may not be aware of any incompatibility between their espoused theories and their theory-in-use (Robinson & Lai, 2006).

The research instruments used in Study 1 aimed to uncover the tacit nature of both espoused theories and theories-in-use in regard to sustaining shifts in practices and gains in students’ literacy achievement made whilst in the LPDP. Examining theories-in-use for the purpose of evaluating their effectiveness requires observation of actual practices and analysis of documentation rather than reliance on explanations of the theories of action. This chapter deals with the latter and reaches some conclusions around
the espoused theories about sustainability held by principals, schools leaders, and teachers.

Principals were asked for their reasons for continuing or changing professional learning priorities in the year after exiting the LPDP in order to evaluate if their participation in the LPDP had impacted on their decision-making. They were also asked to explain their beliefs about sustainability and what they might talk with their colleagues about in relation to sustainability of the LPDP. All participants were asked to describe the key messages they had gained from the project and what indicators they might see if these messages were being sustained in their school. This analysis further explored the notion of instructional coherence as described in the working definition of sustainability offered in this research. Leaders and teachers were asked to describe, if any, the links that they were making between the ideas in the LPDP and the other key professional learning projects that they had been, or were currently, participating in. How these espoused theories impacted on practice and on student outcomes is reviewed in chapter 6.

**Principals’ espoused theories of action**

The 16 schools in the sustainability study all exited the LPDP in December 2005. They, like all New Zealand schools, can choose which national professional learning projects they might want to participate in and, to a certain extent, they can choose when they may want to join that project, depending on provider capacity to meet demand. That is, participation in such projects is not mandatory; the self-governing nature of New Zealand schools determines when and if they participate. Table 4 summarises the choices that the 16 case schools made in the year after their participation in the LPDP and the reasons given by principals about their decision to either continue their literacy focus beyond the formal participation in the project, or to move on to another curriculum focus for development. Shifting professional learning priorities are often cited as a constraint on sustainability both by researchers and practitioners (Hargreaves, 2003; Honig & Hatch, 2004; Okey, 2006). In effect, these decisions marked each school’s initial stand on sustainability and how they would approach grafting the LPDP onto previous and new learning.

---

10 The Education Act, 1989, set up self-managing schools in New Zealand, with devolution of considerable responsibilities to school boards of trustees. The National Education and Administration Guidelines that supported this act of parliament gave each board flexibility to prioritise areas of professional learning for leaders and teachers based on student learning needs.
Table 4
Continuing or shifting professional learning foci for 16 case schools

<table>
<thead>
<tr>
<th>School</th>
<th>Literacy focus in 2006</th>
<th>Reasons for continuity</th>
<th>New key focus</th>
<th>Reasons for shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yes</td>
<td>Consolidation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Yes</td>
<td>Student needs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>No</td>
<td></td>
<td>Numeracy/Gifted and Talented</td>
<td>Availability</td>
</tr>
<tr>
<td>D</td>
<td>No</td>
<td></td>
<td>ICT</td>
<td>Availability/confidence</td>
</tr>
<tr>
<td>E</td>
<td>Yes</td>
<td>Consolidation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Student needs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>No</td>
<td></td>
<td>Gifted and Talented</td>
<td>Availability</td>
</tr>
<tr>
<td>G</td>
<td>Yes</td>
<td>Consolidation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>Yes</td>
<td>Student needs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>No</td>
<td></td>
<td>ICT/PD</td>
<td>Availability</td>
</tr>
<tr>
<td>J</td>
<td>Yes</td>
<td>Student needs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>No</td>
<td></td>
<td>AtoL/Integrated curriculum</td>
<td>Transfer/staff needs</td>
</tr>
<tr>
<td>L</td>
<td>Yes</td>
<td>Consolidation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>No</td>
<td></td>
<td>Numeracy</td>
<td>Availability/student needs</td>
</tr>
<tr>
<td>N</td>
<td>No</td>
<td></td>
<td>AtoL/Numeracy</td>
<td>Staff needs</td>
</tr>
<tr>
<td>O</td>
<td>Yes</td>
<td>Consolidation/resilience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>Yes</td>
<td>Consolidation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td></td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Note. AtoL denotes the national Assess to Learn professional development programme. ICT PD is the national Information and Communications Technology project.

Over half of the schools in the sustainability study continued with literacy as their major professional learning focus in the year following their exit from the LPDP. This is a similar finding to that of the exit survey for all schools in the LPDP (Baretta & English, 2006, p. 100). Over 50 percent of the reasons given by principals for doing so indicated that they wanted to consolidate or “embed” the new literacy knowledge and practices that teachers had gained and that they needed this time to ensure that these practices were consistent across the school. For example, “we needed to ensure all teachers were implementing agreed practices” (Principal, School O). Some of the comments from principals categorised as “consolidation” in this analysis indicated that they needed another year to deepen the learning that had taken place in the project. For example, one principal expressed this as “embedding literacy practices further into [the] teaching and
learning culture” (Principal, School A). Only one of those nine schools continuing with their literacy focus mentioned a factor related to resilience, such as the threat to sustainability of teacher turnover. That school indicated that they had to bring new staff “on board.”

Just under one-third of the reasons, whether for continuation or change in focus, indicated that the schools had taken into account their students’ literacy achievement data in their decision. For example, one school that had a reading focus in the LPDP indicated they needed to continue their literacy focus for another year as “we also identified writing needs and gaps for all children” (Principal, School B). Similarly, School H identified student achievement levels being “still low” as their reason for continuing. Given the emphasis that the LPDP placed on student achievement data supporting all decisions for instruction and professional development, this low percentage would fall short of expectations.

Seven of the schools shifted their major teacher learning focus to other priorities, while still keeping literacy as an ongoing minor focus. Five of these seven schools cited availability of the professional development as the reason for their change of focus, with only one school adding a reference to student learning needs in that choice. The principal of School M commented that “data here suggests how urgent this contract [Numeracy] is.”

Mostly, though, their decisions were related to timing or a new area of interest. Some schools were asked to join a professional learning cluster in another curriculum area. Typical responses included that this cluster learning opportunity “was an area we had not covered” (Principal, School F) or that “we had been approached by other schools in our cluster to join them . . . most of our contributing schools had been involved in numeracy . . .we had focused on literacy since 2004” (Principal, School D). The latter principal was confident that the school had enough expertise to support new staff and sustain literacy practices.

That over half of these 16 schools did decide to continue their focus indicated an awareness of their need to continue to craft sustainability in their schools. Only three of the seven schools that did move on to a new curriculum area focused solely on availability as their reason. The others had either an overview of the needs of their staff and their students and believed this must dominate in their decision to move on or they
had confidence that their teachers would continue to maintain practices in literacy alongside a new focus for learning.

Principals were asked to describe their views of sustainability of professional learning and what this might mean for their school after exiting the LPDP. The coding categories incorporated ideas about inquiry practices and ideas related to coherence of effective instruction, that is, consistency of literacy teaching in the school and transfer of teacher pedagogical practices across the curriculum. Table 5 synthesises 69 “descriptors” of sustainability obtained from the questionnaire responses of principals across all of the 16 schools in 2006. The coding categories for this analysis were derived from various theoretical findings outlined in chapter 2. For example, a focus on learning activities and/or structures to support sustainability is described in the work of researchers such as Fullan (2006), Hargreaves and Fink (2004) and Coburn (2003). Evidence–based decision making and inquiry processes are closely examined in research by Timperley and Phillips (2003), Timperley et al. (2007) and Earl and Katz (2006). Leadership and ongoing learning is central to arguments about improved student outcomes presented by Robinson (2007), while other researchers look at ways that leaders are able to craft coherence of instructional practices across the school (Honig & Hatch, 2004; Newmann et al., 2001) and practices to offset the impact of teacher turnover and reduced funding (Cuban, 2002; Knight, 2005).

Table 5
Principals’ descriptors of sustainability

<table>
<thead>
<tr>
<th>Categories of descriptors</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning activities and/or structures</td>
<td>20</td>
<td>28.99</td>
</tr>
<tr>
<td>Consistency of effective practices</td>
<td>14</td>
<td>20.29</td>
</tr>
<tr>
<td>Evidence-based decision making</td>
<td>13</td>
<td>18.84</td>
</tr>
<tr>
<td>Achievement data-referenced</td>
<td>10</td>
<td>14.49</td>
</tr>
<tr>
<td>Co- and self-regulated improvement practices</td>
<td>3</td>
<td>4.35</td>
</tr>
<tr>
<td>Issues to be mitigated, e.g. teacher turnover</td>
<td>3</td>
<td>4.35</td>
</tr>
<tr>
<td>Accessing external support</td>
<td>2</td>
<td>2.90</td>
</tr>
<tr>
<td>Transfer of skills or knowledge</td>
<td>2</td>
<td>2.90</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>1.45</td>
</tr>
<tr>
<td>No response</td>
<td>1</td>
<td>1.45</td>
</tr>
<tr>
<td>Developing leadership</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Principals in Study 1 typically described the learning activities and structures for teachers as their indicators of sustainability. Almost 30 percent of the descriptors related to the ongoing use of: teacher observations and feedback, modelling of effective practices, professional readings, teacher goals for professional learning, and teacher reflection. The following comment illustrates this predominant focus on maintenance of the LPDP, in particular the structures for learning that they had engaged in whilst part of the project:

[For teachers to] be observed/get feedback – coaching model. Ongoing emphasis on literacy as a school-wide focus . . . setting aside specific staff meetings each term to focus on literacy. (Principal, School D, 2006)

The use of effective literacy practices represented just over 20 percent of principals’ descriptors of sustainability, describing one or all of those practices that the LPDP promoted as effective in raising achievement: linking to students’ prior knowledge, explicit use of learning intentions and success criteria, providing specific feedback to learners, sharing quality literacy models with students, and catering for diversity. The language used by principals was about such practices remaining “evident” and achieving “consistency” across classrooms. For example, the principal of School J responded that he expected to see “learning intentions based on evidence” and that “teacher lessons show effective practice” and “evidence of deliberate acts of teaching.”

The principal of School A was aiming for “consistency of practice across the whole school.”

Almost 20 percent of the descriptors of sustainability discussed a needs-based approach to planning for learning, with classroom teachers using student achievement data to identify the needs of their students so that they could “group” for literacy instruction. Often these responses were limited to the assessment tools used in the LPDP, such as STAR, running records, and asTTle. These responses, together with the 14.5 percent of descriptors from principals that referenced sustainability to ongoing improvements in student achievement, suggest that the notion of using data for making...

---

11 The term “deliberate acts of teaching” comes from Effective Literacy Practices in Years 1 to 4, a teacher handbook distributed to all New Zealand schools by the Ministry of Education. The deliberate acts of teaching include modelling, prompting, giving feedback, questioning, telling, explaining, and directing. It was the key text for teachers in the implementation of the LPDP.

12 Dr Marie Clay developed running records as a means of analysing students’ reading strategies. They provide rich information about students’ decoding and meaning-making strategies and fluency.
decisions at school and classroom levels was again represented in only a third of the descriptors.

Interestingly, only two responses described sustainability in terms of accessing ongoing external support. The other principals appear to have moved away from an externally driven model of sustainability, confident that their school could continue to drive any further work in literacy from within. This idea of “going it alone” is consistent with the findings from the LPDP exit survey (Bareta & English, 2006, p. 99). However, it does not necessarily align with the notion of evidence-based inquiry, whereby expert help may be needed when the data reveal the need for knowledge that cannot be accessed within the school (Lai et al., 2009).

The themes from the principals’ responses were consistent with those found in a similar survey by Fullan and Sharratt (2007) of 70 school principals in York Region School Board in Ontario, also focusing on sustainability of their literacy initiatives. In the Canadian study, principals wrote about shared beliefs, distributed leadership of professional communities, data-driven decision making, resources, and home–school relations. However, in contrast to the models for schooling improvement espoused by the LPDP (Bareta et al., 2006), many of the ideas about sustainability described by principals in the study did not explain co- and self-regulated improvement practices associated with having an “inquiry habit of mind” (Earl & Katz, 2002; 2006), distinct from those that described evidence-based decision making in the school. Only three responses from principals clearly described an inquiry design in the school as evidence of sustainability, where student achievement data were used to test the efficacy of the practices that teachers engaged in and/or to highlight the gaps in their own knowledge. These were clearly distinguished from the more common descriptors about using evidence to support “next steps.” Challenge and justification of teacher practices were qualities judged important in this category. For example, the principal of School B explained that “we continue to identify gaps in our knowledge”, and School L’s principal wrote that “teachers continue to inquire into their own practice”. Typically the responses to questions in the principal questionnaire were no more than two to three sentences, so this complex notion of self-regulated improvement may not have been fully disclosed by principals. When leaders were subsequently interviewed with similar questions, they did enlarge on these practices, although this category was still well below what might be
expected, given the emphasis on inquiry-based approaches in the project milestone reports and project papers (Bareta & English, 2006, 2007; Bareta et al., 2006).

Few principals described sustainability in terms of factors that might work against the achievement shifts that they had gained during the implementation of the project. Only three of the 69 descriptors mentioned the need to bring new staff on board with the key ideas and practices. For example, one principal recorded the following:

Sustainability depends on several factors – the core of teachers remaining at the school, not taking on too many other initiatives. Budget allocation to release teachers to observe. (Principal, School D, 2006)

*Instructional leaders’ espoused theories of action*

The leaders were interviewed as a group in each of the schools over one to two hours, so their responses were often more substantial than those recorded in the principal’s questionnaire. The principal in all but one case attended these interviews. The interview included instructional leaders during the LPDP or those who had since become literacy leaders in the school. The schools’ leaders’ responses replicated the self-limiting notions of sustainability captured in the previous data.

The leaders were asked what they thought were the big ideas or main messages that the school took up from the LPDP in order to assess how the project messages were being mediated by school leaders. Their responses were coded into the categories indicated in Table 6. The categories in this table were specifically derived from the key messages espoused by the LPDP through project milestones and conference papers, as well as research described in chapter 2. For example, the categories of student awareness of their literacy learning and a focus on underachievement is promoted in project tools such as the “Progress through the Phases” document mentioned earlier (Bareta & English, 2006), and are also a key dimension in the literature about self-regulated learning for students (Black & Wiliam, 1998; Paris & Winograd, 2001).
Table 6
Leaders’ views of the key messages of the LPDP

<table>
<thead>
<tr>
<th>Descriptors of key messages of LPDP</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key ideas about literacy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evidence-based teaching</td>
<td>13</td>
<td>24.53</td>
</tr>
<tr>
<td>Teacher pedagogical and content knowledge</td>
<td>8</td>
<td>15.09</td>
</tr>
<tr>
<td>Effective teacher literacy practices</td>
<td>9</td>
<td>16.98</td>
</tr>
<tr>
<td>Student awareness of their literacy learning</td>
<td>4</td>
<td>7.55</td>
</tr>
<tr>
<td>Focus on literacy underachievement</td>
<td>3</td>
<td>5.66</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>53</td>
<td>100.00</td>
</tr>
</tbody>
</table>

| Descriptors of key messages of LPDP                                    |        |       |
| **Key ideas about schooling improvement**                             |        |       |
| Collaborative professional learning structures                        | 7      | 13.21 |
| Self-regulative improvement practices                                 | 4      | 7.55  |
| Developing leadership                                                 | 1      | 1.89  |
| Beliefs                                                               | 4      | 7.55  |
| **Total**                                                             | 53     | 100.00|

Similarly, the need for deep teacher pedagogical content knowledge and processes for professional development that penetrate teachers’ beliefs about teaching and learning are found in the research literature to support sustainability (Donovan et al., 1999; Parr & Timperley, 2006; Timperley et al., 2007).

These leaders’ responses were divided into two dimensions for this analysis: those about literacy teaching and learning, and those related to schooling improvement. This split was intended to investigate whether the leaders would identify this difference and indicate where any transfer of the schooling improvement processes might occur. The ratio of those responses categorised as literacy-focused ideas to those categorised as schooling improvement responses was 2.3 to 1, though none of the leaders made this difference explicit. The schooling improvement descriptors were not always limited to the terms of a literacy context and, without further questioning to confirm this, some leaders may have implicitly intended transfer of these ideas to other areas of the curriculum. Their responses were wide-ranging and did not appear to be retrieved from any single source, such as the LPDP “Progress through the Phases” document, which refers to sustainability as a “habitual inquiry into the effectiveness of the literacy practices” and extensively describes what that evidence may look like. Reportedly,

---

13 This project tool, used with schools in the LPDP, described progressions in relation to the school-related outcomes of the project. See the section on context in Chapter 1.
facilitators and leaders used this at regular intervals to assess progress over the two-year period of the LDPD (Bareta & English, 2006).

The schooling improvement ideas, like those offered by the principals, focused mostly on the processes and structures set up to support the professional communities of practice in the schools. One school described the following organisational features of their learning community:

We’ve videoed ourselves really and taken a portion of that to a staff meeting just to look at a specific part of the lesson. Or we’ve actually been in and observed other teachers. We’ve had systems going where people can go in and observe another teacher’s writing lesson. (Leader, School I, 2006)

The LPDP project leaders first picked up this emphasis on process rather than content in their own analysis of project outcomes for cohort 1 (Bareta & English, 2006). *Professional Learning and Development: Best Evidence Synthesis Iteration* also warns that processes alone will not secure sustainable improvement growth (Timperley et al., 2007). Only seven percent of the leaders’ responses could be categorised as describing co- and self-regulated improvement processes in action. For instance, one leader explained:

You have the teachers who are really keen to get on board, they can see the results through the STAR testing. You also have the teachers who you’ve actually got a lot of work to do, working alongside and sort of moving them forward . . . looking at the sort of things that I want to look at today – “These students moved on; what did I do? What deliberate acts of teaching did I actually perform that made a difference to that particular student?” and on the other side of the coin – “What didn’t I do so that students actually didn’t move on?” . . . and it’s a really good indication to see which classrooms are actively still doing guided reading programmes and there’s a very close link to the results and to those teachers who didn’t take it on board very easily in the two year contract and that we are still trying to move forward. (Leader, School N, 2006)

Clearly, this school understood that they could use their achievement data both as a touchstone for effectiveness and as a lever for change, offering ongoing support and challenge for their teachers.
Interestingly, leaders did not identify developing their own or others’ leadership capacity in instructional literacy as a key idea being promoted in the LPDP. While research suggests that teachers’ literacy pedagogical and content knowledge is critical in raising achievement (Parr, Timperley, Reddish, Jesson, & Adams, 2006), leaders did not commonly identify this as a “main message” of the professional learning project. Only 15 percent of the descriptors identified any professional knowledge, particularly the learning that teachers did around the asTTle indicators for writing and the strategies they might need to teach to support students in reading comprehension. It may well have been a tacit understanding that they would expect all professional learning to include this focus. If this category is combined with two other categories, those of effective teacher literacy practices and student awareness of their participation in literacy learning, then together they represented almost 40 percent of leaders’ descriptors of the main messages in the LPDP. The following two responses are typical of these categories:

The structure of the lesson really, I suppose, making sure that . . . the children are very, very aware of what you’re trying to do in a particular session, maybe the big picture intention, but also the specific learning intention for that particular day. Using quality models so that students have a model that can either be unpacked or shared. Success criteria. Co-constructing success criteria, maybe based on the model. (Leader, School I, 2006)

Importance of making sure that the learning intentions, success criteria were actually buried in the heart of planning and that the children could articulate those back and understand and use them. (Leader, School E, 2006)

So practical teaching strategies were still more prominent in leaders’ minds than evidence-based literacy teaching. Only one-quarter of the fifty-three descriptors described a key message of the LPDP as using literacy achievement data to find the gaps in student learning, to plan for these needs, and to select instructional groups. For instance, the following comments typified the talk:

---
14 This is a similar finding to that in the evaluation of the Literacy Leadership Programme, an earlier national project that did not demonstrably succeed in improving student outcomes as intended (Timperley, Parr, & Higginson, 2003). The LPDP was designed using evidence from this evaluation.
As a classroom teacher, using the asTTle task for example, using that evidence to look at individual needs and being able to group students according to these needs . . . . So actually a change from taking writing as a whole class to using group structure as well . . . . And the analyses of the data, not the data themselves. To give the direction and the focus. (Leader, School A, 2006)

I think that the main message was about making your teaching very specific, making your teaching well planned and focused according to the evidence, the evidence of whole school testing, or evidence from what your students are giving you in the classroom. (Leader, School I, 2006)

Only five percent of the responses identified that the project was asking schools to focus specifically on literacy underachievement. The LPDP has often articulated this particular focus (Bareta & English, 2006; Bareta et al., 2006), and leaders in the study often talked about their teachers monitoring “target groups,” but it was unclear from their interview transcripts who were represented in these groups.

Leaders were then asked to identify what signals they might look for in order to determine if these “key messages” were being sustained in their school. Table 7 collates these responses. Not surprisingly, their main focus as instructional leaders was on classroom practices and the learning activities and structures for teachers that they had led during the implementation phase of the project. Again, a quarter of the indicators mentioned using achievement data as an indicator for sustainability, for instance, “We should see, really increasing levels of achievement” (Leader, School G, 2006).
Table 7
Leaders’ indicators for sustainability of the LPDP

<table>
<thead>
<tr>
<th>Leaders’ indicators of sustainability of LPDP</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement data-referenced</td>
<td>11</td>
<td>25.00</td>
</tr>
<tr>
<td>Consistency of effective practices</td>
<td>9</td>
<td>20.45</td>
</tr>
<tr>
<td>Student awareness of their literacy learning</td>
<td>2</td>
<td>4.55</td>
</tr>
<tr>
<td>Informal teacher talk</td>
<td>2</td>
<td>4.55</td>
</tr>
<tr>
<td>Learning processes for teachers</td>
<td>12</td>
<td>27.27</td>
</tr>
<tr>
<td>Evidence-based decision making</td>
<td>1</td>
<td>2.27</td>
</tr>
<tr>
<td>Self-regulative improvement practices</td>
<td>3</td>
<td>6.82</td>
</tr>
<tr>
<td>Developing leadership</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Awareness of factors that inhibit sustainability</td>
<td>2</td>
<td>4.55</td>
</tr>
<tr>
<td>Transfer of ideas into other curriculum areas</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Other: general statements</td>
<td>2</td>
<td>4.55</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>44</strong></td>
<td><strong>100.00</strong></td>
</tr>
</tbody>
</table>

These results make an interesting comparison with earlier research into the indicators that principal and literacy leaders shared as their success determinants for the Literacy Leadership Programme, the predecessor of the LPDP that showed no gains in student achievement. Leaders in the latter study focused mostly on teacher change and much less frequently on student outcomes. If they did focus on students, the indicators were based primarily on perceptions of student achievement, not achievement data as such (Timperley et al., 2003, p. 34). Despite the frequency in the descriptions of evidence-based practices as key messages of the LPDP in response to the first question for school leaders in this study (such as the use of data for grouping and/or determining learning intentions for students), evidence-based decision-making was not described in the second question about indicators of sustainability of the LPDP in the same terms. Only one indicator gave more detail about how teachers might be using student achievement data to inform their teaching. As with the previous question, most of the leaders’ indicators of sustainability focused on how teachers were learning in informal and formal ways in their school, such as through observations and setting goals for their learning. The principal of School K explained that:

[W]e have also last term trialled peer observations where we modelled the observation sheet with teachers, and we asked them to be looking for things like planning, clear learning intentions – that kind of construction of success criteria – and so on. So there were key bullet points when they went in for observation,
and we hope that that deliberate planning and teaching of reading will be reflected in our achievement. (Principal, School K, 2006)

This school expressed a link, albeit hopeful, between the process and the required impact. The analysis of the responses to the question about what school leaders might expect to see if the LPDP had been sustained mirrored the analysis of the previous question, in that there were far fewer (only 6.82 percent) descriptions of any inquiry approaches being identified as indicators of sustainability in their school. Overall, few leaders in these schools appeared to be able to articulate inquiry practices as a theory of action in the ways that the LPDP project leaders expressed them in their milestone reports over this period.

*Teachers’ espoused theories of action*

Teachers were also asked in a questionnaire what they regarded as the key messages from the LPDP and the same coding was applied to their descriptors. The teachers predictably focused more at the classroom level than leaders, and their responses revealed even more starkly the dilution of the LPDP’s key ideas about effective teaching and learning than those of the school leaders.

Table 8 lists the descriptor categories for 81 teachers who were project participants in both years of the project, for 21 teachers who were new to their school in 2005, and for 24 teachers who arrived in their school in 2006. In accordance with other research about spread of educational ideas (Spillane et al., 2002; Timperley et al., 2003), the less direct experience teachers had with the project facilitator, the less they described the key messages espoused by the LPDP, particularly those related to the larger ideas about schooling improvement, such as collaborative professional learning or inquiry as a process for improving student learning and achievement.
Table 8  
*Teachers’ views of key messages of the LPDP*

<table>
<thead>
<tr>
<th>Descriptor categories of key messages of LPDP</th>
<th>Teachers 04–05</th>
<th>%</th>
<th>NTO5</th>
<th>%</th>
<th>NTO6</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Literacy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General statements</td>
<td>19</td>
<td>13.67</td>
<td>2</td>
<td>6.67</td>
<td>6</td>
<td>18.75</td>
</tr>
<tr>
<td>Evidence-based teaching</td>
<td>36</td>
<td>25.90</td>
<td>10</td>
<td>33.33</td>
<td>4</td>
<td>12.50</td>
</tr>
<tr>
<td>Teacher pedagogical and content knowledge</td>
<td>5</td>
<td>3.60</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Effective teacher literacy practices</td>
<td>56</td>
<td>40.29</td>
<td>16</td>
<td>53.33</td>
<td>18</td>
<td>56.25</td>
</tr>
<tr>
<td>Student awareness of their literacy learning</td>
<td>4</td>
<td>2.88</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Focus on literacy underachievement</td>
<td>4</td>
<td>2.88</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Schooling Improvement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaborative professional learning structures</td>
<td>6</td>
<td>4.32</td>
<td>2</td>
<td>6.67</td>
<td>1</td>
<td>3.13</td>
</tr>
<tr>
<td>Co- and self-regulated improvement practices</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Developing leadership</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Awareness of factors that inhibit sustainability</td>
<td>1</td>
<td>0.72</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Beliefs</td>
<td>5</td>
<td>3.60</td>
<td>0</td>
<td>0.00</td>
<td>2</td>
<td>6.25</td>
</tr>
<tr>
<td>Transfer</td>
<td>1</td>
<td>0.72</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>No response</td>
<td>2</td>
<td>1.44</td>
<td>0</td>
<td>0.00</td>
<td>1</td>
<td>3.13</td>
</tr>
<tr>
<td>Total descriptors</td>
<td>139</td>
<td>100.00</td>
<td>30</td>
<td>100.00</td>
<td>32</td>
<td>100.00</td>
</tr>
</tbody>
</table>

The two most frequent categories were those related to effective literacy teaching and the use of data to inform planning and teaching. Sixty-six percent of the descriptors from teachers who experienced the project in both years and 87 percent from those who experienced one year mentioned the use of learning intentions, or success criteria, or using their assessment data to “diagnose students’ needs” and/or to group their students for instruction and support their next steps in learning. As many as 69 percent of those who had not participated in the LPDP during 2004–05 mentioned either evidence-based teaching and/or particular effective literacy practices as key messages of the project. Therefore, both these messages appear to be embedded, albeit not thoroughly, particularly for those teachers who did not participate in the project.
This response from a teacher who was new to the school in 2006 indicated that the messages about evidence-informed practices had survived and appeared to be a part of the fabric of her teaching:

Needs based teaching (needs taken from STAR). A clear plan of what you are teaching and why… clear learning intentions. (New teacher, School L, 2006)

Clearly teachers were thinking about literacy, and not any larger notion of schooling improvement, when articulating ideas about this project and its impact on their teaching and work.

Coherence as a theory of action for sustainability

A key dimension of the theoretical framework being used in this thesis is how schools might craft sustainability by fending off the demands of constantly changing priorities and/or aligning new professional learning with what has been learned about effective practice already. When principals were asked about why they continued with literacy as a key professional focus or changed to another, only one comment captured a sense of instructional coherence (the transfer of content and processes between curriculum areas) in his/her reason for changing priorities to the Assess to Learn contract, stating, “[We] need to refocus on assessment to improve planning and learning.” (Principal, School K, 2006). Another principal indicated that he had differentiated the responses to the professional learning needs of his staff so that one-third of the staff would continue with a reading focus, another third would move on to writing, and the rest would be involved in an Assess to Learn project since they “had not been through ABLE [Assessment for Better Learning] or AtoL in their career” (Principal, School N, 2006).

Similarly, only a very small number of principals articulated the idea that sustainability of the LPDP might be conceptualised as anything broader than the context of literacy; that the processes and key ideas might be transferred to other curriculum learning areas. Less than three percent of responses in Table 5 were categorised as “transfer of skills or knowledge.” School A’s principal recorded that “systems/processes set up during Literacy Project will continue to run as a matter of course, regardless of school focus.” The “transferring professional knowledge from [the] literacy development to other curriculum areas” was also described by the principal of School O. This same principal also exemplified the notion of instructional coherence, as described in this
research, by indicating that sustainability is about “linking professional knowledge from the literacy development to our school’s teaching levels, values and quality schools criteria.”

Leaders were asked to nominate any links they were making between the LPDP and professional development projects they had recently participated in or were currently involved with. Table 9 categorises the links that leaders offered in response to this question. Five schools nominated the Information and Communication Technologies Professional Development cluster programme (ICT PD) as one of the projects that they were participating in, but only one of these schools could nominate commonalities or connections with the LPDP. The most common connections to the LPDP were made between assessment initiatives and the Numeracy Development Project (NDP), and these were about evidence-based practices and the use of explicit teaching.

The most commonly described connection (26 percent of responses) was the emphasis the various projects each has on explicit teaching, that is, the teacher’s practice of sharing the lesson’s learning intentions and of building success criteria with their students. This explicit teaching was viewed as being based on more precise indicators of learning. For example:

the main one has been learning intentions and success criteria which isn’t a new concept, but we’ve now got it …. [and we are] far more focused on the specific [learning]. And success criteria – we’ve developed [these] with students to form self- and peer assessment [activities] …. Children are far more actively involved in goal-setting as a result of that. (Literacy leader, School K, 2006)

Almost twelve percent of the responses emphasised a connection between projects as the focus on students becoming more involved in their own learning, aware of their progress, their new learning, and their next steps. A similar percentage of the comments recognised the role of professional learning communities in each of the projects the leaders had been involved with.

---

15 A nationwide project aimed to develop teacher knowledge and raise student achievement in numeracy.
For example, the following conversation about links between the LPDP and the Numeracy Development Project took place in School H:

Leader Observations is another thing, too, staff actually asking to be observed.

Researcher And it was happening in Numeracy as well?

Leader Mm, through the action plan.

Leader And that’s that confidence factor of staff being prepared to allow people to see the good and the bad and discuss it, and I think that’s a critical hurdle to overcome. So staff feel confident enough to say “Well, hang on a minute, you know I’m not too hot here, I want to get better here,” and be prepared to front up, because sometimes it’s hidden away. It’s a huge hurdle to overcome.

(School leaders, School H, 2006)

Most leaders did not identify the use of evidence to determine next teaching steps as a connection between learning in different projects, and even fewer saw links about the use of data to measure the effectiveness of teaching practices. Just over 14 percent of the descriptors about links between projects reported non-alignment, and one even described a “clash” between the LPDP, Numeracy, and Gifted and Talented16 projects over the idea of student-led inquiry that might impact on opportunities for the direct teaching of literacy and numeracy skills and knowledge:

No, if anything they sort of clash. Isn’t that terrible? Because you know the inquiry learning thing would almost override if we did it in its pure form, it would override our programmes in numeracy and literacy, and so writing could be just contorted to whatever the inquiry was. (Principal, School F, 2006)

In this instance, the projects were not viewed as competing for time, an issue that is often raised about sustainability, but instead as competing for teaching approaches.

---

16 This is a programme focused on developing schools’ ability to support students identified as “gifted”.
Table 9
Leaders’ views on connections between professional learning projects

<table>
<thead>
<tr>
<th>Connections between projects</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>General statements</td>
<td>5</td>
<td>11.90</td>
</tr>
<tr>
<td>Need for teacher pedagogical content knowledge</td>
<td>1</td>
<td>2.38</td>
</tr>
<tr>
<td>Focus on teacher effectiveness</td>
<td>2</td>
<td>4.76</td>
</tr>
<tr>
<td>Evidence-based decision making</td>
<td>3</td>
<td>7.14</td>
</tr>
<tr>
<td>Explicit teaching</td>
<td>11</td>
<td>26.19</td>
</tr>
<tr>
<td>Student awareness of their learning</td>
<td>5</td>
<td>11.90</td>
</tr>
<tr>
<td>Leadership capacity</td>
<td>2</td>
<td>4.76</td>
</tr>
<tr>
<td>Learning processes for teachers</td>
<td>5</td>
<td>11.90</td>
</tr>
<tr>
<td>Other: specific links</td>
<td>2</td>
<td>4.76</td>
</tr>
<tr>
<td>Non-alignment</td>
<td>6</td>
<td>14.29</td>
</tr>
<tr>
<td>Total descriptors</td>
<td>42</td>
<td>100.00</td>
</tr>
</tbody>
</table>

A further 11.9 percent of responses implying connections between projects were termed in very general ways such as “effective teaching practices” (School B, 2006), “talking school-wide” (School J, 2006), and “resources” (School M, 2006). This general category combined with that of non-alignment means that just over 26 percent of the descriptors did not name specific connections. Overall this is not surprising because the LPDP project leaders had few formal interactions with other key national projects. School leaders may not have been supported to make connections between projects or to identify any perceived conflict between the ideas they were promoting. As Newmann et al. (2001) found, schools had been largely left to make sense of the coherence between projects for themselves.

Some leaders did mention the big ideas of the project, such as “collective responsibility for student outcomes” (School I, 2006) and “having clear expectations makes a difference” (School F, 2006). These were interpreted as expressing the LPDP’s core beliefs and clearly different from any activities or processes that might be replicated in the school. These ideas are powerful (Elmore, 2002a) but, in this study, rarely expressed or linked across projects by the participants.

Self-limited understandings of sustainability

Most principals and leaders in Study 1 described sustainability as operating at more than one level in the school: at the organisational and classroom level. However, in the main, they were concerned about “consistency” and about “ongoing use” of

91
processes. Teachers and leaders focused predominantly on literacy practices as the key message they had gained from the LPDP, but expressed this within a framework of maintenance, expecting a replica of what they experienced in the LPDP (Century & Levy, 2002). Their perspectives for sustainability were concerned with breadth rather than depth. Evidence-based practices were less emphasised, but still prominent across responses from principals, literacy leaders, and teachers. But very few described sustainability within an improvement or coherence paradigm, or even used the term “inquiry” in their responses.

Together these findings pointed to a self-limiting theory of action for sustainability – limited to one area of the curriculum and limited to maintenance of the newly developed practices. How fragile would the shifts in those practices described by leaders and teachers prove to be? Would they be enough to sustain the gains in student achievement that they had made while in the project with a new group of students? The following chapter investigates their students’ literacy achievement data in the two years following their exit from the LPDP.
CHAPTER 5
SUSTAINABILITY OF STUDENT ACHIEVEMENT GAINS

Measurement of students’ progress is a truly complex area for educators. The literature discusses problems associated with measurement such as what might be measured, what tools might be appropriate and how measurement over time is best expressed (Hattie, 1999; Lai et al., 2009; Timperley et al., 2007). As well, there is debate about what expectations are held for students’ progress in learning and how much progress might be required to meet the demands of increasingly multifaceted curricula and competencies for life-long learning (Carr, 2008; Gilbert, 2008). These choices are dependent on the nature of the problem or area of learning that is being studied and all such judgments are problematic in some way. The literature about each of these elements, with a lens on the aspect of progress, is diverse and, to a large extent, follows theories of learning (Carr, 2008). In the first instance, schools may “zoom in” to define specific areas of learning such as reading comprehension and report differences in scores over time, using norm-referenced tests such as those used in the LPDP (asTTle writing and STAR). When learning is viewed as a longer-term endeavour and when outcomes are less certain or defined, such as the outcome of learning to self-manage, then the lens is wider and the assessment less defined by a particular space and time. Such is the complexity of the nature of learning and the notion of its measurement.

That aside, schools deal regularly with micro-level assessments of individuals and cohorts to inform teaching programmes, as well as with more macro-level comparisons of achievement patterns over time. This latter activity, as indicated in chapter 3, has several limitations for New Zealand schools. For example, schools are able to choose from a number of assessment tools to analyse their students’ progress over time and each tool has different characteristics such as the nature of their norming data and growth models. This chapter further examines the literature about how to measure achievement gains for successive cohorts of students over time and reports on related findings in the 16 schools in Study 1. In the absence of substantive research on patterns for long-term improvement in student outcomes in individual schools, these findings are compared to one other New Zealand study of sustainability that involves students’ reading achievement. The final section of the chapter explores the significance of the findings for sustainability of educational outcomes.
In light of increasing demands for accountability, both world-wide and in New Zealand, schools and researchers are looking for ways to measure progress of students over time and to set expectations for improvement that indicate “added value” or outcomes that out-pace the expected gains made by exposure to schooling. This is particularly important in those contexts, like New Zealand, where there are substantial disparities in student achievement and where accelerated and cumulative gains are required for those students in the lowest percentiles to catch up with their peers. Schools in New Zealand are required, under the National Administration Guidelines (NAGs)\(^\text{17}\), to develop goals and targets for student learning outcomes and to report annually on the extent to which they have made progress. Increasingly, Ministry-funded schooling improvement projects, like the LPDP, are required to present data on the progress of students over the period of the intervention, usually collated over a whole cohort of schools and students from different time points during the intervention. Nowhere in the project materials for the 2004–05 cohort of schools were there any explicit discussions of what individual schools might expect in terms of lifts in achievement over time, especially in successive years after their exit from the programme. This is not surprising, given that this was the first set of project data on the intervention and that there were no previous patterns to draw on for discussion with facilitators or schools.

There is ongoing debate in New Zealand about measurement of progress and how to express that progress. This debate adds to the difficulty of defining what sustainability might be in terms of rates of progress in student achievement when schools exit professional learning programmes.

Three sets of expectations for learners are at work in New Zealand schools. Firstly, are students in a particular school performing at similar levels to those of the normed groups? Secondly, are students making similar rates of progress as these normed groups? Lastly, how are students performing in relation to the expectations outlined in *The New Zealand Curriculum* (Ministry of Education, 2007)? The last question is especially problematic since it relies on assessments being calibrated to the curriculum. The STAR assessment is a standardised test that used a large representative sample of students to establish stanine norms for reading skills and knowledge. It was not

\(^{17}\) The National Administration Guidelines are mandatory regulations set by government for all New Zealand schools.
deliberately developed with *The New Zealand Curriculum* (Ministry of Education, 2007) in mind. Developers of the asTTle writing assessment tool did map their tests to the curriculum levels but found that some year groups of students were not achieving at curriculum expectation (The University of Auckland, 2003a). Schools are required by the NAGs to track progress specifically for Māori students so there is increasing need to compare particular groups of students and their progress to that of other groups. There is an expectation that those groups of Māori and Pasifika students, currently overrepresented in at risk categories, shift upwards so that they match the achievement patterns of all other subgroups (Ministry of Education, 2008a; Ministry of Education, 2009).

Studies have contested the notion that schools can dismantle long standing inequalities such as the link between low student achievement and the disadvantaged social backgrounds of their students (Harker, 2004; Nash & Prochnow, 2004). More recently, however, a number of studies confirm that significant disparities for students can be impacted by effective teaching interventions (Borman, 2005; Lai et al., 2009), supporting the claim that low progress, particularly in Decile 1 schools in New Zealand, is “neither inevitable nor immutable” (Phillips, McNaughton, & MacDonald, 2002).

How achievement progress might be calculated is a particular focus for comment among researchers. For example, using simple difference scores is criticised for reliability and for not measuring progress fairly across a distribution (Smith, 2008). In these circumstances, statistical phenomena such as regression to the mean occur and may, for instance, explain negative correlations between times 1 and 2 at the top end of scores (Parr et al., 2007). There are criticisms too of particular assessment tools that have inherent flaws, also impacting on how shifts in achievement might be reported. Parr et al. (2007) highlighted a marked ceiling effect, for those at the upper levels of distribution, in data for STAR tests. There are also potential limitations in the marking of asTTle writing such as moderation processes that are not sufficiently reliable.

Certainly, there are large variations in the ways that progress is expressed. For example, Dylan Wiliam (2007) explains that the regular use of formative assessment teaching strategies raises student achievement, equating to an acceleration of six to nine months’ gain in learning in a year. However, the most common currency for reporting and comparing the magnitude of observed interventions is in terms of effect sizes. Effect sizes take into account the mean shifts in achievement and the amount of variance. For
example, the LPDP reported an effect size of 0.87 \((n = 3,787)\) for reading, as measured by STAR, for all students and 1.97 \((n = 845)\) for the lowest 20 percent of students (Timperley et al., 2007, p. 37). However, these effect sizes represented increases over baseline effects rather than taking into account expected progress, and they recorded the impact after two years of intervention. Expected progress is usually determined by comparison to normed groups of students. That is, any progress that occurs in addition to the shifts in achievement made by normed groups of students is beyond expected or current patterns of progress. This is an important distinction when reporting results, especially where the pace of progress needs to be accelerated for those students who are persistently underachieving (Lai et al., 2009).

One other pattern was noticed in the whole LPDP cohort in 2004–05. Eighty-five percent of those students who began in Stanine 1 in the STAR assessment were still in Stanine 2 or 3 after two years (Bareta & English, 2006). The external evaluation report questioned whether these “at risk” students needed specialist intervention to support classroom teachers to lift their performance (McDowall et al., 2007). While the LPDP results indicate greater gains with lower bands of students, persistent underachievement remained. It is an important point for schools as they think about sustainability of the gains made in such projects. Should schools reasonably expect to move all students out of Stanines 1–3 without specialist support?

Two other phenomena need to be noted in relation to investigating trends and expectations in terms of student achievement data – that of “implementation dip” and “summer effects”. The implementation dip is more of a behavioural effect, where those grappling with new beliefs, ideas, and practices do not have the skills to implement them fully in the first instance. This takes time and practise, with support from others (Fullan, 2001). This dip can be observed as a trend in student achievement patterns as well. For example, in the implementation phase of Comprehensive School Reform (CSR) initiatives in California, those schools involved in partnership programmes with the University of California, Los Angeles (UCLA) experienced a significant lift in the first year and then a levelling off of students’ scores in the second year of implementation, followed by another period of growth. The authors ascribe this effect to an implementation dip (UCLA, 2007) and warn that the impact of comprehensive organisational change is often not linear or incremental.
A second observable trend in longitudinal data on reading scores is what many call the summer effect, where reading scores decline over the vacation period when students are away from regular literacy instruction (McNaughton, Amituanai-Toloa, & Lai., 2007). This effect is particularly significant for those students from lower socio-economic backgrounds who are already at risk of academic failure (Mraz & Rasinski, 2007). Lai et al. (2009) observed a similar effect in their study of seven Decile 1 suburban schools in South Auckland. In addition, there are other observable patterns within particular year groups in studies that track student performance in literacy over time. McNaughton et al. (2007) found in two concurrent studies in South Auckland schools that achievement gains were less for students at the transition from Year 6 to Year 7 than for students in younger year groups. The LPDP student achievement data for cohort 1 schools also had dissimilar patterns in achievement gains for particular year groups (Bareta & English, 2006).

In summary, how progress is measured needs to be considered, as well as patterns for students’ achievement data over time, particularly during the implementation phase of schooling improvement reforms. Research studies indicate diverse impacts amongst individual schools, and differential growth patterns over time. These growth patterns may occur in a step-like fashion rather than being uniform in their progression. Progress above expected levels also needs to be considered in any approach to the measurement of student achievement over time if long-term educational goals are to be met for all students.

**Patterns for schooling improvement post implementation**

Arguably more important for this study is that few reports track student achievement after the implementation phase as schools work with new cohorts of students. This limitation makes it problematic to examine how much improvement over expected gains schools can realistically make over time. While the next section points to promising results in both this study of sustainability in 16 LPDP schools and that of Lai and her colleagues (2009), these schools are in the initial stages of sustainability, just two to three years from an intensive professional learning intervention. Fullan (2006) and also Thomas et al. (2007) ponder the long-term capability of individual schools to successively and persistently improve on gains. They provide patterns of achievement in some large-scale studies where achievement in literacy or secondary examination results has improved over a five-year period only to plateau. Similarly, Gray, Golstein, and
Jesson (1996) investigated five cohorts of students in over 30 schools in one local education authority in the United Kingdom and found that some schools improved their General Certificate of Secondary Education (GCSE) examination results more rapidly than others, but few schools in the study appeared to be able to improve results consistently from year to year.

Thomas et al.’s (2007) study, particularly pertinent for this thesis, considered whether improvement occurred in short bursts or whether it was more continuously sustained. And, if a school improved its results, what were the chances that it would continue to do so in subsequent years? They too were concerned with “value-added” – in their case taking into account students’ prior attainment and background factors when they examined successive cohorts of GSCE results in 63 schools (rather than tracking the impact of any particular professional learning intervention over time). Only four of these 63 schools improved their students’ examination results for more than five years. The modal pattern over ten years was small bursts of improvement each lasting two years.

Fullan (2006) has argued that a ceiling effect or a levelling-off in improvement scores can be explained by the set of strategies in teacher learning not being powerful enough to take on the higher levels of acceleration required by large groups of students. Statistically, there must be limits to the gains that schools can make with their students. In the case of the LPDP schools, it may be that the gains become harder won as the achievement problem becomes more complex. The theoretical model for sustainability described in chapter 2 may also explain a levelling-off effect over time, as the more easily addressed issues are solved and where deeper teacher understandings of content knowledge and co- and self-regulated inquiry practices are required to develop new responses for different and more complex problems.

Konstantopoulos and Hedges (2008) point out that while it is tempting to judge success or failure of reform efforts in terms of the problem that they are setting out to achieve, such as reducing large disparities in student achievement, eliminating these complex problems may not be a feasible goal in the short term for individual schools. What is needed are examples of what is possible in terms of progress so that learning can be gained from these and then applied. This study, along with others on sustainability, will offer a guide to what is possible with new cohorts of students until new learning finds a way to make deeper cuts into the problem.
Benchmarks for sustainability of achievement gains

In the absence of a national testing regime, schools in New Zealand are dependent on the capability of the assessment tools that they employ: whether they have normed populations to measure against, whether the tests are mapped directly to *The New Zealand Curriculum* (Ministry of Education, 2007), and what actual aspects of the curriculum are being tested. Chapter 3 outlined the characteristics and limitations of the two literacy assessment tools that were utilised in the LPDP.

Making comparisons between achievement data sets for different student cohorts in each school also had some limitations. The 2004–05 LPDP data sets were taken at only two points for the same students: at the beginning of the first year of the project and again at the end of the second year. The data gathered for Study 1 and 2 of this research were dependent on when the school decided to test their students. Only nine of the 16 case schools provided test data for their students at the beginning and end of each year over 2006 and 2007, so it was possible to gauge variation on scores between academic years with only these schools. The remaining seven schools either did not provide sufficient data or tested their students in various combinations of beginning- and end-of-year times (see Table 10). Direct comparisons between 2006 and 2007 for students in each school were therefore restricted in the latter schools because end of Year 1 test results may not be comparable to beginning of Year 2 results.

In this study, a benchmark for judging sustainability of student achievement gains was established according to the patterns of normed student populations for each of the tools used by the case schools in the study (asTTle writing and STAR), comparing the net gains made above expected progress by different cohorts of students over two-year periods. These benchmarks were devised for each assessment tool.

If the next cohort of students’ net gain in asTTle writing scores represented over 0.5 standard deviations, then it was judged by the researcher that their achievement results had been sustained. In terms of STAR data, the researcher defined sustainability of achievement as when the next cohort of students had a net mean gain greater than 0.33 stanine. Given the fact that so few “markers” for sustainability of student achievement data are defined in the literature, the rationale for both benchmarks took into account what had occurred during the intervention, the entry point data for new cohorts of students, other effects observed in similar contexts, and the two-year interval for the data being gathered in the 16 schools.
For writing data, the following factors influenced the decision to use a net gain of 0.5 standard deviations as the benchmark for sustainability over two years:

- A mean shift of 2.5 curriculum sub-levels or approximately 1 standard deviation\(^{18}\) was made in the LPDP for all writing schools \((N=49)\) over two years (Bareta & English, 2006). This data were based on the mean shift in asTTle points for all students. The data were not determining shifts beyond those of the normed groups. However, in determining shifts in the writing case schools in 2004–05 and in 2006–07 the researcher recorded the difference of each student’s test score from the asTTle norms at each test time and averaged this for the whole cohort. In this way, progress above expected levels could be judged so a benchmark below the mean shift indicated in the LPDP data was appropriate;

- The case schools also had differential entry points for new cohorts of students. For example, the writing schools had students in the second cohort entering at, on average, 0.4 of a standard deviation higher than students in their 2004–05 cohort. Given that it is often harder to make shifts with higher achieving students and that the case study data was based on progress above expected levels, it seemed reasonable to suggest the benchmark of a mean gain of 0.5 standard deviations.

The following factors influenced the decision to use a net gain greater than 0.33 stanine for STAR data as a benchmark for sustainability:

- Any increase in mean stanine level means students scored beyond their current rates of progress. A mean stanine shift of 0.56 was made in the LPDP for reading schools \((N=42)\) that included all students in all year groups in reading over two years (Bareta & English, 2006). While the ten reading case schools made a mean shift of 0.62 stanine while in the project this average was influenced by one school whose students had mean net gain of over 1 stanine. The benchmark was therefore set at a lower net stanine gain;

- There was also a large variation of entry-level differences in the case schools for reading in 2006–07. Five of the reading focus schools had, on average, students entering in 2006 with scores 0.3 stanine greater than those in the first cohort, while

\(^{18}\) In asTTle, each curriculum sub-level covers between 40 and 50 points, depending on the subject and curriculum level. (asTTle V4 Manual 1.1, Chapter 3, p. 26) A gain of 2.5 sub-levels is therefore approximately 100 points or 1 standard deviation.
four reading schools had students entering with lower scores, on average, of 0.2 stanines. As noted in chapter 3, there is a ceiling effect observed with STAR data (Parr et al., 2007). It appears easier for students to make gains if students enter at scores lower than Stanine 5, alternatively it is harder to make gains if students start at a higher mean stanine. The mean stanine entry score for the students in the 2006 cohort for all reading case schools was below 5 (4.8 stanine) and ranged from 4.3 to 6.2 stanine. Therefore, it seemed reasonable to set a lower benchmark for the new cohort of students than the mean stamine gain for all 91 schools over 2004–05;

- Statistical modelling in one New Zealand study focused on the sustainability of a three-year intervention in reading found that the typical amount of growth in mean stanine level was between 0.36 and 0.49 stanine over one academic year after the intervention. However, there were often large drops in achievement between academic years for students. For example, Year 6 student scores STAR scores reduced between 0.86 and 1.24 stanine when they were tested again as Year 7 students in the new academic year (Lai et al., 2009). For this reason, a mean stanine gain of 0.33 in a two year period would account for any cumulative effects of gains and declines over the holiday period.

As discussed in chapter 3, a second lens was applied to the student achievement data from the case schools. The benchmarks described above provided the main lens for decision-making about sustainability that took into account different entry levels for students and the cumulative effects of any drops in achievement over the holiday period. However, the researcher was also interested in gauging whether there was a pattern of increasing gains being made for each new cohort of students in these schools over time. This second lens would differentiate between those schools that had maintained gains according to the benchmarks for sustainability, and those schools that were improving on their gains for students over time. This check was related to the notion of co- and self-regulated inquiry practices outlined in the conceptual framework proposed in this thesis. Schools that understood inquiry as process for improvement were able to deepen their knowledge about effective practices in order to continue to sustain student outcomes (Timperley et al., 2007). It also relates to the notion of phases of sustainability as described by Century and Levy (2002), where professional learning programmes may be stalled at maintenance of what was developed in the implementation period and not develop the ability to adapt and improve over time. Each school that has sustained their
results according to the benchmarks were then examined for the pattern of the gains for each cohort. To determine whether schools did have a pattern of ongoing improvement, the gains for 2006–07 were referenced against the gains established while they were participating in the LPDP. Improvement in writing was judged as being above the standard error of measurement, 15 points, according to asTTle writing (The University of Auckland, 2004). For reading this was judged to be an improvement in stanine level between the cohorts. For example, if a school had sustained their students achievement results, that is they had a mean gain in reading of over 0.33 stanine or over 0.5 of a standards deviation in writing, as well as improved on their 2004–05 gains then the school was judged as having sustained and improved.

Results for 16 case schools 2006–07

Twelve of the 16 schools in Study 1 submitted sufficient data in the year after their exit to determine whether they had sustained the gains that they made while in the LPDP. Using the benchmarks outlined earlier, ten of these 12 schools sustained their student achievement results with a new cohort of students ($N=1438$ students for 2004–05 and $N=1676$ for 2006–07). In two schools, although improvement continued to be made, the rate of gain for students in 2006 did not meet the study’s benchmarks for sustainability. Four schools in the study did not provide sufficient data for analysis. For example, three schools presented data for only one point of time in 2006, so comparisons in gains could not be made. One school did not provide any data. School F did not collect writing data using asTTle in 2006. At the time of their interview (October 2006), the principal indicated that one asTTle school-wide writing analysis would be completed in November, but this was apparently not actioned when the researcher later requested these data and subsequent data for 2007. After only one year, Schools H, J, P and L had also improved on their gains compared to those made over the two-year period of 2004–05.

Interestingly, after two years these patterns did not change markedly. This time, thirteen schools submitted sufficient data to determine whether they had sustained the gains that they made while in the LPDP. By the end of 2007, 10 of the 13 schools that presented data had sustained their gains (using the benchmarks for sustainability), and three schools had reduced the rate of progress (although students had still improved beyond expected levels of progress). Three schools in the study did not present their data. Only School L had sustained their student achievement gains according to the benchmark and also improved on the mean net stanine gain in 2006–07 compared to that of 2004–05.
Table 10 summarises the results of student achievement data in the 16 case schools for the year after their exit from the LPDP and then a further year.

Table 10

Summary of student achievement results for 16 case schools 2006–07

<table>
<thead>
<tr>
<th>Case Schools</th>
<th>Literacy focus</th>
<th>After 1 year: 2006</th>
<th>After 2 years: 2006-07</th>
<th>Timing of tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Writing</td>
<td>Sustained</td>
<td>Sustained</td>
<td>BOY and EOY 06-07</td>
</tr>
<tr>
<td>B</td>
<td>Reading</td>
<td>Sustained</td>
<td>Sustained</td>
<td>BOY and EOY 06-07</td>
</tr>
<tr>
<td>C</td>
<td>Writing</td>
<td>Sustained</td>
<td>Sustained</td>
<td>BOY and EOY 06-07</td>
</tr>
<tr>
<td>D</td>
<td>Reading</td>
<td>Sustained</td>
<td>Sustained</td>
<td>BOY and EOY 06; EOY 07 only</td>
</tr>
<tr>
<td>E</td>
<td>Writing</td>
<td>Reduced</td>
<td>Reduced</td>
<td>BOY and EOY 06-07</td>
</tr>
<tr>
<td>F</td>
<td>Writing</td>
<td>No data</td>
<td>No data</td>
<td>No data received</td>
</tr>
<tr>
<td>G</td>
<td>Reading</td>
<td>Sustained</td>
<td>Sustained</td>
<td>BOY and EOY 06-07</td>
</tr>
<tr>
<td>H</td>
<td>Reading</td>
<td>Sustained and improved</td>
<td>Reduced</td>
<td>BOY and EOY 06; EOY 07 only</td>
</tr>
<tr>
<td>I</td>
<td>Writing</td>
<td>Insufficient</td>
<td>Sustained</td>
<td>EOY 06-07 only</td>
</tr>
<tr>
<td>J</td>
<td>Reading</td>
<td>Sustained and improved</td>
<td>Reduced</td>
<td>BOY and EOY 06-07</td>
</tr>
<tr>
<td>K</td>
<td>Reading</td>
<td>Insufficient</td>
<td>Sustained</td>
<td>EOY 06-07 only</td>
</tr>
<tr>
<td>L</td>
<td>Reading</td>
<td>Sustained and improved</td>
<td>Sustained and improved</td>
<td>BOY and EOY 06-07</td>
</tr>
<tr>
<td>M</td>
<td>Reading</td>
<td>Insufficient</td>
<td>No data</td>
<td>EOY 06 only</td>
</tr>
<tr>
<td>N</td>
<td>Reading</td>
<td>Sustained</td>
<td>Sustained</td>
<td>BOY and EOY 06-07</td>
</tr>
<tr>
<td>O</td>
<td>Writing</td>
<td>Reduced</td>
<td>No data</td>
<td>BOY and EOY 06 only</td>
</tr>
<tr>
<td>P</td>
<td>Reading</td>
<td>Sustained/improved</td>
<td>Sustained</td>
<td>BOY and EOY 06-07</td>
</tr>
</tbody>
</table>

Note. BOY = data gathered at the beginning of the year. EOY = data gathered at the end of the year.

Key:

- Sustained and improved – those schools that met the benchmarks set for sustainability in 2006-07 (mean gains of over 0.33 stanine for STAR or mean gains of over 0.5 standard deviations in asTTle writing) and improved on their achievement gains when compared to those in 2004-05.
- Sustained – those schools that met the benchmarks set for sustainability in 2006-07 (mean gains of over 0.33 stanine for STAR or mean gains of over 0.5 standard deviations in asTTle writing).
- Reduced – those schools that fell below the benchmarks in 2006-07.
- Insufficient data – those schools whose data were not able to be compared across two cohorts of students.

Each school’s data have been graphed to illustrate these comparisons. Appendix C contains the achievement data for each of the 13 schools for two successive cohorts of students (2004–05 when the school was in the implementation phase of the LPDP and
then again in 2006–07). Figure 3 and Figure 4 represent particular patterns that occurred over the 13 schools. For example, School L, a large suburban intermediate school, increased its students’ mean stanine in reading comprehension by 0.48 over two years ($n=206$ students) in 2004–5. In 2006 a new cohort of students, who began at approximately the same level as the previous cohort, increased their mean stanine by 0.82 within just one year (see Figure 3). The decline in achievement between academic years (time 2 and time 3) was 0.14 of a stanine and only a modest gain was made by students over the second year. Overall, though, the 2006–07 cohort of students in this school had made a mean stanine gain of 0.9, substantially improving on gains made while in the project.

![Figure 3](image-url)  
Figure 3. School L achievement data in reading with successive cohorts of students.

In the first year after exiting the programme, there were often higher rates of progress for students than in the second year. For seven of the nine schools where this comparison can be made, a levelling-off effect was evident. For example, in School P, the reading comprehension student achievement gains made in the project were sustained over the next two years, but levelled off in the second year. In 2007, students made a mean stanine gain of 0.58 compared to the mean stanine gain of 1.1 that they had made in 2006 (see Figure 4). This school also demonstrated the summer effect referred to earlier. In the holiday break between 2006 and 2007, students dropped their mean stanine score by 0.96 stanine. Four of the nine schools where this effect could be investigated
had drops in mean stanine scores of between 0.14 and 1.16. In Schools A and C, this effect occurred between 2007 and 2008.

![Figure 4. School P achievement data in reading with successive cohorts of students.](image)

Ten of thirteen schools that provided sufficient data for comparison were able to sustain their gains after two years according to the benchmarks. A pattern of levelling off in the second year was found in seven schools, indicating that these gains may remain fragile in the longer term. Schools that had sustained their gains according to the benchmark criteria did not always match their earlier rates of progress recorded in 2004–05 when they participated in the LPDP. After two years, four of the ten schools recorded small gains for their 2006–07 cohort but three of these schools had higher student entry levels than in 2004–05. Five of the schools had very similar rates of growth between the two student cohorts. School L was the only school whose students in 2006–07 made marked progress compared to the previous cohort.

**Comparative study**

Given the paucity of comparable studies in New Zealand or overseas, the Lai et al. (2009) study provided an interesting parallel with this study, sharing the same focus on measuring progress in reading, the use of the same assessment tool, and the same school system conditions. In a similar time frame to this study, their research took place over two consecutive years following a three-year intervention that involved a cluster of schools in a Ministry of Education schooling improvement initiative to raise reading comprehension levels. The seven schools in the research group were all Decile 1.
suburban schools, primarily serving culturally and linguistically diverse communities. Lai et al. tracked five year-level cohorts of students over time in the two years following the intervention. Their data had a different unit of analysis than this study, tracking year groups across the cluster of seven schools rather than within individual schools. The intervention itself had resulted in increased rates of achievement over the three years, with stanine effect sizes of $d=0.62$ (p. 5). This represented an average achievement gain, across cohorts followed longitudinally, of one year’s progress in addition to expected progress over that period. As described earlier in this chapter, these data were not continuous patterns of growth, more step-like in form, with rapid gains being offset by falls or plateaus over the summer holiday break.

Gains in achievement in the year following the intervention, as modelled for each year-level cohort, produced mean increases of between 0.36 and 0.49 of a stanine, which were virtually identical to those of the intervention itself. The authors used hierarchical linear modelling to provide a deeper analysis of differences within each cohort of students. They found that generally those students in the “low” and “below average” bands of achievement at time 1, according to the STAR test, made greater gains within the school year than those of their peers who began in the “average,” “above average,” or “excellent” bands of achievement.

Comparing these findings to the data from this study indicates some similarities, at least for the reading schools, despite differences in method and unit of analysis. Both studies found that the same intervention had various impacts on students across schools and that the summer effect was also observed. Two differences were apparent. As noted earlier in this chapter, the achievement patterns for each school in this thesis did not always show similar annual lifts as for previous cohorts and the levelling-off effect in the second year was not mentioned in the comparative study.

**Significance of findings for sustainability**

Both studies provide a promising picture for educators, indicating that sustainability can be achieved after schools exit professional learning projects, at least in successive cohorts. The benchmark criteria for sustainability devised for this research study was the first lens used for analysis. Ten of thirteen case schools managed to sustain achievement gains with a new group of students over two years. A further three schools
in the study did not provide sufficient data with which to apply the benchmark indicators for sustainability.

The second lens for analysis considered whether case schools in this sample had improved on student achievement gains with the successive cohort of students. Some, like School H, made excellent headway in the year following their exit from the LPDP in comparison to gains made 2005–05, only to lose ground in the second year. Other schools, like School A and School C, had sustained their gains, maintaining relatively similar levels of gain over two successive cohorts of students. School L was the only school in the Study 1 to markedly increase the gains for the second cohort of students after two years.

The 16 schools in this study were effectively in the early stages of sustainability relative to their exit from the intervention. However, the longer-term achievement patterns in each school are not predictable from these data, and persistent and constant improvement for each new cohort may be hard won. The broader literature base on longitudinal achievement data portrays a more fragile picture, indicating that any gains for new cohorts may level off over time.

The final chapter of Study 1 compares aspects of school contexts from the 16 schools in order to uncover particular school-based conditions that may have impacted on the variations found in these student achievement data in the same period.
CHAPTER 6
RESULTS AND DISCUSSION: INQUIRY AND COHERENCE IN 16 SCHOOLS

The frame for analysis of Study 1 data was to investigate the depth of understanding and practices about sustainability at each school and then to compare this understanding to each school’s specific patterns of student achievement outcomes, as described in chapter 5 and appendix C. In the first instance, self-limiting notions of sustainability were identified in the espoused theories of action held by most of the schools’ leaders in this study. In most cases, principals and literacy leaders had created their own boundaries around the LPDP, perceiving the project as being solely about literacy. Few made explicit links to how the knowledge about schooling improvement processes and/or effective classroom pedagogy gleaned from their participation in this project might be transferred to new curriculum contexts. Nor did many of the principals or literacy leaders in the case schools articulate their understandings about inquiry as an iterative process for improvement. This chapter uncovers the theories-in-use held by leaders and teachers in the 16 case schools as they forged their own pathways to sustainability in the two years immediately following their participation in the LPDP. The two key dimensions offered in this thesis as a theoretical framework for sustainability are more closely examined in this chapter – that of co- and self-regulated inquiry and coherence of effective instructional practices. Each of these dimensions incorporates the five school-based conditions for sustainability outlined in chapter 2 such as leadership, content and pedagogical knowledge, learning processes for teachers, evidence-based inquiry practices, and finally strategic actions that support coherence of instructional practices.

Adelman and Taylor (2003) remind us that “no single project can transform a school, never mind a school district. At the same time, any project can be the catalyst for change” (p. 20). Whether that change can be durable and still sustain an impact on students’ achievement is largely dependent on the school. Their specific contexts were examined at school-wide and classroom levels for: understandings and actions about individual and collaborative data analysis and use; how professional learning in literacy continued in the schools; what connections teachers were recognising and implementing across professional learning activities; what elements of effective literacy practices were evident in their literacy lessons; how inquiry and coherence were interpreted at the classroom level, and how students perceived their literacy learning. The interview
responses, documents, lesson, and meeting observations were coded and scored according to the method outlined in chapter 3.

The chapter begins by examining the contexts of the sample schools in relation to the factors that are often described as inhibiting sustainability of professional learning such as leadership and teacher turnover. The two dimensions of the conceptual framework are then examined in more depth. Co- and self-regulated inquiry practices are analysed at two levels – school-wide and then at the classroom level. Similarly, coherence of effective instructional practices is described at school and classroom level. The final section of this chapter compares schools’ beliefs and practices to their patterns of student achievement and then the significance of these findings are discussed in relation to sustainability of professional learning.

*Leader and teacher turnover*

Turnover of staff and changes in leadership are often cited as factors that stall sustainability by hindering work to deepen and expand improvement efforts (Datnow, 2005; Hargreaves, 2003; Gersten et al., 2000; Wood, 2007). Leadership and teacher turnover means that schools end up having to spend time retraining staff and may even feel trapped in the first year of implementation as a result (Wood, 2007). Fink and Brayman (2006) found that leadership turnover, in particular, derailed reform efforts. New principals can change the direction of professional learning to new reforms (Datnow, Hubbard, and Mehan, 2002). In addition, high poverty schools tend to be more susceptible to teacher and leader turnover and may lack resources to support ongoing change (Wood, 2007).

The decile and size of the 16 case schools are shown in Table 11. The table also includes data about turnover of staff and school leaders in 2006, the first year of the study. The sample included a range of deciles and school size. Datnow (2005) found no link between the socio-economic factors of the schools and the levels of sustainability of reform in her study of Comprehensive School Reform models in 13 schools in the United States. Similarly, no clear pattern related to the low decile status of schools was found in this study. Three out of six Decile 1–3 schools in the sample managed to sustain their student achievement gains in the two years following their participation in the LPDP, including one Decile 2 school that continued to improve on these gains for their students. At the same time, two of the three schools in the high decile range within the sample
(Decile 7-9) sustained their achievement gains with a new cohort of students according to the benchmark criteria established for this research.

Table 11  
*Case school contexts: decile, size, staff turnover, and sustainability*

<table>
<thead>
<tr>
<th>School</th>
<th>Decile</th>
<th>Size of school</th>
<th>Turnover of teachers %</th>
<th>Turnover of leaders %</th>
<th>Sustainability (using benchmarks for student achievement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>9</td>
<td>20</td>
<td>20.0</td>
<td>0.25</td>
<td>Sustained</td>
</tr>
<tr>
<td>B</td>
<td>9</td>
<td>4</td>
<td>0.0</td>
<td>0.0</td>
<td>Sustained</td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>25</td>
<td>12.5</td>
<td>9.1</td>
<td>Sustained</td>
</tr>
<tr>
<td>D</td>
<td>6</td>
<td>10</td>
<td>40.0</td>
<td>50.0</td>
<td>Sustained</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>7</td>
<td>14.3</td>
<td>0.0</td>
<td>Reduced</td>
</tr>
<tr>
<td>F</td>
<td>8</td>
<td>8</td>
<td>50.0</td>
<td>0.0</td>
<td>No data</td>
</tr>
<tr>
<td>G</td>
<td>4</td>
<td>6</td>
<td>50.0</td>
<td>0.0</td>
<td>Sustained</td>
</tr>
<tr>
<td>H</td>
<td>2</td>
<td>9</td>
<td>0.0</td>
<td>0.0</td>
<td>Reduced</td>
</tr>
<tr>
<td>I</td>
<td>6</td>
<td>9</td>
<td>11.1</td>
<td>0.0</td>
<td>Sustained</td>
</tr>
<tr>
<td>J</td>
<td>4</td>
<td>5</td>
<td>40.0</td>
<td>66.7</td>
<td>Reduced</td>
</tr>
<tr>
<td>K</td>
<td>4</td>
<td>15</td>
<td>33.0</td>
<td>20.0</td>
<td>Sustained</td>
</tr>
<tr>
<td>L</td>
<td>2</td>
<td>19</td>
<td>15.7</td>
<td>0.0</td>
<td>Sustained</td>
</tr>
<tr>
<td>M</td>
<td>1</td>
<td>5</td>
<td>60.0</td>
<td>0.0</td>
<td>No data</td>
</tr>
<tr>
<td>N</td>
<td>3</td>
<td>39</td>
<td>16.2</td>
<td>0.1</td>
<td>Sustained</td>
</tr>
<tr>
<td>O</td>
<td>2</td>
<td>13</td>
<td>15.4</td>
<td>0.0</td>
<td>No data</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>18</td>
<td>44.0</td>
<td>50.0</td>
<td>Sustained</td>
</tr>
</tbody>
</table>

Notes. See explanation of decile in chapter 2. The number of classroom teachers was sourced from the principal questionnaire, 2006. Turnover of teachers is the proportion of new teachers in the school in 2006 to the total numbers of teachers, expressed as a percentage. Turnover of leaders is the proportion of new leaders in the school in 2006 to the total number of leaders, expressed as a percentage.

Half of the 16 schools experienced a 20 percent or greater turnover in teachers immediately after their participation in the LPDP, whereas leadership changes were more static, with only four schools losing more than twenty percent of their leadership strength. These data take account of those leaders described as literacy leaders in each school and include the principal. In New Zealand, the average turnover of teachers in schools in 2005–06 was 13.8 percent for teachers and 15.8 percent for school leaders including principals, with different regions experiencing varied rates of turnover.¹⁹ Schools D, J, and P were the most impacted by high turnover in both teachers and leaders. While schools D and P managed to sustain their achievement gains in 2006–07,

school J reduced their gains for the successive cohort of students after their involvement in the LPDP. School P was a large school and the principal remained in place, while School J was much smaller and the principal had left the school by the end of 2005. School D’s principal had also been part of the LPDP and remained in the school over the period of the study. Datnow (2005) confirms that while there was more continuity of leadership in schools that sustained reforms, this was not always the case. So turnover alone is not necessarily a negative predictor for sustainability of gains made in schooling improvement initiatives, although a combination of high turnover of teachers as well as the principal in a small school such as School J may have impacted on these results.

Co- and self-regulated inquiry practices

The school visits in Study 1 focused on all school-wide elements of inquiry, including whether espoused theories of action articulated by the principal and literacy leaders might be confirmed by school documents, teacher meetings and lessons. School-wide practices are reported first, followed by related practices at the classroom level.

Responses from school leaders and teachers about the sustainability of the practices relating to analysis of student achievement data signalled that they were working in an evidenced-based paradigm, where their focus was primarily on teachers using data to support decisions about next steps for their learners. Evidence-based decision making is a central part of inquiry or self-regulatory practice but, on its own, may only lead to current practices being sustained, even if those practices are newly formed practices from recent professional learning. This distinction between evidence-based practices and an inquiry-based approach was noted in chapter 2 and was described as key idea to support analysis of interviews, questionnaires, teacher meetings and lessons in the methodology section of this thesis. The difference between an inquiry approach and simply using evidence-informed teaching strategies to solve problems of underachievement in reading was observed by Lai et al. (2009). Their study of two case study schools established that the higher gain school had reframed problems in a specific rather than general way, while the lower gain school applied teacher practices that were known to “work well” but were not necessarily related to the identified problem. Lai et al. confirmed the interdependence between student achievement data and teacher practices as fundamental to the inquiry cycle as illustrated in Teacher Professional Learning and Development: Best Evidence Synthesis Iteration (Timperley et al., 2007).
School-wide inquiry

Table 12 categorises leaders’ responses across all 16 schools about how literacy student achievement data were analysed in 2006 and what was being investigated in these data. The schools’ responses illustrate the divergence of understandings about co- and self-regulated inquiry. Nine of the forty-three responses from school leaders articulated evidence-based decision making for teachers as the predominant purpose for data gathering. Nine responses referred to data being analysed collaboratively so that teachers could understand and take ownership of the data and the process. Collectively, both ideas were reflected in just over forty percent of the responses. Comparing sets of data to support calculations about students’ progress over time is a key action when inquiring into school effectiveness and for reporting on targets set for improvement, and this was well recognised by schools.

The notion of inquiry being a knowledge-building process was not well reported in these data. There were major gaps in the leaders’ thinking about data as well as in their responses to similar questions about the purposes of school-wide and classroom data analysis. Only seven percent of the responses to this question were about asking further questions of the data or linking data to effectiveness of teaching. Therefore, variance from targets for achievement may not be questioned or linked to teacher effectiveness, which in the long term, may inhibit schools from sustaining further gains (Copland, 2003; Timperley et al., 2007).

Table 12
 Leaders’ responses about purpose and use of student achievement data

<table>
<thead>
<tr>
<th>Purpose for data analysis</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative ownership and analysis of data</td>
<td>9</td>
<td>20.93</td>
</tr>
<tr>
<td>Comparison with previous years, norms, or other schools</td>
<td>11</td>
<td>25.58</td>
</tr>
<tr>
<td>Asking further questions of the data</td>
<td>3</td>
<td>6.98</td>
</tr>
<tr>
<td>Evidence-based decision making</td>
<td>9</td>
<td>20.93</td>
</tr>
<tr>
<td>Teacher effectiveness</td>
<td>6</td>
<td>13.95</td>
</tr>
<tr>
<td>Co- and self-regulated improvement (school) practices</td>
<td>3</td>
<td>6.98</td>
</tr>
<tr>
<td>Co- and self-regulated improvement (teacher) practices</td>
<td>2</td>
<td>4.65</td>
</tr>
<tr>
<td>Transfer</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Total descriptors</td>
<td>43</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Collaborative problem-solving and using data to inquire into students’ literacy needs were found to be much more evident in the teacher meetings in each school than were examples of co- and self-regulated inquiry practices. Eleven of the 15\(^{20}\) transcripts of the meetings connected their discussions on student achievement data to the specific needs that these data generated for students. Seven transcripts had evidence that the data were being investigated and linked to lack of teacher knowledge, or to the lack of effectiveness of their practices. Only two schools had substantive evidence of using student achievement data to plan and monitor a new inquiry, including identifying new learning for teachers and/or monitoring the impact of that learning. The transcript from School I illustrated several aspects of inquiry as a knowledge-building process. The teachers were working from specific data to decide on their next focus for instruction. This conversation reflects the notion of “challenging talk” outlined by Annan, Lai, and Robinson (2003). Colleagues disputed Teacher B’s interpretation of data and then supported her to think about changes to her practices. As well, Teacher D outlined her theory about how students learn punctuation and her interest in inquiring further into these data:

Principal Is there something that you think you can focus on for the next few weeks before the end of the year?

Teacher A We did talk about the punctuation, didn’t we?

Teacher B It’s quite a hard concept for them to grasp though, knowing what a sentence is, and when we were doing full stops we’d talk about full stops every day and what happens after a full stop – a capital letter. They could verbalise it all, and we can talk about question marks and exclamation marks and things like that we’d use for the big book, but for them to use [them] in their writing, it doesn’t happen.

Teacher A Some of them are doing it. No, it doesn’t “not happen,” because it’s only – it’s 45% who are not, 55% who are.

Teacher B So they practise the full stops at the end of the page, you get a line of things like that first but, yeah.

\(^{20}\) School E’s teacher meeting transcript is missing in this data set.
Chapter 6 – Results and Discussion: Inquiry and coherence in 16 schools

Principal You’ve been grouping your students. Do you think that would be a time to work with those ones, as well as those few?

Teacher C Pick that out as a learning intention and really work on it, and if you need support, you could . . .

Teacher D We’ve got Year 2s, and picking up the punctuation thing, we find this is a significant problem with punctuation up near [year] 2 as well. And I think it would be interesting to compare Year 3 and Year 4 because I feel that children become fluent readers at Year 2 and Year 3, and it’s only when they get in the reading mileage that they really can see for themselves and hear where the pauses come. Because when they’re just reading one sentence or two sentences, and it’s maybe word by word reading, they haven’t got that fluency, and so therefore they’re not picking up on when they need to pause. So the Year 2s, we’ve got 48% achieving below expectation in punctuation, and I just think it would be interesting to see where they’re at Year 3, and see whether that rights itself as they become more fluent.

A broad range of documents was collected from 14\(^2\) schools in the study, including reports to their boards of trustees on student achievement, literacy action plans, annual self-review and variance statements for the Ministry of Education, and inquiry planning formats for individual teachers. Typically, schools reported on shifts made by groups of students over the year and student needs were identified from these data. Targets for student achievement were often set in these documents, but no new inquiry was indicated for those students who were still underachieving in reading or writing. Analysis of these documents revealed a divergence between those schools who engaged with teacher effectiveness and self-regulated improvement practices, and those who appeared to be working in a more recursive inquiry mode, reliant on repeating new practices learned in the LPDP. Table 13 summarises the characteristics of inquiry practices found in documents from each school. Collaborative inquiry was identified in documents from 10 of the 14 schools that submitted them, indicating that school-wide data was being discussed with teachers and used to support shared ownership of data and to discuss issues of achievement and responses to students’ needs.

\(^2\) Schools D and P did not submit documents for the study.
Table 13

*Analysis of school documents in 14 schools related to inquiry processes*

<table>
<thead>
<tr>
<th>Inquiry categories</th>
<th>Schools</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No indicators</td>
<td></td>
<td>0.0</td>
</tr>
<tr>
<td>Collaborative inquiry</td>
<td>A, B, C, G, H, I, J, K, L, N</td>
<td>71.4</td>
</tr>
<tr>
<td>Teacher effectiveness</td>
<td>A, B, C, G, H, I, K, L, N</td>
<td>64.3</td>
</tr>
<tr>
<td>Self-regulated improvement practices</td>
<td>A, C, H, I, K, L, N</td>
<td>50.0</td>
</tr>
</tbody>
</table>

*Note.* Schools D and P did not submit school documents to the researcher.

Only seven of the fourteen schools that submitted documents for Study 1 had evidence of a theory-in-use akin to that of school-based inquiry as espoused by the LPDP. School C illustrates this in their “test analysis report” for the board of trustees. They raised the idea of comparing yearly progress levels by class as a possible next step in their discussion document and also indicated their interest in investigating other models operating in schooling improvement clusters. The document certainly shows a number of instances where achievement results were linked to teacher knowledge and where new questions were posed about current practices and their effectiveness:

Last year our asTTle results also showed a lack of progress in the [Year 8] above [average] group. Some teachers felt confident when working with their below [average] groups but were unsure about next steps when working with more able students. This indicated that professional development in teacher knowledge needs to be ongoing . . . . It would be interesting for teachers to compare and correlate this data [a school-based grammar and punctuation test] with the results of the grammar and punctuation section in the asTTle writing…. As a school we need to compile a list of those students classed as non-readers, critical, and at risk . . . . So they can be more closely monitored and supported . . . . We have noted that in the asTTle writing, the Year 8 cohort’s results appear to have slipped, especially in the “above” band. Does this mean that students have forgotten much of what they were taught, or does it mean that there is a need for increased moderation between Year 7 and 8 teachers? Another reason might be that students have not made progress over the year. (Report to Board of Trustees, School C, March 2006)

The report to the board of trustees on students’ literacy achievement and the literacy action plan for School J was typical of those schools whose use of school-wide
Chapter 6 – Results and Discussion: Inquiry and coherence in 16 schools

data followed a more evidence-informed approach. For example, STAR data were reported with percentages of students who were “below”, “at”, or “above” average, but no new inquiry was posed for those thirty-four percent of students who were still below expectations. Next steps for learners were linked to general explanations of student needs, for example, “having difficulties with visual information to check for meaning,” rather than teaching practices that may need to be changed or adapted. Nor did the report indicate any new learning for teachers that might be required as a result of their students’ reading achievement data. The school appeared to have moved on to writing as their professional learning focus.

Inquiry at classroom level

Study 1 also examined classroom practices for evidence of the “core beliefs and values” (Century & Levy, 2002) of the LPDP and how schools might be adapting these in the year following their exit from the programme. This part of the research involved observations of literacy lessons, interviews with teachers and students, and a teacher questionnaire. New teachers to the school were included in the research so that any discussion around the nature of sustainability would encompass establishing if the LPDP messages had “travelled” in the school (Stein et al., 2008) now that facilitators were no longer visiting them and the school had to manage the knowledge sharing for themselves.

The reporting of results at the classroom level about co- and self-regulated inquiry practices begins by examining the responses that teachers made in questionnaires and interviews to questions about their purposes for gathering data and their practices for supporting students that were underachieving in their class. Then, classroom lessons in each of the 16 schools are analysed for how self-regulated inquiry practices for students were being translated into teacher practices. Students’ responses to these lessons are then analysed for their understanding of the learning objectives and use of strategies for checking on their own progress. Finally, the data related to teacher learning is analysed to assess if teachers are self-regulated in their own approaches to learning.
Table 14 categorises responses from 16 literacy leaders and 16 teachers who were new to their school about their purposes for gathering classroom data on literacy. Evidence-based decisions about next steps for their students were well embedded in teachers’ narratives of their literacy practices. Seventy-one percent of literacy leaders and 68 percent of new teachers to the schools talked of identifying student needs from these data. The following was a typical comment:

My purpose? Well, the first thing is that the children set individual goals with me at the start of the term. (New teacher, School B, 2006)

<table>
<thead>
<tr>
<th>Purposes for analysing student achievement data</th>
<th>Literacy leaders</th>
<th>%</th>
<th>New teachers</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting</td>
<td>1</td>
<td>4.76</td>
<td>1</td>
<td>4.54</td>
</tr>
<tr>
<td>Evidence-based decision making</td>
<td>15</td>
<td>71.43</td>
<td>15</td>
<td>68.18</td>
</tr>
<tr>
<td>Involving students in their learning</td>
<td>0</td>
<td>0.00</td>
<td>1</td>
<td>4.54</td>
</tr>
<tr>
<td>Comparison with previous years, norms, or other schools</td>
<td>3</td>
<td>14.28</td>
<td>1</td>
<td>4.54</td>
</tr>
<tr>
<td>Co- and self-regulated improvement practices</td>
<td>2</td>
<td>9.52</td>
<td>4</td>
<td>18.18</td>
</tr>
<tr>
<td>Transfer</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Total descriptors</td>
<td>21</td>
<td>100.00</td>
<td>22</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Again, the notion of co- and self-regulated inquiry was not strongly evidenced, with less than ten percent of responses from literacy leaders and just over 18 percent of responses from new teachers describing their practices in this way. The following responses were representative of those few teachers who had grasped the notion of assessment as an inquiry into the effectiveness of their teaching and their need to keep learning:

[W]hat areas of my own teaching I need to get up skilling in: if it’s complex sentences that the kids are struggling to get then I read it myself so I know how to teach it, or get ideas on how to teach the areas. (Literacy leader, School A, 2006)

I suppose looking into other ways of putting it across to the children, if they haven’t picked up on things, other angles. And I suppose there’s . . . that slice of accountability that we’re needed for as well. (New teacher, School M, 2006)

One other question is relevant here in relation to the core beliefs and values of the LPDP. An assumption expressed in various written descriptions of the project (Bareta & English, 2006; 2007) and in *Effective Literacy Practice in Years 1 to 4* (Ministry of...
Education, 2003), the core text used with teachers in the first cohort of 2004–05 project schools, is that teachers can make a difference in the face of other external factors that impact on students. Table 15 summarises teachers’ responses to a question about what they thought needed to happen to raise the achievement levels of students that they found really hard to progress. Timperley and Phillips (2003) found that there was a shift in ownership of responsibility for underachievement after professional learning in literacy. Sixty percent of teachers in their study recognised that school-based responses were critical to the issue, compared to over 80 percent of responses, prior to the intervention, naming external conditions that needed to change. Comparing their findings to the 54 responses from 32 teachers in this study (see Table 15 for the categories) highlights where the literacy leaders and new teachers stood in relation to this core belief of the LPDP. About 40 percent of literacy leaders nominated effective teaching as one strategy to reduce underachievement in literacy and just under forty percent of the new teachers mentioned their own role in this endeavour. Both groups also nominated increasing the resources and/or funding for schools to allow greater numbers of specialist support teachers, to reduce class size or to provide more professional development. There was somewhat less focus by these teachers on external supports, such as parental responsibility.

Table 15
Teachers’ responses about strategies for underachievement

<table>
<thead>
<tr>
<th>Responses to a question on how schools can raise achievement levels for those falling behind</th>
<th>Literacy leaders</th>
<th>%</th>
<th>New teachers</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>External support, e.g., parents</td>
<td>6</td>
<td>23.07</td>
<td>5</td>
<td>17.86</td>
</tr>
<tr>
<td>External focus, e.g., resourcing – specialist teachers</td>
<td>7</td>
<td>26.92</td>
<td>9</td>
<td>32.14</td>
</tr>
<tr>
<td>External focus, e.g., student motivation</td>
<td>0</td>
<td>0.00</td>
<td>3</td>
<td>10.71</td>
</tr>
<tr>
<td>Internal focus, e.g., effective teaching</td>
<td>11</td>
<td>42.31</td>
<td>9</td>
<td>32.14</td>
</tr>
<tr>
<td>Co- and self-regulated improvement practices</td>
<td>2</td>
<td>7.69</td>
<td>2</td>
<td>7.14</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Total descriptors</td>
<td>26</td>
<td>100.00</td>
<td>28</td>
<td>100.00</td>
</tr>
</tbody>
</table>

An element of the self-efficacy that Timperley and Phillips (2003) describe in their research was also evident in this study, even in those responses that indicated the need for more time or one-on-one support for these students. Teachers owned the issues and saw their role as central to impacting on the problem:
I think I need to have more specific time with them teaching those strategies, basic strategies – more time. The programme that I need to do is work specifically for them, not anything else. The biggest thing to me – I need more time with them. (Literacy leader, School D, 2006)

Just ongoing professional development for teachers. . . so you know what other teachers are doing who have the same problems. (New teacher, School J, 2006)

One trend was clear across the whole data set from the 16 schools. Few teachers clearly articulated the self-regulated approach that was a central focus of the LPDP, such as recognising gaps in their own knowledge or focusing on the student achievement data to monitor their own impact on students’ learning. The comments below did capture the idea of self-regulation, but they were not typical responses:

I think perhaps we need to look and think, “Now what can we do differently.” We do talk about it all the time, but sometimes I feel that we don’t . . . we know what the problem is, but we haven’t got the expertise, or we haven’t chosen the right teaching method to actually fix it for those children. (Literacy leader, School B, 2006)

The first thing I think is probably to develop my own skills and knowledge, because I mean, I had a classic example when I was working with my lower students. I was getting nowhere with them, and I could see that I just wasn’t hitting the right buttons somehow. Then the reading recovery teacher told me that I needed to . . . go over the whole storyline of the book, and she showed me just how to do that and get the children understanding what the story was about and gave me some steps. (New teacher, School G, 2006)

Only a few commentators have shared the classroom observational data that they used in their research to evaluate whether or not a programme had been sustained (Cuban, 1984, 2007; Lefstein, 2008). Given that the focus of the LPDP was primarily on teachers and their literacy practices, observation of classroom practices was central to this study. The two dimensions of inquiry and coherence once again formed the framework for analysis for the lesson transcripts in this study. Each lesson was analysed as a whole to examine whether ideas about self-regulated inquiry were embedded in teacher practices. The interrelationship between teaching, as interpreted from the lesson transcripts, and students’ learning, as represented in the student interviews after the
lesson, enabled some judgments about the impact of core teaching practices of the LPDP in these schools.

Table 16 summarises the findings about use of classroom-based inquiry practices across the 16 schools, divided again between the lessons undertaken by literacy leaders who had been a part of the LPDP and new teachers to the school.

Table 16

<table>
<thead>
<tr>
<th>Inquiry and coherence in literacy lessons</th>
<th>Literacy leaders ((n=16))</th>
<th>Lessons %</th>
<th>New teachers ((n=16))</th>
<th>Lessons %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit teaching</td>
<td>16</td>
<td>100.0</td>
<td>13</td>
<td>81.3</td>
</tr>
<tr>
<td>Student needs</td>
<td>2</td>
<td>12.5</td>
<td>1</td>
<td>6.3</td>
</tr>
<tr>
<td>Self-regulated improvement practices</td>
<td>9</td>
<td>56.3</td>
<td>6</td>
<td>37.5</td>
</tr>
<tr>
<td>Transfer to other learning</td>
<td>4</td>
<td>25.0</td>
<td>3</td>
<td>18.6</td>
</tr>
<tr>
<td>None of above criteria evidenced in transcripts</td>
<td>0</td>
<td>0.0</td>
<td>1</td>
<td>6.3</td>
</tr>
</tbody>
</table>

Most of the transcribed lessons were small-group instructional contexts, although typically lessons began with the whole class and organisational matters for tasks. All of the literacy leaders were explicit about the learning intention for the lesson, once the students were settled into their instructional group. They usually began the lesson with a discussion about the specific learning that students would be doing in reading or writing, but only nine of the 16 leaders built success criteria with their students so they would be able to self-regulate their learning. Only two literacy leaders mentioned to students why they would need to make headway with their learning by sharing or mentioning achievement data related to the learning in the lesson.

Nor were the key LPDP practices consistently observed in the lesson transcripts from the new teachers to the school. Over 80 percent of new teachers emphasised the learning, rather than the activity focus, for the lesson but only 37.5 percent constructed success criteria with students and supported them to use these during the lesson as strategy for self-regulation (Timperley & Parr, 2009a; Wiliam, 2007). Just one of sixteen new teachers linked the learning focus to their students’ needs as identified in a previous assessment or an informal observation.

Assessment for learning practices are understood to make far greater demand on teachers’ expertise than where judgments are more about students’ ability in relation to others (Marshall & Drummond, 2006). By way of illustration, the new teacher in School
A specifically focused on self and peer assessment in her writing lesson. Students had previously been developing the success criteria for effective explanations in their writing lessons. The lesson transcript indicated that the new teacher was aware of the need to teach self-regulatory skills explicitly:

After your buddy has assessed you, you’re going to go back, and you’re going to look at your own explanation, quickly read through it again, and you’re just going to fill out this little self-assessment which says: “I’m now confident at writing explanations.” And you’re just going to circle one, so you’re only going to say: “Little” (you know, you’re a little bit confident), and maybe you’re around the middle so you’re “Some,” or a “5” (which is lots!), OK? - which means that you’re basically an expert and can teach someone else how to write an explanation! And then, you’re going to write, “Next time I write an explanation, I can do the following things…” so there might be two other things other than what your buddy said. And then you’re going to say, “I am proud of the following in my explanation.” (New teacher, School A, 2006)

The new teacher confirmed that this focus on self-regulation was an embedded part of her practices. Self and peer assessment had been emphasised in her initial teacher training, but it had also been discussed and shared in teacher meetings at her new school. Subsequently, her students were all very sure of the learning they were doing and were very precise about their next steps:

I would say more flow with my writing – like in logical and sequential order. I use the thesaurus well though. (Student 1, New teacher lesson, School A, 2006)

Adverbs, technical terms, because I don’t have technical terms in my explanation – not many. (Student 2, New teacher lesson, School A, 2006)

In School J, there was little evidence of explicit teaching of reading in either the literacy leader’s or the new teacher’s lesson transcript, beyond a general mention of the purposes for the lesson. Students were therefore not able to use the success criteria to judge how they were going in terms of their writing or reading, depending largely on their teacher for any feedback on their progress:

And we’re going to bring in a little bit more about vocabulary and the study of words, and in particular, descriptive words, and we’re also going to pick up on the learning intention we did with Matilda, which talked about understanding
what the story was about, and I’ve put it here – focusing on the text and making
responses to show that we understand the story. (Literacy leader, School J, 2006)

This explanation of the learning was too general and appeared to confuse the
students. They could not, therefore, access any specific criteria to self-regulate, replying
to the question about what their next steps for learning might be with the following:

Read a bit more accurately, cos I don't know if I understand every word. (Student
1, Literacy leader lesson, School J, 2006)

Paying more attention – keep up with the class. (Student 2, Literacy leader lesson,
School J, 2006)

Learning bigger words – learning faster. (Student 3, Literacy leader lesson,
School J, 2006)

Table 17 returns to the cross-school view of the study, with considerably less than
half of the 94 students interviewed being able to articulate any specific learning intention
for the lesson or any precise next steps for themselves as learners. The differences
between those students in the literacy leader lessons and those with new teachers to the
school were not that pronounced in any of the three categories coded in their transcripts.
Students seemed to struggle with the idea of judging their own progress, with only 12.5
percent of students in literacy leader lessons and 19.6 percent in new teachers’ lessons
utilising success criteria to support their perspective of where they were at with the
learning. This was despite 56.3 and 37.5 percent of literacy leaders and new teachers,
respectively, developing some sort of success criteria in their lessons (see Table 16).

<table>
<thead>
<tr>
<th>Student self-regulatory practices in 32 lessons</th>
<th>Students with literacy leaders ($n=48$)</th>
<th>%</th>
<th>Students with new teachers ($n=46$)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can articulate specific learning intention for lesson</td>
<td>18</td>
<td>37.5</td>
<td>14</td>
<td>30.4</td>
</tr>
<tr>
<td>Can use success criteria to describe progress</td>
<td>6</td>
<td>12.5</td>
<td>9</td>
<td>19.6</td>
</tr>
<tr>
<td>Can articulate specific next steps in learning</td>
<td>13</td>
<td>27.1</td>
<td>13</td>
<td>28.3</td>
</tr>
</tbody>
</table>

Finally in this section, the literacy leaders and new teachers were asked about
their continued professional learning for literacy in the year following their participation
in the LPDP. Pedagogical content knowledge was one of the five conditions cited by
researchers as being critical to sustainability and a key ingredient in approaches to co-
and self-regulated inquiry practices (Coburn, 2003; Elmore, 2002a; Parr & Timperley,
2006; Timperley et al., 2007). Table 4 (see chapter 4) recorded the nature of the
professional learning that was continued in each school in the year after their exit from
the project. Nine of the 16 schools had opted to continue their professional learning in
literacy, almost all drawing on internal leadership knowledge to support ongoing learning
for teachers. Seven schools moved to new curriculum areas for professional learning and
were faced with competing demands on teachers’ time for new pedagogical content
knowledge learning in literacy. Table 18 summarises the specific areas of new learning in
literacy identified by literacy leaders and new teachers in each of the 16 schools.

There was some evidence that schools were still committed to ongoing
professional learning in literacy. Sixty two percent of literacy and leaders and new
teachers named pedagogical practices and/or content knowledge for literacy as their
specific learning focus for 2006. However, a quarter of the literacy leaders did not name
a specific focus and over a third of the new teachers either had no specific focus or stated
more general areas of learning.

### Table 18

<table>
<thead>
<tr>
<th>Nature of teacher learning focus for 2006</th>
<th>Literacy leaders (n=16)</th>
<th>%</th>
<th>New Teachers (n=16)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No specific focus</td>
<td>4</td>
<td>25.0</td>
<td>1</td>
<td>6.2</td>
</tr>
<tr>
<td>General literacy focus</td>
<td>0</td>
<td>0.0</td>
<td>5</td>
<td>31.3</td>
</tr>
<tr>
<td>Pedagogical practices and/or content knowledge</td>
<td>10</td>
<td>62.5</td>
<td>10</td>
<td>62.5</td>
</tr>
<tr>
<td>Another curriculum focus</td>
<td>2</td>
<td>12.5</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>100.00</td>
<td>16</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Interestingly, in those schools that had moved on from their literacy focus to a
new curriculum area, there was no link between the school taking on a new curriculum
focus for 2006 and a subsequent lack of specific literacy learning goals for teachers. For
example, six out of the seven new teachers and literacy leaders in schools where there
was a new professional learning focus still nominated specific areas of individual
learning in literacy that they were engaging in. If their focus was specific, then they
nominated aspects of practice such as “links between reading and writing,” “reading
strategies for comprehension,” “working with writing groups,” or how to interpret data as
their individual area of interest.
However, the learning foci did not appear to be strongly linked to student achievement data, indicating that an iterative inquiry cycle was no longer in place. If teachers were engaged in co- and self-regulated inquiry, as described on the LPDP materials, then you would expect them to have linked their learning needs to recent student achievement data. When asked for their reasons for identifying areas of literacy professional learning for 2006, a quarter of the literacy leaders and only 12.5 percent of the new teachers used evidence of student achievement in their current class to decide on this focus. Of those literacy leaders who did have a focus, less than half of the literacy leaders or new teachers mentioned that they were working with their colleagues on a related inquiry. Even more surprising was that, where schools had opted to continue their literacy focus for another year beyond the LPDP, five of the seven literacy leaders did not nominate specific learning foci. These findings mirror the views of school leaders recorded in Table 7 (see chapter 4), where only 27.3 percent of the responses described ongoing learning processes for teachers as an indicator for sustainability.

*Coherence of effective classroom practices*

A key aspect of the inquiry into school leaders’ perspectives on sustainability in Study 1 was related to their understandings about crafting coherence of instructional practices across the school and within classroom programmes. The research instruments examined whether principled knowledge and effective practices gained from professional learning projects such as the LPDP were recognised and amplified in other areas of the curriculum. The coherence dimension of the theoretical model for sustainability being examined in this research, like that of co- and self-regulated inquiry practices, applied to both school-wide practices and individual classroom contexts. The fundamental idea is that for sustainability of professional learning interventions to be achieved, schools need to be able to align messages from the external educational environment about curriculum and effective teaching and learning in order to amplify and embed more effective practices (Honig & Hatch, 2004; Newmann et al., 2001). Firstly, the results related to coherence of effective instructional strategies are reported in relation to school-wide practices in the 16 schools. Then, at the classroom level, teachers’ views on the connections between professional learning projects are discussed. Finally in this section, results from 32 lessons are discussed in relation to how teachers made links across curriculum areas.
School-wide coherence

There were a number of ways that schools expressed understanding about coherence of ideas. This section reports findings from the range of documents collected from the schools in the study and compares these findings to the views about coherence expressed by school leaders. Table 9 recorded school leaders’ views of connections between professional learning projects, and this already has been analysed in chapter 4. Typically, school leaders did not make explicit connections with their teachers as one project finished and another began. Chapter 4 concluded that school leaders in this study ring-fenced their thinking about sustainability to literacy and did not harness the potential of crafting links (such as evidence-based instruction) between projects to enhance improvements across the curriculum.

Table 19 lists the characteristics for coherence of instructional practices found across the documents obtained from 14 schools in the study. The “transfer to or linking sites of learning” category was used to code where documents indicated that processes or practices were being linked to new contexts for professional learning, or where patterns and trends in other student achievement data were used alongside the data being discussed. Six of the 14 schools had evidence in their documents of this sort of transfer occurring. For instance, in School K, a large intermediate school, the 2006 action plan for literacy noted that teachers would visit other primary schools to develop their teachers’ content knowledge of progressions that students make as they learn to read in earlier years. This represented an attempt to develop the big picture of literacy learning, to bring some coherence with respect to how teachers of Years 7–8 might see their students’ learning pathways.

A second category for coherence emerged in the process of analysing these documents. This was where documents compared schools’ performance to research findings or linked school practices with theory. For example, the literacy plan for 2006 in School G included a diagram entitled “Evidence-based Inquiry into Practice,” with a question for teachers to ask daily: “How is the change I am making impacting on students and their achievement?” The document used terms that teachers could apply across the curriculum, such as “metacognition” and “prior knowledge.” These were examples of principled knowledge that transcends literacy and builds the language of teaching (Hiebert et al., 2002). School L’s action plan for literacy, revised in May 2006, set out their intention to discuss the place of an “information skills test” in their
assessment regime and signalled an attempt to better align the school’s assessment practices in the wake of the LPDP and their new understandings of testing. It appeared that School C were transferring their learning about effective assessment gained in the LPDP to other areas of their assessment practice.

Table 19
Analysis of school documents in 14 schools

<table>
<thead>
<tr>
<th>Coherence categories</th>
<th>Schools</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No indicators</td>
<td>B, E, F, D</td>
<td>28.6</td>
</tr>
<tr>
<td>Transfer to or linking sites of learning</td>
<td>A, C, J, K, N</td>
<td>35.7</td>
</tr>
<tr>
<td>Theoretical or principled knowledge</td>
<td>A, G, H, I, K, L</td>
<td>42.9</td>
</tr>
</tbody>
</table>

Note. Schools D and P did not submit school documents to the researcher.

When asked about connections they were making across their professional learning, the school leaders had not expressed the transfer of theoretical knowledge in their responses, yet this was noticed as a feature of the school documents. It appeared that these schools were not cognisant of the explicit ways that coherence of effective instructional practices was actually being crafted across different professional learning initiatives.

Coherence of classroom practices

This section examines teachers’ awareness of connections between their professional learning projects. The other focus for analysis was to examine if the notion of coherence was being transferred to classroom teaching by making connections for students between areas of learning.

Given that school leaders had made few links between professional learning projects, it was not unusual to find that similar patterns emerged when teachers were interviewed after their literacy lesson about links they were making across projects. Table 20 summarises the nature of their responses with the same codes as used for analysis of the leaders’ responses. If teachers made any links between professional learning projects, it was mostly focused on the explicit teaching approaches that they now employed. Almost exclusively, they described the specific and “visible” learning intentions and success criteria, the progressions that students make in their learning, and specific feedback. The next most typical connection recorded was using evidence to plan for teaching and very general statements. They made no mention of catering for diversity or
prior knowledge as a connection between projects. Only one literacy leader suggested the idea of teacher effectiveness as a link between the LPDP and other projects:

That it’s the teachers who make the difference and it’s what we do in our classrooms, and we’ve got to look at our practice, and what we’re doing, and the changes we make for learning and not just saying, “Oh, well, you know, it’s the kid’s fault and he’s not going to learn.” It’s really addressing what we need to do. (Literacy leader, School B, 2006)

New teachers to the schools seemed to struggle to make connections. Either they had not experienced any other projects since they were in their first year of teaching, or they focused on general ideas such as “resources” or “engaging the students”.

Table 20
Teachers’ views on connections between professional learning projects

<table>
<thead>
<tr>
<th>Connections between projects</th>
<th>Literacy leaders</th>
<th>%</th>
<th>New teachers</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>General statements</td>
<td>1</td>
<td>4.00</td>
<td>3</td>
<td>13.04</td>
</tr>
<tr>
<td>Need for teacher pedagogical content knowledge</td>
<td>0</td>
<td>0.00</td>
<td>1</td>
<td>4.35</td>
</tr>
<tr>
<td>Focus on teacher effectiveness</td>
<td>1</td>
<td>4.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Evidence-based decision making</td>
<td>5</td>
<td>20.00</td>
<td>4</td>
<td>17.39</td>
</tr>
<tr>
<td>Explicit teaching</td>
<td>13</td>
<td>52.00</td>
<td>10</td>
<td>43.48</td>
</tr>
<tr>
<td>Student awareness of their learning</td>
<td>1</td>
<td>4.00</td>
<td>2</td>
<td>8.70</td>
</tr>
<tr>
<td>Learning processes for teachers</td>
<td>1</td>
<td>4.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Other: specific links</td>
<td>1</td>
<td>4.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>No response</td>
<td>2</td>
<td>8.00</td>
<td>3</td>
<td>13.04</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>100.00</td>
<td>23</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 16 also provided some evidence about how teachers’ understandings about coherence played out in classroom lessons. In twenty-five percent of the lessons taught by the literacy leaders, explicit links were made between the learning in the lesson and how it might be applied to other areas of learning. Most often this was a link between reading strategies being learned and how they might be used in writing. By comparison, only three of the lessons by new teachers to the school had evidence of these links being made explicit for students. Again, making links across learning areas was not a typical practice in these classrooms.
Comparing schools’ theories of action to their patterns of student achievement

This section compares the findings from these qualitative data with that of each school’s student achievement data for literacy. Figure 5 synthesises the qualitative data for each school from 2006–07. As explained in chapter 3, the data for each school were coded and scored according to whether or not the interview or questionnaire responses, documents, lessons and teacher meetings exemplified the advanced end of the two continua on the conceptual framework for this thesis. Each school’s total score became co-ordinates for the two axes labelled coherence and inquiry. Appendix B records the scoring of responses, documents, lessons, and teacher meetings for each school.

These Study 1 data showed some indicative links between those schools that sustained their patterns of progress in student achievement data in 2006–07, according to the benchmarks described in chapter 5, and their understandings of coherence of effective instructional practices and co- and self-regulated inquiry. Schools A, B C, D, G, I, K, N and P all sustained their rates of progress for a new cohort of students in 2006-07, that is, students’ mean gain for reading were greater than 0.33 stanine, or for writing, students made a mean gain of 0.5 standard deviations beyond normed groups in asTTle. School L markedly improved on their rate of progress for a new cohort of students. Students in the School L 2006–07 cohort made a net gain of 0.9 stanine compared to the net gain of 0.48 stanine made by the students in 2004–05.

Certainly, those schools that scored highest around the inquiry dimension (Schools A and B) sustained their student achievement gains in 2006-07 according to the benchmark criteria. School L best supports the theoretical model in this thesis, scoring the highest in terms of coherence and third highest for inquiry practices, at the same time managing to almost double the improvement gains with students in the two years after their exit from the LPDP.

The graph has a number of schools clustered around the centre (Schools I, G, K, C, N, and P) that all sustained their achievement gains according to the benchmark criteria. These schools did not all necessarily score highly on either the inquiry or coherence dimensions alone, but when considered in relation to others who scored lower on both dimensions (Schools E, M, F, and O), two groups begin to emerge. This lower scoring group either did not sustain their student achievement gains in the two years after their participation in the LPDP or did not submit sufficient data for the researcher to be able to compare their gains.
Chapter 6 – Results and Discussion: Inquiry and coherence in 16 schools

Figure 5. Coherence and inquiry scores in 16 schools.

Three schools offered disconfirming evidence in relation to the conceptual model of sustainability outlined in this thesis. School J was the third highest of 16 scores for coherence, but thirteenth for understandings around inquiry for improvement practices, yet they showed reduced gains for students in reading, relative to when they participated in the LPDP. School D also provided a puzzle in relation to the conceptual model, scoring the second lowest on the inquiry dimension of this study and fifth lowest for the coherence dimension, yet still managing to sustain the gains in reading for their students over successive cohorts. School H was the fourth highest of inquiry scores and, in the 2006 interviews, appeared to have heightened understandings of how effective instructional practices, processes, and knowledge might be transferred across the curriculum. However, their students lost ground in reading achievement in the two years after the LPDP.

The qualitative data were gathered from schools in 2006, the year immediately after their exit from the project. However, the analysis above compares these data with student achievement data gathered over two years, 2006–07, so arguably there may have been some significant changes in understandings and practices about inquiry and coherence for all schools in 2007 that also contributed to these results.
School J had improved gains for students in reading comprehension in the first year after their participation in the LPDP, only to lose ground again in 2007 (see Figure 23, Appendix C for their patterns of student achievement using STAR). A small school with only five teachers, they had lost their principal, and only three out of five teachers in 2006 had participated in the LPDP. The literacy leader, however, became the new principal, and an experienced member of staff took up the role of the new literacy leader, presumably increasing continuity in a situation of high staff turnover. They extended their major professional learning focus on reading for another year, using their internal leadership and the newly published text *Effective Literacy Practice in Years 5 to 8* (Ministry of Education, 2006). The 2006 interview data from this school indicated that the principal and literacy leader completely reorganised the regular staff meetings, setting new structures in place for discussion of achievement data and, in the wake of the LPDP, continuing with provision for teachers to be observed and/or to have lessons modelled for them, based on their learning needs. They had continued with teacher journals for reflection, and target groups for monitoring impact of literacy practices. In 2006, at least, there was strong fidelity to the processes and structures introduced to the school by the LPDP. Over this period the student achievement gains were sustained according to the benchmarks set for sustainability.

From evidence gleaned from the Education Review Office22 (ERO) report (2008) for School J, it appears that adherence to the LPDP processes may not have continued into 2007. Teachers had identified writing as an area where students did not achieve as well as in other aspects of literacy. The school had then joined a cluster of schools engaged with an external provider for professional learning related to writing and became involved in the Numeracy Development Project as well. Interestingly, one area for improvement noted in the ERO report of the following year was entitled “sustainability” and included the following comment:

Quality time is needed to consolidate and embed much of the learning and change in school and teacher practices that have resulted from the professional learning and development. School management needs to consider ways of achieving this

---

22 The Education Review Office (ERO) in New Zealand provides external evaluation of the education provided for school students in all state schools, including integrated schools and kura kaupapa Māori. ERO review reports go to school boards of trustees and to the Ministry of Education. They are publicly available.
consolidation when planning future professional learning and development activities. (ERO report, School J, August 2008, p. 7)

A marked drop in scores was evident in the data gathered from School J between academic years. Students’ mean stanine reduced by 1.16 over the holiday period, and by the end of 2007 they had not made up this loss entirely. In this case, it appeared that a combination of this drop in achievement over the summer break and the change to another curriculum focus in 2007 had impacted on their ability to sustain the student achievement gains they had made while in the LPDP. Their inquiry skills and knowledge were not sufficiently embedded to offset the challenges of underachievement and a new professional learning focus.

School D also offered a challenge to the theoretical model. They sustained the gains in student achievement that they had made according to the benchmark criteria. School D is an urban intermediate but is much smaller than other intermediate schools. The aggregated scores from interviews, lessons, documents, and questionnaires indicated that their understandings about coherence and co- and self-regulated inquiry practices were not as developed as those in most other schools. Their new cohort of students in 2006 was, on average, 0.38 stanine below that of the 2004–05 cohort, but student scores indicated that similar gains had been achieved by the end of 2007 despite the levelling-off effect in the second year (see Figure 18, Appendix C).

School D did take up a new ICT professional learning focus in 2006 and had a turnover of forty percent of their teachers in the same year – both of these factors combined could have derailed efforts to sustain their gains. However, the literacy leader and principal remained in the school, and the ERO report (2007) noted that new appointments had added strength to the quality of teaching and learning since their last visit in 2005. The literacy leader was not allocated a class so that she could work with the new influx of teachers, modelling and giving feedback to them on their practices as well as supporting them to select a target group of students, interpret their test data, and match teaching responses to these data. The decision to have an on-site coach may well have been enough to maintain practices learned during the LPDP and sustain the gains they had made in 2004–05 with a new cohort of students.

Without returning to Schools J and D in subsequent years, it is difficult to explain precisely which conditions may have led to sustainability or reduction of student
achievement gains. The outcomes in these two schools were therefore considered in the framing of the second study and in the selection of case study schools to further investigate the nature of sustainability.

School H was the only one of these three schools that was visited again in 2007 to ascertain the nature of school-based conditions. School H is a Decile 2, urban intermediate school of approximately 230 students. Their focus for the LPDP was reading comprehension, and their patterns of student achievement data are illustrated in Figure 21 in Appendix C. In 2006, School H dramatically improved on the achievement shifts that they had made with students relative to while they were a part of the LPDP. As with School L, the mean stanine gain for students at School H was markedly higher compared to that made by the earlier cohort. In 2006, the mean gain in reading comprehension for Year 7 students was 0.6 stanine, whereas for the two years of the LPDP the mean overall gain was only 0.36 stanine. By the end of 2007, this gain had fallen away considerably to a net gain of just 0.12 stanine for the two years.

School H had not moved on to a new professional learning focus in 2006, opting instead to try to embed the LPDP across the school and, like School D, to strategically position their literacy leader as a full-time mentor for teachers. They had experienced a low turnover of staff in 2006: all nine staff had participated in the LPDP. By 2007, however, School H had three new staff, and in that year their literacy leader went on study leave for 18 months.

School H had continued the patterns of inquiry they had begun during the LPDP into 2006, discussing student achievement data in syndicates and using these discussions to drive decisions about the focus for ongoing teacher learning. In 2007, the literacy leader explained that they had focused on the lowest 25 percent of students, investigating ethnicity data and “pulling apart” the subsets of the STAR test to inquire into particular weaknesses. They had responded to these data in 2007 by focusing on paragraph comprehension and would track the impact later in the year. There was no reference to an earlier inquiry they had mentioned in 2006 about Pasifika students, nor did the literacy action plan in 2007 include focused goals for student achievement. Neither the strategic plan nor the literacy long-term plan mentioned previous data used for setting goals for teacher practices.
By 2007, the principal and literacy leader considered the school to be in “desperate need” of further external literacy support. The literacy leader who had been a full-time mentor and coach in 2006 was granted study leave in 2007, and the new literacy leader/deputy principal did not want to venture into a writing focus without external leadership. School H had also engaged with a local cluster of schools that were focusing on literacy, and leaders used this as an opportunity to reapply for the LPDP. In the year that this cluster took to gather and analyse their cluster-wide evidence, School H’s principal was correct in his analysis that his students had lost ground.

In School H, then, movement in leadership, teacher turnover, and a fragile understanding of inquiry practices appeared to have combined to reduce the longer-term sustainability of the gains and practices established in the LPDP.

Significance of findings for sustainability

The data from Study 1 offered interesting insights into the complexity of sustainability beyond the usual teacher satisfaction ratings and self-report descriptions of practices and their effectiveness for student outcomes. Each of the 16 schools in Study 1 of this research forged its own pathway towards sustaining the learning gained from the LPDP and the gains in student achievement. Most were successful in this pursuit. This final section of Study 1 examines the significance of these findings for understanding the nature of schools’ perspectives and actions in relation to sustainability and then one specific school-based condition that was found that might explain different outcomes between schools.

Across all schools in the research sample, the understandings of sustainability were self-limiting, with only glimpses of any substantive understanding about inquiry as a means for ongoing improvement. There was little sense of an explicit theory for improvement and sustainability emerging from these schools. Evidence-based teaching and effective literacy practices learned as part of the LPDP were seen as the drivers for supporting improved practices and sustained outcomes for students. There were many examples of the idea that teachers and schools must base their decisions for teaching practices on evidence of students’ needs derived from assessment data. However, this fell well short of the LPDP’s core belief that inquiry into student achievement must also be an iterative and self-regulated process, used to reveal further gaps in teacher knowledge in order to support those students who were still not making progress. Chapter 4
concluded that schools had focused more in maintenance of the LPDP practices, rather than continuing to be adaptive to students’ needs, particularly where current practices were not effective. There were many fewer instances in the data of school leaders and teachers using co- and self-regulated practices. Evidence-based decision-making appeared to be the most pervasive idea for school improvement and sustainability. The division between evidence-based practices and inquiry for improvement, as described in the emerging research base about sustainability and accounted for in the conceptual framework for this study, was evident in Study 1.

Schools in the study generally limited their notion of sustainability to a single project, not recognising the principled knowledge or processes they could transfer to the next learning programme. Typically, leaders and teachers named practices related to formative assessment, such as using learning intentions and success criteria, as the ideas that linked their professional learning. These data did reveal the need for different professional development project leaders to converse at system level at least, so that schools might better be supported to build and craft the coherence of the key messages in each professional learning opportunity.

However, there was some indicative evidence that the two dimensions of inquiry and coherence might be important in determining ongoing improvement in the years following an intensive professional learning project. Two years after they had exited the LPDP, ten of the thirteen schools that presented student achievement data sustained their student achievement gains with a successive cohort of students, according to the benchmarks set for sustainability in this study. Those schools that had sustained achievement gains generally scored around the middle or higher on the inquiry axis related to the qualitative evidence gathered in the year after their exit from the LPDP. This aligns with the findings in Teacher Professional Learning and Development: Best Evidence Synthesis Iteration (Timperley et al., 2007) and the Lai et al. (2009) study. These researchers argue that sustained improvement depends on teachers developing self-regulatory inquiry skills so they can collect relevant evidence, use it to inquire into the effectiveness of their teaching, and make adjustments to their practice. As well, the school that had scored highly on the coherence axis had markedly improved the gains made by their students in 2006–07 when compared to the gains made while participating in the LPDP.
It would appear though, in this study, that the outcomes for student achievement over two years was not necessarily predictable from the espoused theories of action and practices about inquiry and coherence observed after only one year. This may, for instance, explain why School D was able to sustain student achievement gains after two years, but in 2006, at least, did not show strong evidence of inquiry knowledge and skills.

There was one common school-based factor that explains more about the differences in student achievement patterns between schools, particularly in the year following their exit from the LPDP. School L had scored highly on both the inquiry and coherence axes, at the same time markedly improving gains in reading with a new cohort of students in 2006. Students in Schools H and D also both made greater achievement gains in 2006 than they had over 2004-05. All three schools were intermediate schools that decided to focus on sustainability in 2006 by having their literacy leaders released from classroom duties in order to support the effective teaching of literacy across the school. In 2007, the literacy leader in School H went on study leave and the deputy-principal took up the role. In 2007, student achievement in reading fell away dramatically in relation to the gains of the previous year. The decision to resource a literacy leader that was focused fully on supporting teachers and students may have been a key factor in the outcomes for these schools.

Study 1 revealed how leaders and teachers perceived sustainability and afforded the development of a theoretical framework for sustainability that has promising indications. Study 2 examined the dimensions of the model more extensively in four of the 16 case schools. These schools were revisited in 2007 and 2008 in order to check and deepen understandings about sustainability of educational reform and to warrant more finely tuned practices that lead to sustained outcomes for new cohorts of students.
CHAPTER 7
STUDY 2

INQUIRY AND COHERENCE AS CONDITIONS FOR SUSTAINABILITY

Introduction: Method Study 2

Valerie Hannon (2008), in her research on innovation in schools, quotes businessman William Pollard in a recent article on educational change:

The arrogance of success is to think that what you did yesterday will be sufficient for tomorrow. (p. 77)

This comment captures the rationale for the second study in this research in two ways. Firstly, 10 of the 13 schools that had offered data in Study 1 had sustained the gains made in students’ literacy achievement, together with key shifts in teaching and leadership practices, in the two years beyond their participation in the LPDP. These gains occurred even when there was high teacher and leadership attrition and/or new priorities for professional learning. Notwithstanding this initial success, would these shifts be sufficient to meet the benchmark criteria for sustainability previously described in chapter 3 over time? Secondly, an inquiry approach requires ongoing knowledge-building and adaptation to new circumstances (Earl & Katz, 2006; Timperley et al., 2007). Study 2 set out to examine in greater depth in four of the study schools whether teachers persisted in adapting their teaching practices to meet different challenges in literacy underachievement and whether schools continued to inquire into student achievement data, particularly for those students at risk of not meeting expected levels over time.

The findings in Study 1 indicated that the 16 schools had mostly focused on maintenance of those practices they had learned in the LPDP, with more emphasis on evidence-based practices than on an inquiry for improvement approach. That is, schools did not fully grasp the significant difference between these two approaches. With an inquiry stance, while there may be better results for all students after just one cycle of improvement, schools would re-engage with the cycle to pose new questions about groups of students that were still not succeeding and then decide on further adaptations to practices to support them to improve. Wood (2007) likens this process to sailing and describes a process of “dynamic stability” being required to keep schools on track to meet goals for improvement, warning that “heeling to the proper direction may require
innovation and adjustment, particularly given changes in context or new information” (p. 3). It is possible, as Century and Levy (2002) contend, that programmes can be “stalled” at maintenance and fail to develop an ability to evolve and adapt further. They describe the hallmarks of a third phase, where sustainability goes beyond establishment and maintenance to evolution, where there is further growth and improvement. Their remarks relate to school-based instructional practices and not necessarily to levels of student achievement, so Study 2 investigated whether the same patterns of student achievement occurred with regard to the latter.

It was important to continue to track these schools to determine whether, in fact, the pathways they had taken to sustainability would continue to be effective and whether particular school-based conditions might allow them to move beyond maintenance to continuous improvement. Study 2 examined four of the case schools in more depth over 2007–08 to test the conclusions formed from the initial analyses and investigate the challenges posed for these schools by high-leverage practices of inquiry and coherence which the theoretical framework of this thesis suggests are necessary for sustainability.

This chapter outlines the method for Study 2. Chapter 8 describes the findings in the four case study schools in the three years following their exit from the LPDP. The final chapter considers the findings from both studies and offers theoretical contributions and implications for the field and areas for further research.

Methodology

This section provides justification of the methodology used, explains the rationale for selecting the four in-depth case study schools, describes the participants and the research instruments and the procedures for data collection. Finally, the processes used to analyse these data are explained.

The selection of and return to four school sites from Study 1 allowed for the “double vision” that Schön (1983) promotes as necessary for avoiding being trapped in particular definitions, categories, or hypotheses:

If the inquirer maintains his double vision, even when deepening his commitment to the chosen frame, he [sic] increases his chances of arriving at a deeper and broader coherence of artefact and idea. (p. 164)
Schön argues that the more that is learned about the problem, the more the way the data are recorded and analysed will change and what the researcher notices will be different. This iterative process connected Study 1 to Study 2 of the research plan. Three schools from Study 1 provided challenges to the theoretical model in that they did not necessarily exhibit a close relationship between sustained student achievement gains and the schools’ understandings and practices in relation to inquiry and coherence. School D had sustained the student achievement gains according to the criteria discussed in this thesis, yet the qualitative evidence did not demonstrate strong understandings of either co- and self-regulated inquiry or coherence. Conversely, Schools H and J had scored highly on coherence of effective instruction in 2006, but did not sustain the gains in student achievement by 2007. Inquiry as a systematic set of routines was much less understood by participants or illustrated in data gathered from all 16 schools in 2006. Few school sites in Study 1 made deliberate connections across professional learning projects. These findings raised new research questions for Study 2 such as:

- How robust were the student achievement gains made by these schools in subsequent years?
- Would schools move on from maintenance to an improvement paradigm?
- How important was coherence as a school-based condition for sustainability?

Case methodology was therefore used again in this endeavour with a similar mixed-method approach as in Study 1. By reducing the number of schools for Study 2, it was possible to use a “multi-focal research” perspective (Borko, 2004), where a wider lens could be placed on school-wide understandings and theories in use about sustainability; at the same time, a “near vision” lens could take more notice of teachers and students working in more than one curriculum area.

A second examination of schools would further strengthen findings and allow for refinement of research instruments to search out any disconfirming evidence. Where there is sequential development in the design, data collected in the first study can reorient the next wave of data gathering and analysis (Borman, Clarke, Cotner, & Lee, 2006). In Study 2, variations were established between cases and placed alongside the theoretical framework to support more generalisability of the findings. Strategies such as triangulation, negative case sampling, and member checking can ward off confirmation

---

bias. In Study 2, the use of student achievement data and multi-methods such as interviews, observations, and document analysis supported triangulation of data across each school site.

**Selection of ongoing case study schools**

The criteria used to select four schools from the original sample were that all were intermediate schools. Intermediate schools in New Zealand have students in years 7–8 only, and these students come from a range of other “contributing” primary schools. There are 123 intermediate schools in New Zealand, representing approximately 11 percent of schools catering for this age group. Most students of intermediate schools spend two years in the school before entering a high school for their secondary education. Typically they have larger numbers of students than other primary schools, so their selection for the second phase of the study had implications for how data were analysed. For example, their literacy achievement data related to similar age levels, which enabled comparability in relation to shifts in student achievement. The larger student body also provided sufficient data to examine achievement patterns for the lowest 20 percent of students.

The selected schools also represented both the reading and the writing focus of the original project. In particular, these schools when presented with different combinations of the two dimensions of the conceptual framework in the Study 1 analyses, had distinctively different school-based contexts and had different outcomes in terms of shifts in students’ literacy achievement in the 2006–07 period.

**Participants**

Table 21 summarises the numbers and types of participants in Study 2 for 2007–08. The key difference between the two studies was the inclusion of numeracy leaders in the interviews about school-wide practices and the observation of two lessons per teacher, again extending the qualitative data to include numeracy classroom practices. Teachers and students were asked about both lessons that they had participated in so that a closer examination of the theories-in-use about connections between curriculum areas could be made. Table 22 indicates the characteristics of the four schools in the period 2006–08.
Table 21

Numbers and types of participants in Study 2 schools

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 22

Study 2 school context variables 2006-08

<table>
<thead>
<tr>
<th>Variables</th>
<th>School A</th>
<th>School C</th>
<th>School H</th>
<th>School L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antecedent variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of teachers</td>
<td>20</td>
<td>25</td>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td>Decile</td>
<td>9</td>
<td>7</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Students</td>
<td>600+ Years 7–8</td>
<td>600+ Years 7–8</td>
<td>200+ Years 7–8</td>
<td>400+ Years 7–8</td>
</tr>
<tr>
<td>Average entry level for students</td>
<td>At asTTle norms</td>
<td>Below asTTle norms</td>
<td>Stanine 4</td>
<td>Stanine 4</td>
</tr>
<tr>
<td>Mediating variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher turnover by 2008</td>
<td>50%</td>
<td>23%</td>
<td>33%</td>
<td>61%</td>
</tr>
<tr>
<td>Leadership turnover by 2008</td>
<td>Literacy leader on leave 2007</td>
<td>Literacy leader left in 2008</td>
<td>Literacy leader on leave 2007</td>
<td>Stable</td>
</tr>
<tr>
<td>Teacher learning foci</td>
<td>Numeracy Inquiry learning TDS</td>
<td>Numeracy Inquiry learning TDS</td>
<td>Literacy</td>
<td>Numeracy Inquiry learning ICT</td>
</tr>
<tr>
<td>Outcome variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student achievement gains compared to LPDP</td>
<td>Sustained (maintained similar levels)</td>
<td>Sustained (maintained similar or improved levels)</td>
<td>Gains reduced until participation in LPDP in 2008</td>
<td>Sustained (improved levels)</td>
</tr>
</tbody>
</table>

Note. Number of classroom teachers was sourced from the principal’s questionnaire 2006. TDS is the Teacher Designed Schools Network project. The 2007–08 student cohort in School C had improved gains compared to 2004–05 by 24 asTTle writing points.

School L had relatively stable staffing when they moved into 2006, with all of the leadership team still in place and a relatively low turnover of teachers. By 2008, only seven of 18 staff had actually experienced the LPDP. However, there was a great deal more consistency in leadership, with the literacy leader, deputy principal, and principal all remaining in the school over the same period. While the school had lost some other
leaders who led particular teams of teachers, the literacy leader remained as an in-school coach, supporting all teachers in their literacy professional learning and also providing some specialist support where needed for individual students. Students were taught in one of six learning “academies.”24 The data collected in 2006 indicated that School L had scored the highest in terms of espoused theories of action about coherence of instructional practices when compared with other schools in the sample. At the same time, School L had the third highest aggregated score for inquiry practices. This school had almost doubled its student achievement gains in reading comprehension over 2006–07 when compared to the gains made while in the project.

In contrast, School H was selected because it had sustained students reading comprehension achievement results in the first year after exiting the LPDP, but their students had lost considerable ground in the second year. It presented as a negative case in terms of the theoretical model, scoring relatively highly on both inquiry and coherence in the 2006 data. The school had opted to try to embed the LPDP across the school in 2006 and strategically positioned their literacy leader as a full-time mentor for teachers. In 2007, the literacy leader went on study leave and the deputy principal stepped in to provide leadership in literacy. She did not seek further specialised support for this leadership role over 2007 and had not been formally mentored in the observation and coaching methods as part of the LPDP. In 2008, the literacy leader returned but as a classroom teacher with additional literacy leadership responsibilities. The deputy principal left the school, and the new deputy principal was also soon to leave to take up a new principalship. Although there was no staff turnover in 2007 and 2008, there was considerable turnover in leadership. School H was involved in a cluster-wide process in 2007 to build a case for further schooling improvement funding and, unusually, the school was successful in its bid to rejoin the LPDP that in 2008 was working with its third cohort of schools nationally. This came about because the regional cluster that they were now a part of had applied for the project as a cluster.

School A had scored highly on the co- and self-regulated inquiry dimension, had a twenty percent turnover in staff in 2006, and also had a new principal, who had previously been the deputy principal during the implementation of the LPDP. The school had deliberately chosen not to work in a cluster of schools or move on to a new professional learning focus in 2006. Over 2007, the literacy leader was granted study

24 The learning academies were set up in 2007 and are explained further in Chapter 8.
leave and on her return also took up new management responsibilities. School A had managed to sustain progress in writing achievement with a new cohort of students over 2006–07. In 2007, only 12 out of 20 staff had participated in LPDP, and by 2008 there had been a 50 percent attrition rate.

The final school was School C. It was selected from those in the middle group of schools as represented in Figure 5 (see chapter 6). School C had maintained similar levels of student achievement gains in writing over 2004–05 and 2006–07 and was the only school in Study 2 to have moved immediately into a new professional learning focus in the year following their exit from the LPDP. They had relatively stable leadership and staffing up until the literacy leader left in 2008 to take up a principal’s role at a neighbouring school.

The selection of these schools was, therefore, consistent with Schön’s “double vision” notion, providing a further opportunity to examine those school sites where there may be disconfirming evidence in relation to the theoretical framework proposed in this research.

Data gathering instruments and procedures

As with Study 1, mixed methods design were employed in the data gathering process. The design of the research instruments for Study 2 was again based on the dimensions of inquiry and coherence but was also extended to observe more of the theories-in-use by teachers.

The set of research instruments for collecting data in this new phase of the research included:

- a semi-structured interview schedule for the principal;
- a semi-structured interview schedule for the principal and literacy leader(s);
- a semi-structured interview schedule for the literacy and numeracy leaders;
- the observation schedule used in Study 1 to analyse a literacy and numeracy lesson of a teacher who had participated in the LPDP and a teacher who was new to the school;
- a semi-structured post-interview schedule for both teachers involved in the classroom observations;
- a semi-structured post-observation interview for three students from each classroom (the teachers were once again asked to select three students who represented the upper, middle, and lower bands of literacy achievement in the class but did not represent the extremes).
The research instruments aimed to check practitioner understandings of sustainability now that there had been a longer interval between the end of the LPDP and the establishment of new professional learning priorities, and where there might be new leadership and fewer numbers of teachers who had experienced the project.

Study 2 research tools focused more closely on the school’s use of student achievement data at classroom and school-wide levels, to ascertain whether repeated rounds of inquiry were being undertaken with more specificity of focus and differentiation of targeted responses (Copland, 2003). Had new problems emerged out of the latest sets of literacy data, and had these problems been theorised so that different responses might be adopted than those devised while in the LPDP? These data would provide another check against findings from concurrent research on sustainability in similar contexts (Lai et al., 2009).

As a follow up to findings about the co-and self-regulated practices in Study 1, the students were also asked if they could share the learning objective of the lessons and how they knew whether they had learned this in their lessons. These questions enabled the researcher to compare results from Study 1 with enacted practices. A range of documents, similar to those collected in Study 1, were gathered in Study 2 and analysed for evidence of the steps in the inquiry cycle.

In Study 2, the principals were interviewed separately from other school leaders so that their role in sustainability might be better evaluated. In the earlier study, principals had completed a questionnaire and then participated collaboratively in the school leaders’ interviews in 2006 and 2007. In 2008, the semi-structured interview with the principal pursued how he or she viewed problems in literacy that the school faced and whether he or she had developed a theory for improvement to address these problems. The principals were also asked what their role was in communicating to the board of trustees and parents in relation to student achievement: what professional learning their school was undertaking; and what, if any, interventions he or she had made to adjust how professional learning was delivered if the school had taken on a new focus with new providers. These questions were related to how the principal was thinking about student achievement and whether there was with a longitudinal view about ongoing improvement in the school.
In each year of this research study, the leaders and teachers in the four schools were interviewed about the tools and routines they had used to support sustainability of literacy outcomes. These routines may or may not have been established as part of the LPDP, but their purpose and impact were checked again so that any alignment or adaptation to the project’s core principles and practices could be noted. Tools and routines are viewed by many as important mechanisms for scaling up changes in schools (Coburn, 2003; Coburn et al., 2008; Cohen & Ball, 1999; Robinson et al., in press).

A second line of inquiry in Study 2 related to the second dimension of the conceptual framework, that of coherence of instructional practices. Findings from Study 1 indicated that few schools were making coherence an explicit strategic action to support sustainability. However, when participants in Study 1 were asked to draw out these connections, there were some thoughtful responses about the use of similar strategies across the curriculum and the transfer of processes thought to be helpful for professional learning in the school. Study 2 deepened this inquiry by inviting numeracy leaders to be part of the research. The leaders’ interview, for instance, asked the literacy leader and the numeracy leader what, if any, connections they were making about effective pedagogical practices between their domains. Classroom transcripts were analysed to ascertain whether teachers recognised and articulated the links between their teaching strategies in different curricula and/or made learning links explicit to promote metacognitive learning so that their students could see where each lesson was positioned within the bigger picture of their learning.

In the interviews that occurred after numeracy and literacy lesson observations, teachers were asked what, if any, strategies they had used in both lessons and why. Students were interviewed immediately after the two lessons and asked if their teacher did anything in a similar way in both lessons. The schools documents were examined for any evidence of where coherence of instructional practices across the curriculum was being made explicit. The logic for this was that any one professional learning initiative could contribute to the overall coherence of effective instruction in a school and, in doing so, amplify the ideas about effective teaching and learning, in turn supporting spread and depth of new changes to practices.

Procedures

In 2007, the researcher visited each of the four case study schools for one day in the period October 13 to November 30 to interview the principal and literacy leaders, to
gather schools’ documents related to literacy, and to collect 2007 student achievement data for the same literacy focus that they had undertaken when participating in the LPDP. In 2008, the visits were extended to two days in each site so that lesson observations for literacy and numeracy could be scheduled as well as interviews with students and teachers, the principal, and curriculum leaders. These school visits took place in the period August 20 to November 25. Student achievement data for 2008 were gathered after each school visit so that a third cohort of students could be included in the data analysis. Although the time frame for the 2008 cohort was shorter than for previous cohorts, these data allowed for a longitudinal lens on the school with regard to student achievement patterns over time.

Each research tool was adapted to the individual school so that differences in school contexts could be more specifically probed at the time. The semi-structured interview format allowed this flexibility. The researcher conducted the interviews and observations with all participants in both 2007 and 2008. A research assistant taped the interviews and lessons and arranged for the collection of students’ literacy achievement data subsequent to site visits. Instructions for the introduction of the data collection instruments accompanied each tool to ensure consistency across school sites.

The school leaders were sent the findings from Study 1 so that they could engage with the ideas, and in the 2007 and 2008 site visits, they were also given graphs of their students’ literacy achievement, compared with those of 2004–05, to comment on. In this way, member checking occurred as part of the process.

Analysis of data

The data sets in Study 2 were firstly analysed by case, rather than across the sample of four schools. This approach was to enable the researcher to probe different aspects of each school’s pathway to sustainability, to uncover as much as possible about their individual context, and to notice any disconfirming evidence. The nature of the semi-structured interviews still provided enough consistency across sites to later compare and contrast school contexts and their outcomes in relation to sustainability.

Qualitative data

The method used to analyse the interviews followed similar routines to those described in chapter 3. All of the interviews and lessons were transcribed. The codes developed in Study 1 for similar open-ended questions were used again in the subsequent
analysis of responses. Similar analysis frameworks for the school documents and lesson transcripts were applied again in Study 2. The documents and lessons were used to cross-check, where possible, information offered in the interviews to test whether the theories of action were, in fact, theories-in-use. The coding schedules in Study 2 captured disconfirming evidence directly for each question, and this evidence has been considered in the analyses of data in the following chapter.

Reliability checks were not applied to the coding in these data since the data were not aggregated in the same way as they were in Study 1. Instead, a cross-case analyses framework was used to analyse the school-based conditions over the three years of the study. By comparing sites, the researcher can establish the range of generality of a finding or explanation and, at the same time, determine the conditions under which that finding occurred (Miles & Huberman, 1984). The theoretical framework derived from the research literature, and explored for its significance in Study 1, was revisited in Study 2 in order to look for fresh perspectives and contradictions. As well, the school site dynamics were compared across schools by examining both antecedent and mediating variables over the time span of the study. The antecedent variables included the demographics of the school and the key actors and their advocacy roles in supporting sustainability. Mediating variables that occurred during the post-implementation phase such as teacher and leadership turnover, reprioritising of new professional learning, and moves to codify and institutionalise practices, were used in conjunction with the conceptual framework to compare and contrast the four schools in Study 2, particularly in relation to their outcome variables – shifts in students’ literacy achievement.

Quantitative data

The analysis of the schools’ 2006–2008 student achievement data was aimed at describing whether each of the schools was able to sustain the shifts in students’ literacy achievement or whether the patterns of improvement had subsequently increased or reduced over time. In the first instance, the same criteria used in Study 1 to judge whether achievement gains had been sustained with new cohorts of students were used again in Study 2. Their rationale was explained in chapter 3. Study 2 achievement data included three cohorts of students over two years (2004–05, 2006–07 and 2007–08). A fourth cohort of students included students who began their Year 7 in 2008. These data therefore indicated one year of progress in either reading comprehension (using STAR) or asTTle writing rather than the two years used for the previous cohorts. The fourth
cohorts allowed some ongoing patterns to be gauged, even if the trend in the second year was for the growth in achievement gains to level off. The same student assessment tools had been used over this period as in Study 1 and their validity, reliability, and limitations have been already discussed in chapter 3.

As in Study 1, a second lens was used to judge sustainability in relation to the achievement data. To determine whether schools did have a pattern of ongoing improvement, the gains for each cohort were also compared with the gains established for students while they were participating in the LPDP. For example, if a school had sustained their students’ achievement results, that is, they had a mean gain in reading of over 0.33 stanine or over 0.5 of a standards deviation in writing, but had not improved on their 2004-05 gains then the school was judged as having sustained and maintained. If a school had increased writing or reading scores for a new cohort of students then this was judged as being sustained and improved (in writing this was judged as any increase in mean gain score above the standard error of measurement, 15 points, according to asTTle writing [The University of Auckland, 2004] and for reading, any increases in mean stanine). This second lens enabled more finely tuned analysis of school-based conditions for sustainability in relation to these data.
CHAPTER 8
RESULTS AND DISCUSSION: FOUR CASES

Introduction

This chapter reports the results from four case studies over 2006–08 and is organised school by school in relation to the two key themes of inquiry and coherence. Three schools did sustain their student achievement gains in the three years (according to the benchmark criteria in Study1) following their participation in the LPDP, and one did not. The chapter begins with School L and School H, who participated in reading-focused initiatives during the LPDP but had different outcomes in sustainability. The findings for School A and School C are then discussed in relation to sustainability of their focus on writing. The final section of the chapter discusses the implications of these findings for sustainability of professional learning initiatives.

School L

The results for School L are separated into three subsections. Firstly, the patterns of achievement in reading comprehension for three successive cohorts of students are compared to rates of progress made by students while the school participated in the LPDP. The dimension of co- and self-regulated inquiry practices for improvement is then examined across school and classroom practices, followed by analysis of coherence of effective instructional practices.

School L’s results had conformed closely to the theoretical model when the first set of qualitative data were gathered and analysed in 2006. They had scored the highest on the coherence dimension and third highest on the inquiry dimension in relation to the other 15 schools in the sample and had continued to improve their students’ literacy achievement results in the period 2006–07. School L appeared to be acting in an improvement rather than maintenance paradigm. Study 2 aimed particularly to reassess this conclusion, looking for any disconfirming evidence that might adjust or even negate the conclusions from the analyses.

Patterns in student achievement

School L had sustained their student achievement results over 2006-08 according to the benchmarks for sustainability discussed in chapter 5. Figure 6 illustrates successive cohorts of students in School L over 2004–08, each line tracking mean stanine results in STAR reading comprehension for a cohort of Year 7 students from the beginning of their
first year in the school until the end of their Year 8 year. The 2004–05 cohort of students
was tested only at the beginning and end of the school’s participation in the LPDP, but
the cohorts following were tested at the beginning and end of each of their two years in
the school. The 2008 cohort, however, represents just Year 7 students over one year and
has been included so that three successive cohort patterns of achievement could be
compared to that of 2004–05.

In order to compare gains in student achievement, Table 23 records the mean
stanine gains in STAR in 2004–08. In 2006–07 the mean stanine gain for students almost
doubled compared to the previous cohort (0.91 stanine gain \([n=136]\) compared to 0.48
stanine in 2004–05 \([n=206]\)). After that, successive patterns of student achievement for
new cohorts of students beginning in 2007 and then 2008 remained relatively static, with
mean stanine gains of 0.82 and 0.52 (the latter after one year) respectively. Students in
Year 7 had made improvement well above previous rates of gain, approximately 0.8
stanine in each case, lost ground over the summer holiday period, and then made more
moderate gains in their Year 8 year. Additionally, the entry point levels of Year 7
students in the 2007 and 2008 cohort were on average 0.35 stanine lower than for the two
previous cohorts, so the challenges for teachers were even greater in these circumstances.
School L had sustained and improved their achievement gains over the period 2004–08.
Table 23
School L STAR mean stanine gains 2004–08

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=206</td>
<td>n=136</td>
<td>n=126</td>
<td>n=189</td>
</tr>
<tr>
<td>T1</td>
<td>4.41</td>
<td>4.32</td>
<td>3.97</td>
<td>3.97</td>
</tr>
<tr>
<td>T2</td>
<td>5.5</td>
<td>0.83</td>
<td>4.75</td>
<td>0.78</td>
</tr>
<tr>
<td>T3</td>
<td>5.01</td>
<td></td>
<td>4.61</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>4.89</td>
<td>0.48</td>
<td>5.23</td>
<td>0.91</td>
</tr>
</tbody>
</table>

At risk students in reading comprehension were defined in the LPDP milestone report to the Ministry of Education as those students who were in the lowest three stanines (Bareta & English, 2006). School L had 30 students in Stanine 1 and 2 in 2004 and they made approximately double the progress over two years when compared to mean gains for all students in the school (overall a mean shift of 0.94 stanine compared to 0.48). A further 47 students in Stanine 3 lifted their performance in the STAR test over two years by 0.6 of a stanine. However, subsequent cohorts of at risk students did not continue this pattern in relation to all students in the school (see Figure 7). From 2006, their mean stanine gains were very similar to those made by all students in the school. The number of students in Stanines 1–3 was 77 in 2004–05 and only 39 in 2006–07, so care needs to be taken with interpreting these results. The decline in achievement over the summer break was more noticeable with the at risk group than for the whole student population and may also have impacted on their results over two years.

Figure 7. School L student achievement gains for at-risk students 2004–08.
In summary, School L had improved on student achievement gains for the cohort in the year subsequent to the school’s participation in the LPDP and sustained similar gains for two more cohorts. At risk students within each cohort had similar rates of progress but did not repeat the pattern, identified in 2004–05, of accelerated gain compared to the rest of the cohort.

Inquiry for improvement practices

This section of the School L case study is organised chronologically for site visits over 2006–07 and then 2008 and reports findings related to co- and self-regulated inquiry.

Findings from Study 1 suggested that schools typically limited their inquiry process to evidence-based practices, where teachers used student achievement data primarily to inform their teaching decisions. They may also have been “stalled” in the maintenance phase that Century and Levy (2002) describe, dependent on newly formed practices to support new cohorts of students and, more importantly, to support those students who were still not making adequate progress. In chapter 6, evidence-based practices were described as a subset of a fully fledged inquiry approach, and this boundary became a key idea for further analysis in Study 2.

Evidence-based versus ongoing inquiry practices in 2006–07

Few school leaders or teachers in Study 1 even mentioned the word inquiry in their responses to questions about the key messages of the LPDP. School L was a key exception. In 2006, the literacy leader had set expectations about individual classroom inquiry by designing an “Inquiry into Practice” template for teachers that scaffolded the entire process, linking it to student data and requiring a set of monitoring criteria. Teachers had used the tool as part of their goal-setting process for appraisal. However, by 2007 this tool was not being used as systematically as it had been the year before. The literacy leader explained, “We spent quite a lot of time setting it up but, at the end of the day, teachers felt that it was just too much, and so we didn't pursue it.”

Evidence-based practices, however, were found across the school in 2006. The literacy leaders and principal had extensive knowledge of their actual data. They talked about entry points for their students and the progressions of the cohort during the year, indicating that their monitoring was not just about measurement of achievement but also about planning responses to these data. There was more evidence in 2007 that School L
was refocusing on those students who were still not achieving at each iteration of their data collection, and this reflected more of an inquiry for improvement approach:

Our STAR data, we used that as a sieve . . . so anyone below a Stanine 3 . . . in any class has a running record and that’s specifically for teachers to find out more about . . . these students’ reading behaviours, why things are happening, why they are not happening. (Literacy leader 1, School L, 2007)

The leaders reset their reading targets in 2007 to reflect stanines rather than chronological age and debated this difference with teachers. Leaders were interested in focusing more on data about actual comprehension skills and making teachers understand that they were responsible for supporting their students to achieve these targets.

There were some great philosophical arguments . . . . Because the teachers have become, actually: “I need to know what you want us to do because we are going to have to do it. Because we know that you are going to actually come and ask us about it.” . . . . A new teacher who’s come in mid-year [asked], “Where do you actually expect me to have got these children? We have just done a STAR test yesterday, and I am just marking them.” (Literacy leader 1, School L, 2007)

Inquiry as a set of questions about practice was beginning to be used as an explicit principle and as an expectation for all staff in School L. The induction and staff booklet for reading instruction for 2007 focused explicitly on underachievement and the need to fully interrogate their student achievement data:

The questions, discussions, and “solutions” are what make the difference, not the data . . . . We need to fully understand why the data is the way it is – what is the real issue – so that the right intervention is used. (School L, teacher induction materials, January 2007. p. 3)

Teachers were supported to set more finely tuned student goals as a result of data analysis, and leaders recognised that teachers needed professional support for this task. A memo to staff about STAR testing (not dated, 2007) provided a template for STAR data and their translation to learning intentions. The memo also provided information about links between data and the needs of students for each of the subtests. Other templates provided in the teachers’ school reading instruction booklet scaffolded more informal individual inquiry and included ways to monitor impact of new practices, suggesting daily checks such as undertaking a “quick survey of target students’ work,” having a
“mental focus on one student,” or “interviewing two students during the lesson.” So, after two years, inquiry as an individual pursuit had been adjusted from a formal project requirement to being integrated into more informal classroom routines.

In 2007, a more differentiated approach to School L’s school-wide goals was also established. Goals focused on shifts for particular groups:

Year 7 students with a STAR Stanine 1–3 in February will shift 1 or 2 stanines by October (we need to accelerate these students). In the STAR subtests, our aim is to: move “critical” and “at risk” students to “typical,” move “typical” students to “above typical.” (School L 2007 Resource Book about Reading Instruction, p. 4)

Nor were these goals static over time; they were revisited and debated with teachers. Goals for students in Stanine 8 and 9 were added as well in 2007.

**Leadership of inquiry practices in 2008**

The two-day visit in 2008 to School L revealed much more about the breadth and depth of co- and self-regulatory practices at school and classroom level. The school’s literacy leader from the time of the LPDP still remained in place. Her role was to support teachers in their classrooms, teach students who required extra interventions, and lead staff in their literacy professional learning. Her persistence and sense of urgency around improving achievement was pervasive in the school documentation and interviews. She had identified a worrying trend in their Year 8 student data over two years in reading comprehension. These students were not making the same progress as they had in the previous year, “particularly in the second half of the year”. In 2007, a school-wide target was set to ensure that Year 8 students “remained focused on their goals.” This had not impacted on their subsequent results, so the report to the board of trustees and a special memo to teachers in March 2008 raised the issue again and provided some theories as to why this may have been occurring. Her aim in both reports was to begin the conversation with teachers but not necessarily to focus on factors such as student motivation or learning behaviours. Instead she raised possibilities about teaching practices, stating, “Maybe as students become better readers, progress is harder to achieve (we’ve done the easier changes in Year 7)” (memo to Year 8 teachers, April 2008).

The literacy leader then required teachers of Year 8 students to develop an action plan for each classroom and provided them with some teaching strategies that they could action and monitor. There was no explicit suggestion in this memo that teachers should
learn more about reading instruction. It may be that the extensive “handbook” for teaching reading, devised by the literacy leader and supplied to each teacher, provided that new learning. It contained; research articles; sets of deliberate teaching responses to particular assessment findings from the STAR and other reading tests; school action planning for literacy; and the school and teacher inquiry models. Much of its content had been gathered from the LPDP, but school ownership was obvious with its up-to-date action plan and personalised comments from the literacy leader. One of the teachers interviewed in 2008 referred to this resource as “their bible” for reading instruction. Without any prompting by the researcher, this resource was also mentioned by the other teacher who was interviewed.

The school had earlier initiated learning “academies” that specifically grouped students according to their learning needs and interests. Those students with high literacy needs (STAR Stanine 1) were placed in the “adventure” academy, where they had guided reading instruction five days a week for 40 weeks. This development was a major alteration to the school’s structure and management and also the result of an inquiry-minded approach to improving learning and achievement at the school. School leaders designed the academy structure on the basis of research and collaboration with their community and reportedly tracked their students’ achievement and attendance figures systematically to evaluate the success of the innovation. By 2008, the adventure academy was made up of 60 students, 80 percent being Māori boys, supported by four teachers and three teacher aides. The academy approach was not a routine response to the problems they were observing: it required new learning for all involved and new ways of collaborating within and across academies, and it was staffed by teachers who had been specifically selected for their impact on student learning. However, achievement results for students who began in Stanine 1 and 2 in 2006 and in 2007 had gained, on average, 0.93 and 0.72 stanines, respectively, over two years. They were still some way short of the school’s target for this group of 1–2 stanines per year.

In School L it was clear that the deputy principal, working in concert with the literacy leader and numeracy leader, was reinforcing and amplifying an inquiry-minded approach in the school. For example, the deputy principal had established new routines for monitoring instruction:

That's a requirement, that all weekly planning comes in to me either digitally or in paper form every week . . . . And I comment on specifically the meeting of
group needs in literacy and specifically the meeting of group needs in numeracy.

. . . I'll do informal walk-throughs where I'll just check that the group work and the planning have been followed. (Deputy principal, School L, 2008)

Meanwhile, the literacy leader undertook to do all of the classroom observations, rather than this being shared with other leaders, as it had been during the LPDP. Expectations of what should be observed in every classroom were published in memos to teachers before the visit. In this way, the teachers all received individualised feedback on their practice from a specialist literacy practitioner, which continued to deepen their content and pedagogical knowledge. Teacher data were also collated in the same manner as they had been in the LPDP so that a school-wide lens remained as a measure of the quality of literacy instruction; these data were triangulated against student achievement data.

These findings in 2008 represented a step forward from evidence-informed instruction (where the focus was mostly on the students’ needs and where teacher responses were limited to what they already knew) to an inquiry approach, where new responses to problems were constructed at the school-wide level and monitored for effectiveness. There was also strong evidence of effective distributed leadership to support systematic inquiry that focused on what teachers were doing in response to their students’ achievement data.

Self-regulatory classroom practices

At classroom level, too, there were strong signs that the LPDP practices had been sustained. The literacy leader reported that teachers were still requesting observations of their practice, and this was confirmed by the two teachers who were interviewed in 2008 and whose lessons were audiotaped and transcribed. Both teachers emphasised an explicit focus on student data when asked about how they judged their own effectiveness in teaching literacy. One also confirmed the regularity with which the school achievement targets were revisited each year, set at the beginning of each term and reviewed halfway through each of the four terms. While the new teacher did not talk to other teachers about the specific data for her class, as she would have probably done while a part of the LPDP, she called on the literacy leader for support. She understood the need to adapt and change her practices because of the challenge and expectations of the school-wide targets for improvement.
However, there were differences between the two teachers. The new teacher was very aware of not having participated in the project and of needing to learn what they had experienced. She was highly self-regulated and often referred to the reading manual that the literacy leader had collated to improve her practices. She commented:

[People would say,] “when I was on the project” . . . or “when we were on the project” . . . I heard people saying how they were using . . . the data from STAR, and I thought oh gosh, because I knew I wasn’t, so then I got the bit that told me about STAR and I just started to work it out really. (New teacher, School L, 2008)

The teacher who had participated in the LPDP relied more on the literacy leader to provide “ideas” than on any formalised learning. He was able to describe a theory of practice about inquiry as asking questions of practice and “always refining and improving,” yet this was inconsistent with other responses that he offered. He admitted that his personal learning focus had moved on from literacy to school organisational management. So one teacher was observed working in an improvement paradigm and the other appeared to be operating within a maintenance model.

In terms of fidelity to the LPDP classroom practices, both teachers shared specific learning intentions with their students but variously built success criteria so that students might be able to self-monitor their learning. Neither teacher consistently shared any previous achievement data with students to justify their learning focus, but both did require students to explicitly explain their strategies for knowing. One out of two and two of three students, respectively, could identify their next learning step in their interviews following the lesson.

Diversity of classroom practices was also reported in the literacy leader’s analysis of 16 teacher observations in October 2008. She wrote that most teachers showed “some” evidence of each of the five “elements of best practice as used in the LPDP.” She commented that those teachers with “strong” evidence were among those teachers who were “here for all or some of the LPDP.”

**Coherence of effective instructional practices**

This section of the School L case study is structured in a similar way to the previous section. Evidence related to coherence of effective instructional practices is reported firstly for 2006–07 and then for 2008 from various participant groups.
Transfer of learning from the LPDP in 2006–07

In 2006, the principal had articulated that a key reason for moving to a new numeracy focus in 2007 was to build links across the two projects about learner needs and teacher responses. Literacy leaders had reiterated that this transfer of ideas was already occurring amongst teachers:

One of the big outcomes [of the LPDP] was that teachers started to say . . . “I want my maths teaching to be as informed as my [literacy] teaching” . . . [and] “I am finding it so difficult to know what to do with these kids when I don’t have data to support my teaching decisions.” (Literacy leader, School L, 2006)

School leaders also identified another element of transfer. They reported that teachers were now feeling more comfortable with revealing what they did not know in their new learning focus as a result of their involvement in the LPDP. Despite this, School L had taken a passive stance to the transfer of ideas from one project to the next in 2006. The leaders had not discussed their experiences in the LPDP with the Numeracy Development Project providers or what learning they might transfer. Nor were the leaders and teachers asked to discuss this. Reportedly, the two external providers of professional learning had not actively supported the handover by promoting coherence and alignment of the ideas inherent in each intervention.

A third area of transfer considered by the leaders in School L was how new learning in numeracy would need to be monitored for its impact on achievement. One leader explained:

There’s actually a whole philosophy, isn’t there, behind the literacy, that we have kind of just embedded. So that whole evaluation [idea] . . . [if] we have something that’s not working now, it’s just automatic: “Oh this is not working, what are we going to do about it?” (Literacy leader, School L, 2007)

By 2007, school leaders appeared to be more explicit about the links between professional learning projects. They had voiced concerns about the Numeracy Development Project with their facilitators, although the talk was more about how the initiative was being delivered than about teaching and learning approaches between the two projects. Nonetheless, some learning from the LPDP was being integrated into expectations about teacher professional learning.
In thinking about improving their writing instruction, school leaders returned to earlier learning that they had undertaken in using the national exemplars for English, and they also blended approaches from the LPDP in their design for further professional development in literacy. The literacy leader commented that teachers were “far more astute in terms of their levelling of the writing samples and then basing their teaching on next steps, so there has been a direct transference of those skills [in reading] into writing” (Literacy leader 1, School L, 2007).

Alignment of school-wide and classroom practices in 2008

Elmore (2002b) explains how “after living with the same wallpaper for a certain number of years, people cease to see it,” in the sense that schools fail to examine what has become the norm or to question assumptions that they may have made for many years. In 2008, the “wallpaper” at School L could hardly be missed. One whole staffroom wall was dedicated to a display about literacy, but the messages went a long way to establishing a coherent set of principles for effective teaching across the curriculum. Quotations from New Zealand researchers and strategies for classroom management sat alongside specific references to effective teaching and learning, such as links to prior knowledge, shared learning intentions, effective feedback, catering for diverse needs, and the use of deliberate acts of teaching. One quotation suggested:

We need to move away from considering achievement data saying something about the students and start considering achievement data saying something about our teaching, as cues for teacher action, particularly teaching in a different way.

(School L’s staffroom wall, 2008)

In this way, School L had made their own links to the theoretical principles that they were supporting in their practices as the “way we do it around here.” The latter was a phrase heard often in the interviews with leaders and teachers in 2008, indicating a strong sense of ownership of such practices and illustrating one of Coburn’s (2003) key dimensions for scaling up reforms.

By 2008, the school had embarked on another new professional learning initiative, one that other schools in Study 1 had become confused about in regard to the terminology being used. Inquiry learning processes for students were being established in each classroom, where the changed focus for teacher practices was on students’ generating their own questions as a way to integrate and deepen curriculum knowledge.
When the researcher asked the literacy leader about inquiry as an approach to professional learning, she immediately articulated the links between teacher inquiry and student inquiry. In addition, it had become apparent in her informal teacher observations that the student inquiry learning approaches were, in some cases, replacing explicit reading instruction. The literacy leader then undertook her own inquiry, surveying staff to see how this new learning had impacted on literacy practices. She published her results in a memo to staff:

> It appears that many teachers moved from directly teaching to the needs [that were] indicated in student data to teaching [the] skills students need [in order] to use the SAUCE [pneumonic used for inquiry learning model] . . . . Guided reading was used less and shared reading and even whole-class approach[es] were more prevalent than previously. (School L, Memo to all teachers, July 2008)

The results had shown that 10 of 14 teachers surveyed had not remembered their literacy professional goal and that learning intentions shared with students were now about the SAUCE model rather than being derived from student data. In a sense, this new professional learning project was a direct challenge to the sustainability of the LPDP, and the literacy leader’s response was to highlight this tension and quickly resolve the conflict between teaching practices:

> It is expected that teachers will use guided reading as the main approach (research has shown this to be effective in raising student achievement), refocusing on explicitly teaching student skills they need to become fluent readers with excellent comprehension. Some of these skills will be those needed to follow the SAUCE model, but [they] need to be balanced with any other skills missing . . . . Refocus on the plan to get these students achieving the school-wide target, or better, by the end of Year 8. The SAUCE model may be a vehicle that helps us do this. (School L, Memo to all teachers, July 2008)

One other example highlighted this drive from leaders to align the support for effective learning across the school. A regular report sent to parents about literacy aimed to make learning activities at home and school more aligned, included descriptions about what good readers do, and explained the STAR test and their individual child’s results. It also provided strategies for parents to support their child’s reading.
Although deliberate actions were being taken to realign new learning foci to previous shifts in literacy practices, this was seldom matched with other explicit leadership actions to promote coherence across other curriculum areas. The numeracy leader, like the literacy leaders in previous interviews, described sustainability of the Numeracy Development Project entirely in terms of the project boundaries, the eight stages of progression, the one-on-one diagnostic tests, and the materials and equipment used in teaching.

The numeracy and literacy leaders were asked whether they were making connections between the projects for teachers and students. They had not talked explicitly about any links, but could describe a number of practices that were now being shared across these subject domains. For example, they reported that the “modelling book,” first used in the Numeracy Development Project, was also being used in literacy contexts as a way of keeping track of learning intentions and teaching exemplars for each instructional group. When prompted, the deputy principal outlined similarities across numeracy and literacy, such as evidence-based teaching, understandings of the progressions that students make in their learning, and teaching to groups with similar strengths and needs.

Coherence of instructional practices at the classroom level

In School L, there was more substantive evidence, in the new teacher’s lessons, of coherence of effective instruction. In the numeracy lesson, she made real-world links for students as they practised their skills in using decimal place values. Similarly, in the literacy lesson, she reminded students about a previous narrative they had read, as they practised looking for good examples of characterisation. She made a link from the reading comprehension skills they were learning to their applying similar models to their writing. There was a striking similarity in the language she used with each group across her numeracy and literacy lessons. For example, she asked one student, “What was the guide that you had in your head when you made that decision to go to this number?” Then, in her reading lesson, she asked her instructional group, “What sort of pictures came into your mind when you had a title like that in your story?”

Not surprisingly, when asked about what their teacher did that was similar in both lessons and what they did to help themselves learn, her students could identify these connections across literacy and numeracy lessons and the links made by the teacher
between what they were reading and their writing skills. The following connections were identified in interviews with two of her students:

To do imagery and to do things in your head. (Student 1, New teacher lesson, School L, 2008)

To have strategies . . . to help us understand and make it easier for us to remember them by. (Student 2, New teacher lesson, School L, 2008)

The teacher who had participated in the LPDP was far less obvious in the links that he made for students in each lesson and this, in turn, was reflected in his students’ responses to similar questions. They did, however, identify that he always helped them understand the main purpose of the lesson, writing WALTs\textsuperscript{25} on the board in each lesson.

Leaders and teachers in School L had forged some connections across their professional learning and applied them to some aspects of teaching practice. The literacy leader, in particular, had played an important role in realignment of new learning in the slip stream of the LPDP. Where they were explicit, these links were mostly limited to the literacy domain. However, there was a strong sense of “how we do it around here” being deliberately designed into classroom routines. The teacher who had participated in the LPDP was asked about similarities across his teaching in both literacy and numeracy:

There is certainly the expectation [that] this is what we do at [School L]. I think it would be very hard to go into any classroom, including . . . four to five PR [provisionally registered] teachers, even in their classrooms, they weren't here for that [the LPDP], but this is how we do it at [School L]. (LPDP teacher, School L, 2008)

Discussion

While some aspects of the LPDP had not endured in their same form, fidelity to the core beliefs and principles of the project remained evident in School L, despite the fact that only seven teachers out of 18 had actually experienced the learning firsthand. Spillane (2006) is careful to explain that distributed leadership as delegation of leadership tasks is not, in itself, a model for improvement. Rather, it is how leaders act in concert to bring about change that makes the delegation of tasks effective. There was

\textsuperscript{25} WALT is an acronym for “we are learning to” and is a strategy adopted from the assessment for learning literature to identify explicit learning intentions for each lesson that are formed with students and focused on their needs as learners. It was not explicitly used by the LPDP, but learning intentions were emphasised as integral to effective literacy practice.
strong evidence of systematic and collaborative monitoring of student achievement data and teacher practices by leaders in this school. In particular, key tasks for the literacy and numeracy leaders were related to their individual strengths and roles in the school, but improvement targets and teaching to student needs were prioritised, and the deputy principal connected the two curriculum areas and reinforced teachers’ actions with regular monitoring and feedback to them.

Strategic actions to promote coherence of effective instruction were not necessarily explicit to teachers, but were more associated with consistency and breadth of effective practices in literacy and numeracy. The LPDP practices had been explicitly codified in school documentation, and leaders continued to set clear expectations of their being employed in all classrooms. School documentation on numeracy materials also supported teacher practices, and teachers referred to these documents and used them to support their individual knowledge gaps.

Since the LPDP, School L had moved beyond a maintenance phase, operating mostly in the sustainability phase that Century and Levy (2002) describe. Leaders had evolved their inquiry into underachievement into a routine school-wide practice and were consistently monitoring achievement targets and engaging teachers with ongoing discussion about their relevance (Copland, 2003). School leaders and teachers were “evidence aware”: that teaching practices had to be supported by research evidence of their effectiveness and that assessment evidence must be used to inform them of their students’ needs. School L approached achievement issues with innovation in mind (Hannon, 2008) and were intent on moving students with low achievement entry levels towards national norms at each iteration of school-wide data analysis. Their academy approach to structuring learning and responding to underachievement could be considered equivalent to those “adaptive” responses to specific problems reported elsewhere (Lai et al., 2009; Wiliam, 2007). However, while all students continued to make achievement gains at a higher level than even the LPDP cohort, the impact for at risk students was still not enough for them to reach Stanine 5. The independent evaluation of the LPDP argued that teachers may need to have access to specialist reading intervention teachers to support them to lift achievement of these at risk groups, particularly those students in Stanines 1 and 2 (McDowall et al., 2007). This could explain why School L was unable to continue their acceleration of at risk students beyond the levels they had made for all students.
Chapter 8 – Results and Discussion: Four cases

School H

On two counts, School H presented an interesting challenge to the conceptual framework for sustainability. The school had scored relatively highly in the qualitative data for both co-and self-regulated inquiry and coherence of instructional practices gathered in 2006, the first year after the school’s participation in the LPDP. Students in School H also made a net mean gain of 0.61 stanine by the end of 2006 but by the end of 2007 this gain was reduced to 0.12 so it did not meet the benchmark criteria for sustainability. School H’s acceptance into the LPDP again in 2008 also challenged the notion of “going it alone” that has traditionally been associated with sustainability. The school’s professional learning focus in the 2008–09 LPDP shifted from reading to writing, although the LPDP had changed from its earlier orientation of focusing on only one aspect of literacy to include content and pedagogical knowledge for both reading and writing. This meant that facilitators supported schools to examine their data in both domains and to make links between reading and writing scores in their analysis of students’ strengths and needs. Teachers would also be supported to make links between their teaching practices for reading and writing and to emphasise with their students where their learning in reading could be transferred into their writing tasks.

Patterns in student achievement

School H, like School L, had relatively low student entry levels, with a mean stanine score of 4.52 in reading comprehension for their Year 7 students as they entered the school in 2006 (see Table 24). Within one year they had made a mean stanine gain of 0.61 (n=80), almost twice the gain that students in the previous cohort had made over two years. However, that gain was almost completely eroded in 2007 when these students were in Year 8. Overall, the cohort achieved a net mean stanine gain of only 0.12 after two years. A third cohort of Year 7 students (n=92) were tested at the end of 2007 and managed, on average, an increase of 0.28 stanine by the end of Year 8, slightly less than the gains made by students in the 2004–05 cohort, when the school originally participated in the LPDP.
### Table 24

*School H* STAR mean stanine gains 2004–08

<table>
<thead>
<tr>
<th>Time period</th>
<th>2004–05</th>
<th>2006–07</th>
<th>2007–08</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=85 Net gain</td>
<td>n=80 Net gain</td>
<td>n=92 Net gain</td>
</tr>
<tr>
<td>T1</td>
<td>4.26</td>
<td>4.52</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td></td>
<td>5.13</td>
<td>0.61</td>
</tr>
<tr>
<td>T4</td>
<td>4.62</td>
<td>0.36</td>
<td>4.64</td>
</tr>
</tbody>
</table>

The STAR data in School H had been gathered at only two points of time for the 2004–05 and 2007 cohorts, so it was difficult to evaluate any patterns within each period, such as possible summer effects. Figure 8 illustrates the successive cohorts of students in School H over 2004–08. These results were in sharp contrast to those of School L, where gains had been sustained and improved while independent from any external support and with a significantly higher turnover of teachers.

![Figure 8. School H reading comprehension results 2004–08.](image)

Figure 9 indicates the patterns of progress in reading for those at risk students in Stanines 1 to 3. Whilst teachers were participating in the LPDP in 2004–05, these students made slightly less gain than that made by all students in the same period, a gain of 0.3 stanine compared to a mean stanine gain of 0.36 by their peers in the school. This did not match the trends shown in the 91 schools in this first cohort of the LPDP, where at risk students benefited most with a mean shift of 1.1 stanine over two years. (Bareta & English, 2006). In 2006-07 there were three data collection points that served to more
graphically highlight the reduced gains in the second year for those at risk students 
\( (n=21) \) who had been in Stanines 1-3 at time 1. By the end of 2006 at risk students had 
gained, on average, 1 stanine, compared to 0.6 for all students in their school. These 
gains fall away dramatically in the second year: almost all students still remained in the 
at risk band at the end of 2007. At risk students \( (n=33) \) did return to earlier levels of 
progress when the school was once again participating in the LPDP.

\[
\text{Figure 9. School H student achievement gains for at risk students 2004–08.}
\]

\textit{Inquiry for improvement practices}

In Study 1, School H had continued the patterns of inquiry they had begun during 
the LPDP into 2006, discussing student achievement data in team meetings and using 
these data to drive decisions about the focus for ongoing teacher learning. They had 
intended at the end of that year to refine their investigation of literacy data into Pasifika 
students’ achievements in the school, following a co- and self-regulated pattern of 
refocusing on those students who had still not made sufficient progress.

\textit{Evidence-based versus ongoing inquiry practices in 2007}

School H leaders expected to achieve “similar” shifts for students in reading 
comprehension in 2007 as they had when participating in the LPDP. They did not talk 
about improving on those shifts, as might have been expected if co- and self-regulated 
inquiry were embedded as a habit of mind. The literacy leader explained that they had 
focused on the lowest 25 percent of students, investigated ethnicity data, and “pulled 
apart” the subsets of the STAR test to inquire into particular weaknesses. Teachers had
responded to these data in 2007 by focusing on paragraph comprehension and would track the impact later in the year. There was no mention of the earlier inquiry about Pasifika students, nor did the literacy action plan in 2007 include focused goals for student achievement. Neither the strategic plan nor the literacy long-term plan mentioned previous data used for setting goals for teacher practices. Evidence-based practices may have endured through to 2007, but the idea of generating new questions about further teacher learning needs for reading comprehension had apparently not occurred to school leaders.

In 2007, the principal and literacy leader viewed the school as being in “desperate need” of further external literacy support. The literacy leader did not want to venture into a writing focus with their teachers without external leadership. School H had also engaged with a local cluster focused on literacy, and leaders used this as an opportunity to apply again for the LPDP. In the year that this cluster has taken to gather and analyse their cluster-wide evidence, School H’s principal was correct in his analysis that his students had lost ground:

There is a cluster initiative going on . . . and it’s taking a lot of time, so we’ve stood still really in the sense that this . . . reading stuff is here, but we need to be into this writing and we’re still are where we are at now. LPDP has been allocated to four of the schools, and we are not one of them, and so we are in a bit of a catch 22, and three of the other schools in the cluster have already done LPDP writing. (Principal, School H, 2007)

Leadership of inquiry practices in 2008

In 2008, the principal and literacy leader both regarded their cluster involvement as a potential source of expertise to support ongoing sustainability of their collaborative learning as well as a way to increase accountability. There were school-wide targets to lift student achievement in reading that were in place. “Moving students one to two stanines” in reading was described in interviews with school leaders, although this target was not mentioned specifically by teachers participating in the study. The principal commented that he was not even sure the target was realistic now that they had compared themselves to others in their cluster for writing data and judged that their students’ achievement was “not too dissimilar” (Literacy leader, School H, 2008). It seemed that the principal, at least, was setting lower expectations for improved achievement.
The teacher inquiry and knowledge-building cycle diagram from *Teacher Professional Development: Best Evidence Synthesis Iteration* (Timperley et al., 2007) had been pinned on the wall by the LPDP facilitator. When asked how this concept might be working in the school, the principal confused this with inquiry learning approaches for students. However, the school’s leaders were aware of the significant issues that they faced with underachievement. They talked about their data with specificity and were actively investigating ways to make improvements. In 2008 they had created an extra class, utilising some extra funding that they had available.

[W]e created the situation where we pulled out kids who we thought would benefit from smaller numbers and would benefit from the teaching and would not necessarily be behaviour [problems], but that’s proven a bit of a myth to some extent and, of course, it lowered the ratio in the other classes, and so for the second half of the year, the ratio in the classes has been really good, and I would expect that to be making a difference . . . . So we’re putting extra staff in there to work with them, and [name of teacher] is also a good teacher. (Principal, School H, 2008)

The numeracy leader was asked whether they were analysing the impact of this structural change on their students’ achievement. Her response indicated that this was not formally planned, “but we could do, school-wide data, yeah.” The principal did not have a formalised improvement theory but described the ongoing development of teacher knowledge as his priority for supporting improved student achievement. He did not link this set of beliefs to any systematic inquiry, nor did he describe improved content and pedagogical knowledge in relation to the setting up of the new classroom. The cluster that the school had joined in 2007 was just beginning to develop their theory for improvement as a collaborative process, again with external support.

**Self-regulatory classroom practices in 2008**

At the classroom level, both teachers illustrated evidence-based approaches in their lessons. For example, the new teacher used several strategies in her lesson on number patterns to check on students’ understanding as the lesson progressed, adjusting her instruction as a result. For example, she instructed her students to:

Give me a thumbs up about how confused you are right now. So this is “I think I’m alright,” “I’m a little bit fuzzy”. . . . Can you give me another thumbs up to
show me how well you understand [name of student]’s strategy, and we’ll test it, and we’ll see if it works. (New teacher, School H, 2008)

The LPDP teacher discussed recent achievement data with her students as she introduced the learning for her lesson. In general, though, there was less emphasis on student co- and self-regulation in the LPDP teacher’s lessons than in those of the new teacher. The new teacher created multiple opportunities for students to contribute their strategies for working out the numeracy problem or to identify use of implicit and explicit characterisation in a text. Consequently two of her three students interviewed after the lessons could identify their learning and had specific next steps. One student used a peer to support her learning in numeracy:

I got confused when other people said their ideas, and I knew my one and then I went to my table and he taught me it . . . . I got to know his pattern, so I could use that as well. (Student 2, New teacher lesson, School H, 2008)

The post-lesson interviews with teachers confirmed that the collaborative talk about achievement data in teams had fallen off around reading, but they were both able to talk about their own students’ shifts. The LPDP teacher identified that she had sustained the explicit teaching focus from her earlier involvement in the project but struggled to remember the achievement targets she was working towards in reading. The new teacher did not have the “extremely low group that other classes have,” but she had targets and was mindful of rates of progress and links to her effectiveness as a teacher. She explained:

[T]here’s been a couple of times since I started where I have really been focusing on something and then I’ve tested the water and they haven’t moved . . . . It happened twice, I thought that they would, you know, I’ve done so much. I’ve actually gone to someone and said, “Why isn’t this working?” And they’ll help me . . . . And then we have observations. (New teacher, School H, 2008)

Coherence of effective instructional practices

According to analysis of qualitative data from 2006 in Study 1, School H had described sustainability of the LPDP as extending into other areas of the curriculum, a broader view than that of most other schools in the original sample.
Transfer of learning from the LPDP in 2006-07

The 2006 literacy action plan for School H had revealed that one of its goals was for teachers to have an “understanding of theories of learning” as the basis of their work in literacy, while still linking to specific knowledge about literacy teaching and learning. School leaders had specifically identified learning processes for teachers that would be transferred to new professional learning.

However, in the following years, School H did not strengthen coherence of instructional practices. The 2007 data revealed more espoused theories than actual theories-in-use about coherence of effective instructional practices. For example, when asked about the key ideas that the school was sustaining in 2007, the literacy leader focused only on the context of reading instruction and not how these could also be transferred to writing or across other curriculum, as had been the case in the interviews with leaders and teachers in 2006. Although the principal had recognised that the model was transferable to other areas in 2006, he appeared to regard it as an externally driven model in 2007 and one that was not intended to be adapted:

Now we know that the LPDP model is . . . a structure that can be used – and for us it was a successful structure . . . but as part of our cluster discussion we are now aware that there actually are other models, and LPDP is just one model, and there are other providers, and there are other people out there. (Principal, School H, 2007)

The literacy leader went on study leave in 2007, and School H appeared to have relied more on teachers’ informal talk than on specific learning opportunities to establish coherence of instructional practices across the school. Leaders reported that there had been some discussions at parent evenings about the Numeracy Development Project but that since then there had been no formal attempts to bridge home and school opportunities for literacy learning for students. Nor were the literacy practices from the LPDP codified in the same way as School L had established, with its explicit expectations around practices, its reading “handbook,” and its observation frameworks. There may have been other accountability mechanisms at work in School H, such as appraisal and observation frameworks, but these were not identified in any documents gathered at the time of the visits. Sustainability appeared to have been left to chance over 2007 in the absence of the in-school literacy leader, whose role had been specifically designed to support and mentor teachers in effective literacy teaching and learning.
Alignment of school-wide and classroom practices in 2008

The school had participated in the Numeracy Development Project in the years prior to their involvement in the LPDP. As with other schools in the study, leaders had not talked explicitly with providers or their teachers about how the two projects complemented one another in terms of their underlying principles of effective teaching and learning. In 2008, the Numeracy leader commented:

I think we’re actually quite private about our own pedagogical stage . . . . We do talk about things in teams, for example . . . specific things, you know, about a reading group or a strategy that might work or a text that might work really well, but . . . not pulling back and doing it more generally, somehow. (Numeracy leader, School H, 2008)

School leaders had been able to reflect on similarities in classroom pedagogy, though they had not worked with teachers to discuss any commonality of these practices across curriculum areas. They nominated conferencing with individual students, working with similar-ability groups, using explicit learning intentions and success criteria with students, and sharing data so that:

[students are] able to see what areas they’re doing well in and areas they need to work on . . . . It allows them to be able to set goals with the teacher. (Literacy leader, School H, 2008)

At the classroom level in 2008, the lesson transcripts confirm that these approaches were visible across numeracy and literacy lessons. Both the new teacher and the LPDP teacher had used explicit learning intentions in their introductions to students about their learning in reading and in numeracy. Both teachers made connections from that learning to other areas of the curriculum and worked with small groups to focus their instruction and monitor students’ understandings as the lessons progressed. However, when asked, students from each classroom identified very few commonalities between their literacy and numeracy lessons. Mostly their responses were about working in groups or more generally about how the teacher explained the learning. For example:

She doesn’t use a lot of complicated speech, she just, she tells it how it is, just like that, yeah. (Student 1, School H, 2008)

The teachers both identified the same ideas as their school leaders in terms of similarities in their teaching across the curriculum. Interestingly, the new teacher, who
had joined the staff in 2007, perceived her “newness” as supporting sustainability of the LPDP:

It’s probably good for them to have people that didn’t do it because then they get to explain to us . . . . They were all doing it the same way, and we were quite confused. So we got quite a bit of help and everyone . . . was still quite enthusiastic about [the LPDP]. [Name] was my team leader last year. . . And he was really enthusiastic about everything he’d learnt. . . I wasn’t really getting mixed messages, which was nice. . . There was lots of talk going on about the deliberate acts of teaching. Lots about WALTS . . . and think alouds. (New teacher, School H, 2008)

By 2008 the school was once again participating in the LPDP, and leaders had noticed some different approaches in their relationship with the LPDP facilitator. The literacy leader explained:

I suppose this time, it’s been more of a focus on trying to develop everybody in the role of a leader. . . so it’s not a matter of someone actually coming in and leading it per se, it’s matter of developing those around. . . We’ve observed with her and she’s looked at our observations, and we’ve given feedback, and she’s observed feedback, and so on, but it’s more working together collectively. . . part and parcel of that shared ownership or shared leadership. (Literacy leader, School H, 2008)

Once again teachers were focused more formally on effective literacy practices, now in both reading and writing. Lesson observations became more consistent and formalised. School H’s return to the LPDP had, of course, “muddied the waters” in terms of identifying which school-based conditions may have been impacting on the sustainability of their previous learning in the project. However, the earlier indications from 2006 around breadth and consistency of reading practices had not been sustained. In 2007, the principal had recognised this lack of consistency, explaining there “was some slippage I felt this year, where different classes were doing different things . . . when I would have expected them all to be doing the same things, but obviously different lessons within that context” (Principal, School H, 2007).
Discussion

Movements in leadership in School H and a fragile understanding of inquiry practices appear to have combined to reduce the longer-term sustainability of the gains and practices established in the LPDP. Nonetheless, there was still evidence across school-wide and classroom practices of the continued use of student achievement data to inform teaching programmes. Any earlier sense of coherence of effective instruction was confined to specific practices, particularly explicit sharing of learning intentions and awareness of the need for students to understand the strategies they were using to support their learning. Whether these practices can be attributed to learning in the first round of the LPDP could not be confirmed since the teacher new to the school had stronger indications of such practices than the teacher who had participated in the project. Teachers did make links to their effectiveness by using the student data, but the knowledge-building component of the inquiry cycle was not sufficiently supported in 2007 to make headway with student achievement levels. The talk was more about sharing and employing new ideas in practice than systematic inquiry into their effectiveness. The antecedent and mediating variables at work in this school context, particularly in 2007, appear to have combined to produce a situation for teachers that the principal described as “treading water.” The student achievement data reflected this same pattern in that year.

School H spent all of 2007 trying to access further external support rather than focusing on ways to generate learning from within. In 2008, the school was able to access further support from an LPDP facilitator within a cluster schooling improvement initiative, and this established another potential boost to pedagogical content knowledge. Rates of progress for the school’s students in 2008 were akin to those in 2004–05.

School A

School A was selected for Study 2 because it had scored highly on the inquiry dimension in Study 1 but data indicated much less evidence of coherence of effective instructional practices. The school also had high leadership turnover during the period of the study.

Patterns in student achievement

In 2004, Year 7 students in School A began, on average, 37 points below the mean score for New Zealand students in asTTle writing. Table 25 indicates School A’s mean gains in asTTle writing points above normed year groups for 2004–08.
years, students were, on average, 127 points above the mean for their year group. A new cohort of students in 2006 began their year at about the New Zealand mean for Year 7 students but higher than the previous cohort of students. However, they ended their first year 50 points above the mean. By the end of 2007, these same students, now in Year 8, had slightly slowed their pace of progress but still managed to be, on average, 0.8 of a standard deviation above their previous levels of progress. In 2007, students made a similar first year gain and then reduced that rate of progress in 2008. The 2008 cohort of Year 7 students managed a mean gain of 36 asTTle points above the expected gain calculated from the test norms in one year. Figure 10 illustrates patterns of achievement in writing for three successive cohorts of students. School A had sustained and maintained their student achievement gains since participating in the LPDP.

Table 25
School A net gains in asTTle writing points above mean gains calculated from asTTle norms

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n=200 Net gain</td>
<td>n=237 Net gain</td>
<td>n=266 Net gain</td>
<td>n=277 Net gain</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>-36.6</td>
<td>-3.3</td>
<td>-12.4</td>
<td>-11.27</td>
</tr>
<tr>
<td>T2</td>
<td>47.3</td>
<td>50.6</td>
<td>57.94</td>
<td>70.34</td>
</tr>
<tr>
<td>T3</td>
<td>56.8</td>
<td>29.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>91.0</td>
<td>127.6</td>
<td>74.6</td>
<td>77.9</td>
</tr>
</tbody>
</table>

Note. 100 asTTle writing points = 1 standard deviation

Figure 10. School A asTTle writing results 2004–08.
Students in School A had made large gains during the two years of their participation in the LPDP. The next two cohorts of students had higher entry scores in writing but still reached similar mean differences from asTTle norms after two years. Both later cohorts had made a greater than 0.5 standard deviation gain, the benchmark for sustainability of gain in writing for asTTle arrived at for this study. The rates of progress for the lowest 20 percent of students (at Time 1) at School A over three cohorts are shown on Figure 11. The at risk student in each cohort made similar gains over 2006–08, almost double the mean gains in asTTle writing recorded for all of School A’s students over the two year periods.

![Graph showing School A asTTle writing results for at risk students 2004–08.](image)

**Figure 11.** School A asTTle writing results for at risk students 2004–08.

*Inquiry for improvement practices*

School A had the highest score on the self-regulated improvement practices dimension of the Study 1 data. The interview with one of the new teachers in the school in 2006 was typical of responses to questions about the link between data and teachers’ effectiveness. Despite having only participated in the LPDP for one year, she understood that student achievement data could inform her of her own learning needs. When asked about how she used classroom literacy data, she responded:

I need to be flexible, too, and say, “Maybe there’s something I’m not doing that I need to change to help the kids become writers,” so basically it’s [data] for the kids and for me as a growing teacher as well. (New teacher, School A, 2006)
Evidence-based versus ongoing inquiry practices 2007

As the literacy leaders discussed their 2007 data, it was apparent that they were disappointed with their results and recognised that their student achievement had levelled off in the second year for this next cohort of students. Further probes about their inquiry in 2007 revealed that School A had certainly not embedded practices associated with inquiry as routines. After undertaking an inquiry into students performing at the higher levels of asTTle writing, school leaders had made a decision to place these students with “effective teachers.” They did not describe whether they had evidence that these teachers were effective in raising student achievement, or how they would monitor regularly for the effectiveness of this decision.

The literacy plan for 2007 appeared to be driven by actions rather than outcomes. Leaders voiced frustration at the considerable time lags evident between data collection, analysis and working with staff on next steps for students. There was no evidence in their interview responses about using student achievement data to reveal what teachers needed to learn, yet this had been strongly emphasised in their Study 1 transcripts in 2006. The student achievement targets in the annual review documents for the board of trustees remained the same as they had been in 2006 and were framed in terms of the asTTle sub-levels. The targets did not describe literacy knowledge, or skills that would represent these sub-levels, nor did the review documents refocus on inquiry into those students who were still not achieving adequately in the school. The review described the current best-fit level and commented that this was “comparable to the same time last year.” If inquiry had been an embedded practice in School A, new questions would need to have been asked about those students at risk, particularly where targets had not been met. Leaders’ responses for improvement were more about conformity to approaches than about considering the problem more thoroughly:

This is the second year that we haven’t met our targets. It is obvious that quality guided reading programmes still need to occur in classes. The provision of such programmes needs to be carefully monitored in both Year 7 and 8 classes.
(Literacy leader, School A, 2008)

A report to their board of trustees in 2007 outlined a general need for better leadership and focus on effective literacy pedagogy. Individual classroom data had been collated, but there was no attempt, at least publicly, to analyse what actual teaching had occurred in those classrooms that had supported students to exceed the school’s targets
for writing. Minutes of teacher meetings indicated that actions such as testing and literacy observations were discussed but that there were no links to any larger school-wide inquiry into literacy practices and/or achievement. School A had a stronger sense of inquiry than other schools in Study 1, but this appears to have been eroded a year later.

**Leadership of inquiry practices in 2008**

In 2008, another close lens was placed on the nature of co- and self-regulated inquiry practices in the school. The principal and literacy leaders once again had precise knowledge of trends in their students’ literacy achievement over that year and certainly looked to effective teaching as the lever to solve issues in underachievement. They drew on national research evidence to discuss ideas for improvement and talked of measuring effectiveness with reference to their student data. Their theory of action for inquiry was aligned with much of what was identified as an inquiry habit of mind in the LPDP (Bareta & English, 2006). The principal explained their beliefs about how to ensure ongoing improvement in student achievement as:

> Having teachers who are effective practitioners . . . . [W]hen I have an effective practitioner it means someone who takes cognisance of the professional development that’s put into them. They take cognisance of the data that’s passed on to them in regard to their own performance as a teacher. They take cognisance of what the student achievement data is telling them. (Principal, School A, 2008)

While leaders claimed that they were “challenging” their teachers by holding them accountable for their results, the literacy leader indicated that the analysis of student achievement data with staff was investigating the outcomes for different subgroups of students rather than individual classroom results. She also reflected that “we’re still inquiring at the student level rather more than at the teacher level.” While some teacher goals may have been set using their achievement data, this was not a systematic practice in the school. The identification of target students and their needs as a means of supporting collaborative learning between teachers had also fallen away in the 2008 period. Teachers had identified a need to better understand how to analyse and respond to reading assessments being used in the school, and external expertise had been utilised to support them in 2008. While these data certainly indicated a strong belief in evidence-based practices as part of the inquiry process, there were important gaps in relation to co- and self-regulated inquiry. For example, focused inquiry into underachievement and
knowledge-building were not connected or systematic. This became more evident in the data gathered from their two teachers participating in the study.

*Self-regulatory classroom practices in 2008*

The lesson transcripts in the writing and numeracy lessons all followed similar formats. Lessons began with a short organisational episode, conducted with the whole class, in order to introduce the learning activity, and this was followed by focused instruction with smaller groups, based on their specific needs. The teachers either had specific learning intentions displayed or talked with students about what the focus was for their learning. The new teacher built success criteria with her students for writing a “catchy headline” for their newspaper article, but students were not asked to use these independently. Two out of three students from each class could not identify their next steps and were reliant on their teacher for evaluating their success in the lesson. The LPDP teacher was aware of the targets set for improvement but felt that these were not always “realistic,” especially for those students who were achieving at the higher levels in writing. The teacher new to the school was also concerned about her low-achieving students and had made that her focus. However, she did not link that focus with goals that she had set for her appraisal as a provisionally registered teacher.

In 2008, School A had begun to inquire into the levels of achievement for their Māori students, using particular research tools to measure students’ engagement with their learning. The analysis was yet to filter down to individual teachers, since neither of the teachers participating in the study mentioned any particular shifts in practices that they were making as a result.

*Coherence of effective instructional practices*

School A had scored in the lowest quarter in relation to other schools in Study 1 on the coherence dimension. The interviews with leaders and teachers, the lesson observations, and the meeting and document analyses had not indicated that strong and explicit connections were being made between the learning from the LPDP and other areas of the curriculum.

*Transfer of learning from the LPDP in 2006–07*

School leaders had been very deliberate in 2006 about their decision to delay the start of a numeracy focus so that staff could embed their learning from the LPDP. However, when the school did take up this new focus in 2007, school leaders were
critical of their involvement with this project. The Numeracy Development Project, in their view, had not met their expectations of delivering effective professional learning for their teachers. This new focus was mostly perceived as competing for time with professional learning in literacy rather than an opportunity to leverage new learning across two or more knowledge areas. Leadership changes had occurred in the school. The principal who had led the school while it was part of the LPDP had moved on, and the deputy principal took up the role in 2007. The literacy leader went on study leave in 2007, so leadership of literacy was dependent on a group of leaders within each teacher team who had not necessarily been groomed for the role.

Alignment of school-wide and classroom practices in 2008

In 2008, the new principal described the potential of the schooling improvement model implicit in the LPDP. He talked about this project as a “blueprint” for “how we do things here.” He then gave examples of how teacher observation data were used, in both their numeracy and their literacy action plans, to provide evidence on where the “individual gaps” were. The numeracy leader confirmed that the classroom observation formats used with teachers for numeracy had the same procedures and communications with staff as those used in literacy. Although such structural components were being used to develop alignment across the curriculum, the numeracy leader had a view on the differences between the two projects in terms of content and pedagogy. He explained:

It seems to be that the literacy was more about . . . how the teacher is teaching, whereas the numeracy is about what the children are learning, so it’s very much about looking at the different stages, looking at the different strategies, so it’s . . . teacher knowledge of the content, rather than teacher knowledge of effective practice to the extent that the literacy [project] was, I think. (Numeracy leader, School A, 2008)

The principal described a theory for improvement that called for a deliberate plan to sustain the shifts in teacher practices:

But the problem is [that] we forgot about the sustainability part. I think there’s an assumption that because we’ve done it for three years it would stay stuck in people’s practice, but we know for a fact, and I’m sure it’s everywhere else, it hasn’t happened that way . . . . But it wasn’t deliberate enough . . . . So to me sustainability’s not a year after you finished the project . . . . It’s five, seven, it’s
just always moving, moving on, and part of that process is building sustainability into what you’ve done. (Principal, School A, 2008)

School A had just become part of a new project called Teacher Designed Schools Network. The new project’s focus was not linked to a particular curriculum learning area but was more akin to the coherence dimension of this study, developing a coherent set of effective approaches across the school. All teachers were involved in “design teams,” which were beginning to inquire into specific innovations that worked in other contexts. The principal hoped that this initiative would bring about the culture shift that was needed for what he described as a “cruiser school.” He said that any new focus would now have to tie in to what they were doing for this initiative. He reported that, later in 2008, there was to be a move in the school to shift leadership of the teacher teams from more administrative leaders to a model of pedagogical leadership. This new principal and his team had begun to think about the similarities and differences between their areas of professional learning and were being more deliberate in crafting coherence in teaching and learning across the school as a mechanism to improve performance.

At the classroom level, though, this coherence was not evident. Neither of the lesson transcripts, from the teacher new to the school or the teacher who had participated in the LPDP, illustrated substantive evidence of links being made for students to other curriculum learning, or to the bigger picture of their learning within literacy or numeracy. Students also struggled to make these links when asked. Both teachers viewed their literacy and numeracy teaching as largely separate functions during their day and could only identify surface similarities, such as learning intentions and questioning. The more deliberate moves to support alignment of effective practices across the curriculum at strategic levels in the school in 2008 had not yet impacted on teachers’ perceptions of their teaching.

Discussion

School A had sustained a strong ethic of evidence-based instruction in the three years since its involvement in the LPDP. In that time, it had also maintained similar patterns of student progress above normed populations, without being at the far end of the more systematic and rigorous co- and self-regulated inquiry continuum. The principal talked about thresholds in relation to these gains:
Maybe there’s a threshold where you can take kids, maybe there is a cognitive level where you won’t be able to take them any further, we may be miles away from that yet . . . . There’s going to reach a point where it’s going to be pretty constant and that’s up to us to build that value added on to that. (Principal, School A, 2008)

Leaders were paying greater attention to building more coherent and consistent understandings in the school about effective instruction across the curriculum, but these understandings were not yet manifested in teacher theory or practices. A threshold of evidence-based instruction appeared to be enough to establish levels of gain similar to those made earlier in the LPDP, for new cohorts of students in School A.

School C

The final case study school had presented with neither strong inquiry nor strong coherence indicators in Study 1, yet student achievement gains in writing had been sustained in the period 2006–07. Study 2 aimed to reveal any other contributing factors that may have led to these outcomes.

Patterns in student achievement

Students in School C made a net mean gain, above the mean gain calculated from asTTle norms, of 74 points in writing during the implementation phase of the LPDP (see Table 26). This shift was based on a sample of students ($n=34$) across the school and represented around an additional year’s growth above expected rates of progress for students (asTTle, version 4, manual 1.1).

Successive years of new cohorts of students entered the school at similar entry points and achieved similar levels of progress. Students in the 2007–08 cohort, though, had a slightly increased rate of progress, almost an average movement of 1 standard deviation, compared to 0.8 of a standard deviation in 2006–07. The last cohort of Year 7 students in 2008 had made a similar rate of progress to that of the 2004–05 cohort in just one year, but may still have experienced a levelling off in their second year, as Year 8 students. Fig. 12 illustrates these data for School C. School C had sustained their student

---

26 In the first cohort of the LPDP, a 10 percent random sample of students was used in each school to monitor their progress in writing, since the assessment tool, asTTle, had not been used previously in the schools. Facilitators were trained in marking these scripts.
achievement gains in writing since their participation in the LPDP according to the benchmark criteria, as well as improving on gains by 2007–08.

Table 26
*School C net gains in asTTle writing points above mean gains calculated from asTTle norms*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=34</td>
<td>n=236</td>
<td>n=268</td>
<td>n=272</td>
</tr>
<tr>
<td>T1</td>
<td>-56.97</td>
<td>-52.0</td>
<td>-62.0</td>
<td>-70.7</td>
</tr>
<tr>
<td>T2</td>
<td>-7.5</td>
<td>44.5</td>
<td>-5.23</td>
<td>56.77</td>
</tr>
<tr>
<td>T3</td>
<td>-3.4</td>
<td>-25</td>
<td>-5.23</td>
<td>3.63</td>
</tr>
<tr>
<td>T4</td>
<td>17.21</td>
<td>74.18</td>
<td>29.5</td>
<td>81.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>36</td>
<td>98</td>
</tr>
</tbody>
</table>

*Note.* 100 asTTle writing points = 1 standard deviation.

Figure 12. School C asTTle writing results 2004–08.

Figure 13 illustrates the mean difference from asTTle writing norms for the lowest 20 percent of students in School C. The at risk students who began their Year 7 in 2006 (n=47) and in 2007 (n=51) had, on average, gains twice that of all students in the school over two years.
Chapter 8 – Results and Discussion: Four cases

School C mean differences from asTTle norms showing rate of progress of lowest 20% from beginning of the year to the end of the following year:

Figure 13. School C asTTle writing results for at-risk students 2006–08.

Inquiry for improvement practices

The Study 1 data had placed School C around the middle of the 16 schools in terms of the outcomes for the inquiry dimension. Certainly the idea of monitoring target students had survived, but there was little evidence of any systematic inquiry in place in these data.

Evidence-based versus ongoing inquiry practices on 2006–07

In 2006, the literacy leaders had indicated that they also wanted to focus their efforts on those students in the high-ability range and had allowed teachers to choose their own target group, but one that was linked to their classroom data. Literacy leaders continued the LPDP practice of teacher observations and modelling with target groups of students. By 2007, this emphasis on monitoring teachers and particular students had fallen away as a systematic school-wide inquiry:

A number of our classroom teachers, when we identified target groups, identified the target group of high ability as opposed to traditionally going back to the lower ones. I don’t know that it was necessarily [an] across the board thing . . . . In terms of the actual feedback as to how the target groups actually got on, I can’t recall. (Principal, School C, 2007)

We don’t have any data on those top students, but I can remember conversations that I had. (Literacy leader, School C, 2007)
At the beginning of 2007, their asTTle results had revealed a need to focus on spelling. School C had then engaged an expert to run sessions with their teachers, and the literacy leader intended to analyse the impact in the end-of-year writing test. Reportedly, leaders set expectations that data were to be used regularly, insisting that these data should be evident in weekly planning and not just at school-wide testing times. Teachers handed on writing samples to the next teacher of their students, and leaders had rescheduled school-wide testing so that more use was made of these data from the previous year.

There were other indicators in 2007 that School C was still actively engaged in evidence-based practices around data as well as making some stronger links between these data and engaging in theory testing and further knowledge building. For example, a new practice in 2007 was to include classroom data (the averages for school-wide tests at different testing points) rather than just cohort data in the school’s annual literacy review:

This would enable each class’s progress to be tracked over the year to . . . drive its staff and school development focus and classroom programmes. (Annual Literacy Review, School C, 2007)

The recommendations in school documents were specific about what teachers needed to do with their data, but were not necessarily focused on those groups who were actually underachieving. The statements about literacy achievement compared cluster results with similar percentages at School C. School leaders had analysed what they had noticed over time in student achievement patterns and had begun to theorise about the reasons for a dip in results over the holiday break:

Student results appeared to be negatively affected by at least three factors: the Christmas Holiday Break, a new teacher and a new class level. Over the last five years of testing a “dip” of 20–30% was noted in the results gained at the end of the year testing by Year 7s and their results four months later when they were in Year 8. This means that the February results may not be an accurate picture of their achievement. (School-wide Assessment at [School C], 2007)

However, there was no indication of a response to this theory in terms of adapting practices other than by shifting dates for testing. In 2007, the Literacy Committee meeting notes posed inquiry questions and the targets were differentiated, focusing once again on those students who were not achieving:
What is our literacy target going to be? (To move as many low-achieving Māori students from the below band [Stanine 1, 2, 3] to the middle band [Stanine 4, 5, 6] or above.) How might this committee monitor this? . . . . Why are so many Year 8 students performing poorly in writing samples? (28 May, 2007)

Leadership of inquiry practices in 2008

The 2008 site visit revealed that school-wide data were being further analysed so that gender and ethnicity were also discussed as variables. In 2006, student achievement data had not been published or analysed by class but, by 2008, the linking of teacher appraisal to student outcomes was being considered. The new literacy leader in 2008 took on a mentoring role so that teachers could access her for support, including for modelling of literacy practices. She was often described as the “booster teacher” since she was able to work individually with particular groups of students who had specific needs identified in reading comprehension data, and she also worked alongside teachers in their classrooms. Originally her role was only guaranteed to last for a 10-week period, but the progress achievement data that she supplied to the school’s board of trustees had convinced them to confirm her role for the whole year.

The leadership team had a strong sense of the trends in their student achievement data in literacy and, in 2008, had already created their own versions of those shown in Figure 12. Both the principal and the deputy principal recognised that the school was maintaining similar shifts over two years for each new cohort of Year 7 students and wanted to be able to push students even further in that period. Though they described constraints, such as staff turnover and other priorities, as impacting on their capability to do this, they had begun to inquire into these patterns in more depth in 2008. Like School L, they had found that they were able to make greater shifts with their Year 7 students and noticed the levelling-off effect in Year 8. School leaders had examined classroom data to evaluate whether some teachers were making more impact than others but found that this was inconsistent. The literacy leader herself had observed that her own Year 7 students had made larger shifts in one year than the following year, despite similar practices and similar entry points. By 2008, leaders in School C were tackling the complex issues of linking teacher effectiveness to student outcomes. They had differentiated their achievement targets between year groups and compared school cluster patterns to the performance of their students.
The deputy principal admitted that the LPDP practice of teachers monitoring a group of underachieving students in their class and using these data to promote collaborative professional learning had “dropped off” in the wake of new professional learning foci. Nor were observations used as systematically as they had been in the LPDP. However, there were still strong indications of evidence-based practices being used to decide on teaching content and strategies and to test assumptions. For example, the literacy leader explained that one teacher had a “professional hunch” that her students had a listening issue and had selected an appropriate assessment to create “concrete evidence” and more specific knowledge of her students before embarking on a teaching response. This same approach was being used with the 14 percent of students, identified though their testing in numeracy, as not progressing from their reliance on manipulatives for problem solving. They had purchased new resources to trial with these students and were actively looking for other ideas and knowledge, realising that what they currently knew and did was no longer sufficient. Similarly, the literacy leader had more closely examined the STAR reading data and found that while students as a whole had achieved well in comparison to national norms, a large group of students were still in Stanines 3 and 4:

I picked up a group of kids that were just below [at Stanine 3 and 4] at the end of last year from the Year 7 grades, and I have made a big push in 8 weeks, and they are having 4 sessions a week, and I am really pushing them, and over half of them went from “at risk” into the “typical” band in March STAR testing. We have also tested another group for the commonly misspelt words and also essential word lists in Year 8, and they have also gone up dramatically. (Literacy leader, School C, 2008)

*Self-regulatory classroom practices in 2008*

At classroom level, the two teachers in the study had an evidence-based approach to their teaching. The new teacher and LPDP teacher combined their formal and informal assessment with students to decide on next steps:

I analyse a piece of writing probably twice a term, and the kids identify their next learning steps. . . . [W]e look at the piece of writing [together] once I have looked at it, and they say, “Yes, and I need to focus on that.” (LPDP teacher, School C, 2008)
Probably the biggest way [is] their draft writing books because that formative feedback all the time is a way that I can see that the children are “getting it,” and then it leads on to the next area that we are working on. (New teacher, School C, 2008)

Both teachers confirmed that they did have a target group of students but that their data were not discussed at team meetings to allow the link between individual teacher practices and student achievement to be examined more closely. These meetings were mostly about the performance of whole cohorts of students. While student needs may have been identified by these teachers, these needs were not consistently shared with students in their lessons. Most often the lesson transcripts revealed little about where the learning intentions were derived from, nor how they linked to strengths or previous learning. Both teachers had established routines associated with student self-regulation, but there was no check on whether students actually used the strategies. However, one out of three students questioned in one classroom and both students in the other classroom actually utilised the success criteria in some way to evaluate their success in the lesson or to identify a next step for their learning.

The two teachers interviewed in 2008 grasped the idea of teaching as inquiry in different ways. The new teacher described inquiry as “part of the whole searching for people that are good at what they do” when she called on the literacy leader to model effective practice. The LPDP teacher interpreted inquiry as using the students’ achievement data to plan her lessons, limiting it to the notion of evidence-based teaching, without the element of knowledge building. However, when prompted, she did add that she had searched out teacher material on listening skills when she found that her students had a specific need in this area of literacy.

*Coherence of effective instructional practices*

In the initial data collection gathered about the first dimension of the theoretical framework, there had been few significant indicators of explicit learning transfer in School C. The school immediately took on the Numeracy Development Project in 2006, and the principal had supported staff to recognise that there was always confusion and upheaval in new learning and that they would build on their knowledge and experiences of the LPDP:
So, therefore, they came to the Numeracy Development Project, I think, with a
different mindset in respect of saying, “Okay, well, we may not have a real
handle on this at the beginning, but we have confidence, as a result of the work
that has been with us previously, that we will get there.” (Principal, School C,
2007)

By 2007, though, leaders had more deliberately considered coherence of effective
instruction as a means to supporting sustainability of the LPDP.

*Transfer of learning from the LPDP in 2006-07*

When school leaders inducted their three new staff in 2007, one literacy leader
located professional learning materials from a previous focus on assessment for learning
and blended these into supports in literacy for these new teachers. Leaders had been
supported in the LPDP to observe in classrooms, and they had recognised they could
transfer their experience and the tool from LPDP to the Numeracy Development Project:

> [W]e had a lot of professional development in observations . . . those observation
sheets and the different aspects we were looking at, but pretty much those aspects
were really aspects of good teaching . . . so when we were doing our
observations, even though we had a numeracy focus, we were able to use many of
those techniques that we had learned. (Literacy Leader, School C, 2007)

Indeed the principal had espoused a strong belief that effective teachers in literacy
could be effective in all teaching areas as long as they built content knowledge in each
area. He had started 2007 with a professional learning session for teachers to reinforce
the coherence of “those big ideas and what makes a difference.” He also articulated a
keen belief about school ownership of the project and had viewed LPDP as a fixed
model. He had debated approaches with the LPDP facilitator in 2004–05 in order to align
them to the school’s culture for professional learning. He commented:

> [O]ften we had to say “We don’t want to do it that [the LPDP] way,” or
sometimes we even said, “It’s not going to happen that way.” That wasn’t
particularly easy sometimes. (Principal, School C, 2008)

The principal of School C, in contrast to the principal of School H, had been more
deliberate in crafting the LPDP model to the needs of the school.
Alignment of school-wide and classroom practices in 2008

In 2008, School C had also joined the Teacher Designed Schools Network project. Eight teams of teachers had been established across the school to inquire into areas of professional learning that they had prioritised as effective teaching and learning approaches. For example, one team had recorded their inquiry probe as “What processes and systems would need to be in place to ensure that all [School C] teachers, administrators, and students can get the rich daily feedback they need for ongoing growth and development?” Another group were inquiring into innovative and creative classrooms, examining ideas such as “buddy teaching” and sharing of knowledge. These inquiries acknowledged the alignment between leadership, teacher, and student learning and signalled that the school was aiming to build coherent systems across curriculum, classroom, and school-wide structures. School C was the only school in Study 2 explicitly to connect inquiry learning for students to how teachers might also use inquiry approaches to design questions and investigate other contexts where learning had been proved to be effective. When leaders described the new project to the researcher, they did not present a clear rationale for their choice of the project on the basis of any specific achievement data, but such a rationale may have been part of the process.

The Literacy Committee, chaired by the literacy leader, discussed the alignment of new professional learning themes to what the committee understood to be effective literacy teaching and learning. For instance, inquiry approaches for students had also been a focus in 2007 for staff, and the committee had discussed at this time whether guided reading instruction had fallen off in the wake of this new learning approach. While leaders in School C had articulated a strong belief in demanding practical ideas from any professional learning, they had also noticed when these ideas might undermine or clash with current theories-in-use.

In 2008, School C leaders talked much more about connections they were making over time as they worked in the slipstream of their previous projects. For example, the deputy principal articulated a view that was very closely aligned to the idea of professional learning projects contributing to the overall coherence of effective instruction in a school rather than being sustained as discrete projects:

27 This term is used to describe a form of team teaching whereby planning is shared with, and teaching practice is critiqued by, a colleague in the school.
I think also that in introducing the [N]umeracy [Project], we made it quite clear that we have already gone through this understanding, we have gone through the WALTs and the TIBs and the WILFs, and we understand through ATol earlier on, before we did writing, that sustained through. We talked about grouping, and that will continue through. There were going to be no major changes to that basic quality teaching that we would be wanting to see when we did numeracy. We weren't going to start from scratch and now say, “Here is a whole new set of rules” . . . It all comes back to quality teaching. . . whether you are teaching a maths lesson or a writing lesson, and so the thinking now is there has to be a purpose, there has to be grouping, you have to be able to analyse the data and know specifically where the group is at or the child is at, and your delivery must be pertaining to where the need is. . . We get the students to . . . self-evaluate. . . including peer assessment, so I think the whole quality of the teaching aspect has been. . . the thread which has tied them both together, and. . . it does overflow into all aspects. (Deputy principal, School C, 2008)

The lesson transcripts confirmed that lessons were indeed structured in a similar way across literacy and numeracy by both the new teacher and the teacher who had participated in the LPDP. All four of their lessons had established routines such as: explicit learning intentions shared with students; success criteria developed for self-regulation; think, pair, share strategies; and a review of learning towards the end of the lesson. The LPDP teacher could link these strategies to principled theory, using phrases such as “scaffolding learning” and “creating student independence” but the new teacher commented that:

I don't know the theories behind it though. So, I know how it is meant to run and what are some of the systems and structures put into place for try[ing] to cater for all of your kids. (New teacher, School C, 2008)

While these routines were common across numeracy and literacy lessons, there were fewer connections being made between learning contexts. The students who were interviewed after the lessons also identified routines that were similar, such as teacher

---

28 These are all acronyms used with students in the assessment for learning projects to support metacognition and self-regulation. WALT is “What I am learning to”, TIB is “This is because”, and WILF is “What I am looking for?”

29 Think, pair, share is a learning strategy involving students working in pairs to discuss their learning.
modelling of examples and strategies. So two years from the initial probe into elements of coherence in instructional practices, School C appeared to have a much stronger set of shared practices across the curriculum that were both articulated and in use in classrooms.

Discussion

As in School L, distributed leadership roles in School C had combined to support teachers to sustain the learning that they had gained while part of the LPDP. The new “booster teacher” role in 2008 provided opportunities for teachers to deepen their literacy knowledge and practices, while the deputy principal and principal attended to normative coherence (Coburn, 2003), broadening expectations for shifts in practice by promoting discussion of the principles of effective instruction across the curriculum. Leaders had described strong ownership of the professional learning in the school and, despite fragile understandings of inquiry being revealed in Study 1, inquiry processes appeared to be strengthened across the school in the Study 2 data. Evidence-based practices were more systematic at both school and classroom levels, and new questions were being raised about achievement in successive cohorts of students. The knowledge-building component of the inquiry cycle was more evident in the 2008 data although developing teacher goals based on data was not a systematic enterprise.

School C had consistently sustained students’ achievement gains in writing for successive cohorts of students, even slightly improving these gains over time. At risk students had been accelerated in comparison to all students in the school. These improvements occurred in the context of relatively stable levels of teacher and leadership turnover, combined with strong evidence of evidence-based practices and strategic decisions to strengthen coherence as a lever for sustainability of professional learning.

Implications for sustainability

In Study 2 it was possible to compare more closely the conditions in those schools that had sustained their student achievement gains according to the benchmark criteria for sustainability set for Study 1. Schools C and L had sustained their student achievement gains in the three years after their involvement in the LPDP according to the benchmark criteria. Both schools had also increased their gains for students in subsequent cohorts. However, School A had sustained their student achievement gains according to the benchmark criteria, maintaining but not necessarily improving these gains over time. Seven school-based conditions promoting sustainability of professional
learning and ongoing gains in student achievement emerged from the findings in Study 2: four in relation to coherence of effective classroom and three linked to inquiry practices for improvement.

**Coherence: breadth, consistency, and connectedness**

The notion of coherence in this study builds on the models presented by Coburn (2003) and Newmann et al. (2001). Breadth of programme reforms requires institutionalisation of new practices across the school, and this is usually achieved by the deliberate creation of structural, procedural, and cultural norms (Datnow, 2005). Another form of coherence has also been investigated in this study for its importance in promoting sustainability, that of connectedness of effective instruction across the curriculum. Study 2 findings indicated that the following conditions were present in Schools L and C, where student achievement gains were improved over time. School leaders crafted coherence deliberately in Schools L and C, but more in the sense of fidelity to the original LPDP literacy practices than in connecting these practices to effective instructional practices across the curriculum. While leaders in School A were paying greater attention to coherence of effective instruction across the school in 2008, these understandings had not filtered down to teacher theory or practices in explicit ways. Leaders in School H had espoused theories of practice about coherence of literacy practices across the school in 2006, but this condition was not demonstrated in subsequent site visits. School H had not sustained their student achievement gains in the years following their participation in the LPDP.

1. **Prioritising breadth and consistency**

Study 1 found that schools recognised the value of consistency of effective practices in regard to sustainability but that school leaders limited this notion to maintaining their new practices and making sure that all teachers used them in their literacy lessons with students. In Study 2, Schools L and C, in particular, regularly monitored classrooms for evidence of LPDP practices in literacy lessons. Particular practices related to breadth and consistency were found in these two schools that had sustained and improved their student achievement gains:

- Literacy leaders were released from classroom duties to mentor and observe individual staff and to support those students who were not reaching targets set by the school.
• Observations of classroom teachers in their literacy practices were continued, with opportunities for feedback on strengths and next steps.

• Literacy leaders recognised and alerted teachers to non-alignment of ideas in their professional learning and supported teachers to engage with only those ideas that were known to be effective. For instance, professional learning about inquiry learning for students appeared to be diverting teachers from the regular and focused literacy instruction that they had learned in the LPDP. Both literacy leaders acted on this threat to sustainability.

• School A also focused on consistency of effective literacy practice in the years following their participation on the LPDP. Leaders’ actions included follow up workshops for particular literacy topics and moderation meetings to discuss scoring of writing manuscripts. These actions continued to build teachers’ pedagogical content knowledge but were not as systematic as School L and C in their application. In contrast, School H did not build on the earlier signs of coherence from Study 1 and largely relied on informal checking of practices across the school, particularly after the literacy leader was awarded study leave in 2007. Sharing of ideas about effective practices became the norm in School H, rather than any opportunities to have someone observe the impact of those practices as they occurred in classrooms.

2. **Tools and routines used deliberately for sustainability**

Tools have been identified in the literature as being influential in promoting coherence of effective instructional practices (Coldren & Spillane, 2007; Robinson et al., 2009). There were specific tools and routines used to promote continued use of effective practices in each of the two schools that had sustained and improved their students’ achievement. These tools and routines scaffolded and amplified learning from the LPDP. For instance:

• School L codified the LPDP practices in a reading “handbook” for all teachers. In this way, teachers new to the school were well supported to replicate the first-hand learning experienced by those that had participated in the LPDP. Articles in the handbook helped them link theoretical perspectives to classroom strategies;

• Routinely over the year, School L engaged all teachers in discussion of the achievement targets for their class. Teachers were referred to the reading “handbook” for strategies to raise students’ achievement in particular dimensions of the STAR;
• Leaders in School C also utilised materials from previous projects to use with their new teachers in order to support their learning.

None of these specific tools or routines were described by teachers or observed in School H or School A in 2006–08.

3. **LPDP processes as a blueprint for schooling improvement**

Timperley et al. (2007) outline the learning processes for teachers that are shown to be effective in raising student achievement. These include multiple opportunities for teachers to learn and apply information, collaborative sense-making and knowledgeable expertise to challenge any existing practices that may be ineffective in raising achievement. Processes for teacher learning, introduced in the LPDP were continued in each of the case schools, with various degrees of fidelity. For instance:

• Leaders in School C and School L recognised the value of teacher observation, modelling, and feedback as means of supporting changes to classroom practices and, when these schools engaged with new learning projects, the leaders found that they challenged other forms of professional learning as a result. Schools C and L also more consistently found ways to keep these LPDP processes in place for teachers even with new learning priorities. School C, for instance, transferred the teacher observation schedule used in the LPDP to observations for numeracy, reinforcing effective instructional strategies across the curriculum;

• Observations and feedback on literacy practices in School A continued in 2006-08 but they were not consistently led by leaders who were sufficiently knowledgeable;

• School H did not take up a new focus for professional learning in 2006-07 but neither did they consistently use the schooling improvement practices in the year after their literacy leader went on study leave. The LPDP schooling improvement processes for teacher learning were therefore not embedded.

4. **Distributed leadership for breadth, consistency and connections**

School leaders in those schools that sustained and improved their gains in student achievement worked interdependently to ensure coherence of effective instruction across the school (Spillane, 2006). For example:
In School L, the deputy principal established a weekly planning routine across literacy and numeracy, while the literacy and numeracy leaders focused on supporting teachers’ literacy knowledge;

The deputy principal in School C collated and analysed school-wide data in literacy and numeracy, while the literacy and numeracy leaders supported teachers to gather and analyse data to inform their teaching programmes;

The principal of School C had utilised a staff meeting to talk about the “big ideas” of effective practice, particularly before embarking on a new professional learning project in 2006.

These findings, related to consistency of literacy practices within the school and connectedness of instructional practices across the curriculum, suggested that challenges to sustainability, such as teacher turnover and changing professional learning priorities, could be offset with attention to coherence. School L had undergone a 61 percent teacher turnover by 2008 and School C had moved immediately after the LPDP into the Numeracy Development Project, yet both schools had sustained and improved their student achievement gains.

Inquiry: different facets

The findings from Study 1 had led to a refinement of the research in Study 2 in order to check whether all facets of the inquiry cycle were being actioned. The distinction between evidence-based practices and co- and self-regulated practices was again evident in the case study schools.

5. Fidelity to evidence-based instruction a sufficient threshold for sustainability

Evidence-based decision making was observed in teacher practices across all four case study schools. This appeared more systematic, though, in Schools C, A, and L, where teachers more regularly engaged with their literacy data to inform their teaching programmes. Leaders in these schools had begun to look for trends and patterns in their achievement data. All three schools met the benchmark criteria for sustainability of gains in student achievement with new cohorts. In contrast, School H had not been as systematic with evidence-based instruction and had a particular context at work that may have resulted in reduced gains for the cohort of students that followed the LPDP. Their literacy leader went on study leave, and the school could not access external expertise to support ongoing knowledge building for teachers.
6. **Teacher mentoring deepened learning for teachers.**

Teachers need sufficient depth of pedagogical learning to be able to impact positively on student outcomes (Coburn, 2003; Elmore, 2002a; Timperley et al., 2007). All four of the schools in Study 2 opted, at different points, to fund a literacy leader to simulate the role of the LPDP facilitator after the project was finished. These literacy leaders were freed from their responsibility of regular classroom duties to coach teachers in their literacy practices. In this way, teachers could continue to learn, even if this was not in the same format as previous learning activities in the LPDP. In School H, in particular, this was not continuously supported over the three years. This appears to have impacted on their students’ achievement results.

7. **Inquiry and coherence maximised student achievement outcomes over time.**

In Schools L and C, particularly, teachers were differentiating their achievement targets, asking new questions of their data, experimenting with new ways to lift literacy achievement even further, and monitoring these initiatives for their impact on student outcomes. Schools L and C advanced beyond the evidence-based practice threshold to routinely examine teacher practices associated with students’ achievement data. They were able to improve on rates of progress even further despite high teacher attrition and/or new professional learning priorities.

Schools L and C had demonstrated various components of coherence as described above, but they had also demonstrated understandings and practices that would place them on the advanced end of the continuum for co- and self-regulated inquiry practices. Coherence, therefore, appeared to be a necessary, but not sufficient, condition for ongoing improvement sustainability. These two case study schools supported the argument, proposed in this thesis, that inquiry and coherence together may be significant dimensions for determining ongoing improvement in the years following an intensive professional learning project. In Schools L and C, strong distributed leadership models were demonstrated with both support and accountability systems in place. Both schools had moved from programme maintenance into an improvement paradigm (Century and Levy, 2002) and better exemplified the iterative inquiry and knowledge-building cycle as described by Timperley et al. (2007).

In contrast, School A’s espoused commitment to an inquiry approach in the year after their participation in the LPDP fell away over 2007–08. While teachers still referred
to student achievement data in their responses about approaches to teaching and learning, there was less talk about using that data as a means of identifying new areas of professional learning. School A managed to maintain similar rates of progress in students’ writing. In 2007, School H did not demonstrate consistent use of data-gathering and collaborative discussion about new responses to ongoing issues about underachievement. School H did not sustain the gains made while in the LPDP in students’ reading skills and knowledge with their next cohort of students.

Persistent and continuous improvement over time in student achievement is required for many students in New Zealand. For this to be a realistic goal, a shift from evidence-based approaches in schools towards more systematic inquiry and knowledge-building practices will be required. This study has therefore confirmed the “conjectures” offered by Timperley et al. (2007) and recent evidence presented by Lai et al. (2009) about sustainability being supported by conditions that promote teacher self-regulation and ongoing knowledge building. It has also provided new insights into building normative coherence by institutionalising new practices in contexts where there were conditions of change (such as teacher attrition, reduced funding, and new learning priorities) that arguably serve to limit sustainability.
Sustainability of reform has been characterised in various forms in the literature, but most often as some sort of endurance test for educators. School leaders and teachers are left on their own to weather the “tempests” of reduced funding, shifting educational priorities from new policy regimes, and the constant turnover of teachers and leaders (Wood, 2007). New classroom practices are also easily “derailed” by “turbulent” internal forces within schools, such as the micro-politics of the staffroom, or competing beliefs about effective learning models (Datnow, 2005; Spillane et al., 2002). So, there are many warnings but few guidelines for schools to follow after the implementation phase of any reform. In fact, Timperley and her colleagues’ (2007) synthesis of professional learning initiatives concluded that sustainability was simply an “article of faith” because most studies used teacher reports of continued practices as their only evidence of effectiveness. There are even fewer descriptions of how schools “bridge or buffer” these external forces (Datnow, 2005; Honig & Hatch, 2004) or jump the next hurdles without the support of external experts. Nor does the research literature provide guidance about the desirable size of quantitative measures related to progress of new cohorts of students so that schools might have some reasonable expectations to at least steer them in these feats of endurance.

There appear to be significant gaps in the evidence base at the same time as there is increasing political urgency for making inroads into the large-scale underachievement patterns detected in many educational jurisdictions (Konstantopoulos & Hedges, 2008; Raudenbush, 2005). Considerable public investment has been made in wide-ranging responses to solving disparities in outcomes for diverse populations, but without significant attention as to whether they might be sustainable over time. Therefore, research into the nature of sustainability of professional learning that promotes deep and consequential change, particularly within projects that aim to lead to ongoing improvement in students’ achievement, is both timely and important.

The question in the title of this thesis and the research questions that were examined in this study are addressed in this final chapter to offer some conclusions to the research community and some insights to educators working towards shaping sustainability of outcomes from their educational initiatives. The chapter begins with a
brief review of the methodology and results of this study and then examines the conceptual model for sustainability that underpinned this research as a contribution to the research literature, then three main conclusions about sustainability of professional learning are described. These conclusions are drawn from the findings of this research in relation to the two dimensions of the conceptual framework for sustainability. Finally, areas for future research are recommended and implications for the discourse around policy and practice are discussed.

Two interdependent studies

This research was undertaken in a purposive sample of schools in the three years after an intensive two-year literacy intervention. The LPDP had been demonstrably successful with 91 schools in the period 2004–05 with gains made for students in writing or reading comprehension. Sixteen schools in Study 1 had agreed to participate in a site visit in 2006 and were asked to submit their students’ assessment data over 2006–07 using the same focus as teachers had worked with as part of the LPDP.

Study 1 investigated the theories of action held, and practices demonstrated, by school leaders, teachers, and students related to establishing sustainability. In Study 2, four schools were selected from the sample to participate in ongoing site visits in 2007–08, providing a feedback loop in the research design and allowing an opportunity to check on previous data. Mixed methods were utilised in a case study approach in both studies so that influences of real-life contexts and participants’ beliefs and values could be revealed as part of the analyses of conditions in these schools. By reducing the numbers of schools to four, it was possible to use a “multi-focal lens” (Borko, 2004) to widen the focus from just literacy and the LPDP to the teaching and learning of numeracy as well. Each of the schools in Study 2 had presented with different combinations of the framework in the Study 1 analyses and had distinctively different mediating and outcome variables.

The patterns of students’ literacy achievement were tracked in each participant school, using either the STAR assessment or the asTTle writing data, over 2004–07 for Study 1 and over 2004–08 for Study 2. Taking into account the particular nature of each of these assessment tools and varying entry points of achievement for new groups of students, a benchmark was established to designate whether each school had sustained the rates of progress of successive cohorts of students. A second lens was applied to these
data to differentiate further between those schools that had sustained their students’ achievement gains – whether they had maintained similar rates of progress or markedly improved on their rates of progress for new groups of students. In Study 2, where schools had greater numbers of students, it was also possible to compare the patterns of progress for the lowest achievers.

Ten of the twelve schools that submitted sufficient data in the first year after the intervention sustained gains in literacy with a new cohort of students according to the benchmark criteria. In two schools, while improvement continued to be made, the rates of progress for students in 2006 did not reach the benchmark criteria. Three schools presented data for only one point of time in 2006, so comparisons could not be made. One school did not provide data. A similar pattern was observed after two years. By the end of 2007, ten of the thirteen schools that presented data sustained their gains, three schools reduced the rate of progress according to the benchmark criteria (although students had still improved beyond expected levels of progress), and three schools did not present their data. Three of the four schools in Study 2 had sustained their student achievement gains according to the benchmark criteria after two years. After three years, two of these three schools had improved on rates of progress for new cohorts of students.

These results were impressive given the nature of the mediating factors at work in each school. All but two schools had experienced some turnover of staff or leaders, seven out of 16 schools had embarked on new professional learning priorities in the year immediately following their participation in the LPDP, and none had access to the extra funding that they had utilised to release teachers for observation, modelling and feedback learning experiences. External demands also impacted on schools. The Ministry of Education had released a new national curriculum and a set of literacy learning progressions for students in Years 1–10 in the same period. So these schools had sustained their gains in the face of all of those factors that researchers warn may derail them. Most school leaders had taken deliberate actions to maintain and spread the LPDP practices in those years.

What was surprising was that the school that experienced the highest teacher turnover of the four schools had ongoing and improved achievement gains for three cohorts of students in reading comprehension. As well, the only school that took up a new professional learning focus immediately after their exit from the LPDP also improved on the student achievement gains they had made in 2004–05. The school that
had continued their focus on literacy professional learning in both 2006 and 2007 actually reduced their gains after two years, only picking up gains in student achievement when they re-entered the LPDP in 2008. Students’ progress in 2006 had outstripped the previous cohort but fell away dramatically in the second year. The loss of a literacy leader without classroom responsibilities and the external pressure of involvement in school cluster data gathering appeared to delay momentum and impact heavily on sustainability.

This research has found that sustainability can be deliberately crafted, albeit in different ways. Quoting the poet Antonio Machado, Hargreaves and Fullan (1998) explain:

There is no one answer to the question of how one brings about change in specific situations. You can get ideas, directions, insights and lines of thought but you can never know exactly how to proceed. You have to beat the path by walking it. (p. 27)

In a sense, the schools in this study have forged the path to sustainability by walking it for themselves. There is much to learn from their experiences for all those involved in educational reform.

Reframing sustainability as two dimensions

In this study, the notion of sustainability was reframed in order to offset a number of issues found in the research literature. The review of this research literature exposed sustainability of educational reform as a problematic notion on a number of fronts. Firstly, the term itself is undertheorised, with insufficient definition or shared agreement on the nature of its meaning in educational settings. Mostly, studies related to sustainability are framed in terms of the durability of a discrete set of project practices for school leaders or teachers and these are tracked as they pan out in particular sites over time (Century & Levy, 2002; Datnow, 2005; Lai et al., 2009; Miles & Huberman, 1984). Rarely have researchers looked closely at sustainability in terms of more than one intervention, or at the point of intersection as one project melds into the previous one. Indeed, this latter lens is more often used to explain how multiple interventions and ever-changing policy agendas threaten sustainability (Hargreaves & Fink, 2006; Honig & Hatch, 2004). Research on the effects of leadership practices on schooling improvement suggests that a positive relationship exists between improving coherence and improved
student achievement (Newmann et al., 2001). This body of literature has the potential to lessen the “disjuncture” between discussions of sustainability as applying to programmes or interventions and of sustainability as applying to the school or school system (Levin, 2008).

Secondly, a broad range of school-based conditions is offered in the research on sustainability. These include: leadership for improvement (Fullan, 2006; Hargreaves & Fink, 2006; Robinson, 2007; Spillane, 2006); inquiry skills (Copland, 2003; Earl & Katz, 2006; Reid, 2004; Robinson & Lai, 2006; Sutherland, 2004); sufficient depth of content and pedagogical knowledge (Bishop & O’Sullivan, 2006; Elmore, 1996, 2002a; Parr & Timperley, 2006); communities of practice that enable ongoing and shared ownership for improvement in student outcomes (Elmore, 2002a; Lima, 2001; Little, 1993); and learning processes for teachers that penetrate teacher beliefs about learning and the efficacy of the practices they engage in so that principled knowledge can be transferred to different contexts and curricula (Timperley et al., 2007). However, this substantive literature does not offer a framework with which to integrate these conditions so that they do not become a single focus or a list of actions to support sustainability of reform. The literature on sustainability typically focuses on one level for reform such as system (Datnow, 2005; Levin et al., 2008; Supovitz & Taylor, 2005) or school level, (Lai et al., 2009; Timperley & Phillips, 2003) rather than a model that impacts on thinking at all levels of the system.

Thirdly, very few studies connect these conditions to ongoing achievement of student outcomes (Timperley et al., 2007). Researchers do agree, however, that adaptations of any new practices inevitably occur (Century & Levy, 2002; Spillane et al., 2002) but, while superficial changes may exhibit the “core beliefs and values” of a particular reform, they may not necessarily secure improved outcomes for students (Timperley et al., 2003). More recent definitions of sustainability, therefore, add evaluative capability to the mix so that schools can regularly check the efficacy of altered teaching practices on student outcomes (Timperley et al., 2007).

With these issues in mind, this research reframed sustainability through a broader lens in order to strengthen examination of accountability and measurability in its design. The model that is promoted in this thesis combines a multiplicity of conditions into two overarching principles for thinking about sustainability and for framing actions at various levels of the system. The research also aimed to fill another gap in the research literature
on sustainability by presenting patterns of student progress in successive cohorts and annotating them with analysis of the two dimensions found in these sites.

Coherence of instructional practices and co- and self-regulated inquiry practices for improvement are the two key dimensions of sustainability that form the theoretical framework underpinning this research. These dimensions have been conceptualised as interdependent and dynamic continua, each melded from the myriad of school-based conditions that researchers have argued are critical for sustainable outcomes. They work in concert and can be applied to more than one level within the school or system. Recasting these conditions two-dimensionally enables connections to be made between a number of school-based factors and increases their potential to be used at multiple levels within an education system.

The first dimension of the theoretical framework is co- and self-regulated inquiry practices to nurture the depth of ideas and beliefs that educators hold to transform practices (Coburn, 2003; Elmore, 2002a). The teacher inquiry and knowledge-building cycle found on the inside front cover of *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration* (Timperley et al., 2007; see also chapter 2) is a relatively new depiction of a schooling improvement schema. It combines a number of important ideas about inquiry. Firstly, as in many other substantial works, it privileges student data as the “touchstone” (Timperley, 2003) for effectiveness of educational changes (and their sustainability). Secondly, the cycle begins at the top with student achievement data as the primary source of information to inform teaching and teacher learning. The cycle requires teachers to utilise knowledge and strategies that are already proven to be effective to support the needs identified in the student data. The cycle also indicates the link between the student data and what these demonstrate to teachers about their own knowledge gaps, about what they know and do that is effective and what they might still need to learn about and do differently. This is where depth of understanding is explored. Inquiry may also occur across learning contexts or be confined to one problem or learning area.

The study investigated each stage of this cycle in the case schools and revealed a distinction between practices that were evidence-informed and those that followed an inquiry for improvement model, where success was followed by new questions about persistent underachievement, and where innovation and next practices were monitored for their effectiveness (Hannon, 2008; Reid, 2004). Case schools were mostly stalled at
the evidence-based step in the inquiry cycle, where leaders supported teachers to use student achievement data to plan lessons but not to interrogate the data for where they may need to develop new knowledge to ensure improvement in student achievement for those students at risk of not meeting expected levels over time. Schools had mostly followed the cycle in a recursive manner, without a commitment to deepening knowledge and adapting and evaluating practices for students who had not progressed after the implementation of the first cycle.

The second dimension of the conceptual framework is coherence of instructional strategies. Sustainability is defined in this research as being broader than one professional learning initiative, encompassing the idea that each new project contributes to the overall coherence of effective instruction and to the knowledge and practices related to schooling improvement (and is not sustained solely as an entity in itself). In this way, educators have to take strategic action in the design and implementation of any professional learning that takes place in the slipstream of previous learning, making this coherence explicit to all involved so that the negative impact of “Christmas tree” innovations (Newmann et al., 2001), which may fail to impact on outcomes, are challenged. Sense-making processes for coherence require educators to search out broader principles rather than mechanistic routines or strategies and to surface and debate beliefs about teaching and learning.

Coherence of instructional practices is an aspect of educational theory that is not well illustrated in terms of how it might manifest itself in schools. More emphasis is placed on why coherence may be needed for effective reforms and what it does not look like in schools or in school districts (Adelman & Taylor, 2003; Honig & Hatch, 2004; Newmann et al., 2001). This study teased out the different manifestations of coherence in the contexts under study in order to establish their importance in determining sustainability of change. Two characteristics of coherence were found in those schools that did improve on their impact on students’ literacy achievement over time, and both required school leaders to take strategic and explicit actions. The first characteristic was the crafting of coherence within the domain of literacy itself: a form of breadth and scale but across individuals in one school site. The second form of coherence, observed in far fewer schools in this study, was establishing connections across learning domains in order to amplify and embed what are known to be effective classroom pedagogies. These connections may be identified in beliefs about learners or as teaching and learning.
pedagogies: how students and teachers learn effectively and how teachers and leaders guide that learning so it impacts on students’ outcomes. In either form, coherence involves multiple layers in the school and supports shifts in ownership of change and institutionalisation of practices (Coburn, 2003; Elmore, 2002a).

Figure 14 illustrates these two dimensions working together in a school context. This conceptual framework contributes to the current literature on sustainability by broadening a definition of sustainability to embrace multiple professional initiatives and by focusing attention on the relationship between breadth and depth of reforms.

![Figure 14. Sustainability of educational improvement as two dimensions.](image)

The framework can be used at multiple levels for considering policy and strategy for sustainability reforms aimed at improved outcomes for students. The next section draws out the key conclusions about sustainability of reform in relation to the case schools in the study and this conceptual framework.

**Conclusions**

Century & Levy (2002) described three stages of development of sustainability. Implementation of science programmes in their research was followed by a maintenance and then improvement phase, where teacher practices evolved and adapted practices based on the core principles of the original reform. This notion was taken up to support the analyses of the qualitative and quantitative data from the case study schools. Would schools move from a maintenance to improvement paradigm and what would this look like in terms of student outcomes? The three main conclusions of this research that

204
follow link the idea of maintenance and improvement to the two dimensions of the conceptual model.

_Schools hold self-limiting perceptions of sustainability_

This research examined the spread and ownership of practices within school sites as new strategies were practised and adapted and when new teachers joined the school in the wake of the LPDP. Data analysed and synthesised across 16 schools in Study 1 indicated that most principals and school leaders did not have a macro-view of sustainability. Although they described indicators of sustainability operating at both organisational and classroom level, their lens was predominantly on literacy, confined to one instructional domain. The school leaders and teachers in Study 1 mostly conceptualised sustainability as about the maintenance of those practices and processes learned from their participation in the LPDP. Leaders were typically seeking consistency of literacy practices, breadth and ownership of reform, across their school rather than continuing to deepen teachers’ knowledge about literacy. There was little sense of a theory for improvement and/or for sustainability emerging from these schools. Talk was practice based, although often with some reference to achievement data. No responses were specific about how much improvement they actually expected now that they were no longer supported by external facilitators.

Coherence of effective instructional practices across curriculum areas did exist in these schools but was recognised as, and limited to, a core set of formative assessment practices. These practices included explicit instruction (such as the use of learning intentions derived from student needs and the development of success criteria), focused feedback, and students’ awareness of their learning strategies. Formative assessment was not necessarily articulated as a core set of practices that would support sustainability, but participants made such links between curriculum learning projects when specifically asked. Few participants in Study 1 recognised or actioned transfer of principled knowledge about schooling improvement to other curriculum contexts. In other words, most schools in the study did not reinforce ideas such as inquiry as a process for schooling improvement, or the use self-regulation of the effectiveness of teaching strategies across professional learning projects, nor did it appear that those who were leading new professional learning in the school explicitly supported leaders and teachers to make connections across particular projects.
Together these findings have some significance for sustainability of educational reform. A self-limiting theory of action for sustainability – limited to one area of the curriculum and to maintaining the newly developed practices – may increase the fragility of the student achievement gains over time and lessen the opportunity to amplify and institutionalise effective practices across the curriculum.

_Evidence-based practices led to maintenance of outcomes for students_

Inquiry as a model for schooling improvement was not yet fully embedded in the case schools but, for the most part, evidence-based approaches were. Study 1 indicated that there were gaps in schools’ understandings of inquiry despite it being a central belief and value of the LPDP.

Data from Study 1 indicated significant differences between the LPDP inquiry model (Bareta & English, 2006) and what was understood and evidenced in these schools. A distinction between evidence-informed practices and a fully fledged inquiry model emerged as the study progressed. Schools sat at different stages of the inquiry continuum in Study 1, with most having paid much more attention to evidence-informed decision making than to co-and self-regulated practices for improvement. The inquiry process was used in a recursive manner in most instances, with few new inquiries into underachievement being carried out after one cycle. Schools were relying on what they had learned previously to impact on those students who were still not making adequate progress. Therefore, most schools appeared to be working in a maintenance regime rather than an improvement paradigm in the period after their implementation of the LPDP.

However, schools in the study survived the endurance test, developing resilience to competing demands and institutionalising the changes that they made to leadership and teacher practices as part of their participation in the LPDP. Ten of the 13 schools had sustained their student achievement gains according to the benchmarks developed for the study. The “lowest common denominator” for those schools that did sustain student achievement gains in subsequent years was a central core of coherent ideas around what were largely formative assessment strategies. For example, participants described similarities such as sharing specific learning intentions in each lesson, building success criteria with students, and using data to plan lessons that matched their students’ needs.

Where these practices were more systematically applied, they appeared to be a sufficient threshold to support sustained outcomes for their students. Thompson and
Wiliam (2007) also found that formative assessment strategies, without specific attention to deepening content knowledge, yielded marked improvements in student achievement. They added that further research was needed to determine whether these boosts in achievement might plateau and whether any further gains in student achievement may be achieved subsequently with a focus on strengthening teachers’ pedagogical content knowledge. This study confirms the latter idea.

*Inquiry and coherence practices led to a pattern of ongoing improvement for student outcomes*

The practices around inquiry described in Study 1 were again observed in Study 2. Two of the four schools were engaged in rounds of data analysis that did not necessarily deepen their pedagogical content knowledge. They were not asking new questions of their student achievement data to refocus on persistent underachievement, instead applying their current and newly acquired practices to the next cohort of students. These schools were not necessarily systematic in linking their student achievement data to teacher practices. Schools L and C, however, had more deliberately pursued coherence of effective instructional practices and progressed from evidence-based practices to using a more systematic inquiry approach as they moved into 2008. Their student achievement data also reflected improved gains for students in the three years after the LPDP. Persistent low achievement was certainly monitored, and new ways to solve this problem were theorised and enacted. Teachers understood the expectations for improvement. Both schools moved beyond a maintenance regime to an improvement paradigm that more clearly represented the inquiry and knowledge-building model discussed earlier.

In Study 2, the two schools that improved on their student achievement gains best represented the drive for breadth and consistency of practices within one domain, as described in the findings from Study 1. They were deliberate in their strengthening of school-wide coherence in literacy programmes, providing teachers and parents with tools and processes that codified and preserved fidelity to the LPDP practices, routinely checking their occurrence with observations, “walk-throughs,” and requirements for planning.

In 2008, leaders in both schools expressed a view that each professional learning project was indeed contributing to the broader corpus of knowledge about effective change processes and effective classroom teaching. This led them to a macro-view of sustainability. One of these schools, in particular, more deliberately crafted coherence in
this way. The principal led discussion with teachers on the links they were making across professional learning projects. They had participated in an assessment-for-learning initiative, the LPDP, the Numeracy Development Project, and then inquiry learning approaches for students, all in quick succession. They wanted to “consolidate” and chose to look for a process that would help them take a “big picture” view of effectiveness. In 2008, the school engaged in a revisioning process that involved learning about and prioritising factors known to be effective in galvanising staff and raising student performance. Conversely, Schools A and H appear to have stalled at a maintenance phase and to have failed to develop an ability to evolve and adapt further (Century & Levy, 2002).

**Directions for further research**

This study has a number of limitations which have been outlined earlier, but it has raised new possibilities that will require new and different studies. Hostetler (2005) reminds researchers of the problem of just ending with answers. Certainly further research is required to identify just how convergent the “core beliefs and values” are across professional learning projects in New Zealand and other countries, particularly their understandings about inquiry and knowledge building for improved student outcomes.

Codifying practices have their own dangers, however, when practitioners do not check on their effectiveness or do not regularly look to new practices that may have more impact on persistent underachievement (Lai et al., 2009; Timperley & Phillips, 2003). The emphasis on normative coherence as a lever for sustainability is certainly promising in the literature (Coburn, 2003; Datnow, 2005; Levin, 2008; Newmann et al., 2001). That is, teachers and schools are more likely to sustain reform when new district policies are aligned with the next set of initiatives that they take on (Coburn, 2003). However, the literature lacks narrative of how this is manifested effectively in schools and how it might be combined with an “inquiry habit of mind” (Earl & Katz, 2006).

There were only glimpses of substantive understanding of inquiry as an iterative and systematic approach. Inquiry is heralded as a key determinant for improvement at school and system levels (Reid, 2004; Lai et al., 2009; Timperley et al., 2007), but the results of this study indicate that much more evaluative capability is required in schools for them to fully gain its benefits as a process for ongoing knowledge building. Shaffer
and Serlin (2004) argue that studies can build on each other to provide increasingly nuanced understanding of complex notions. There is further research needed to identify which particular skills and knowledge about evaluative capability are not well supported in the current environments for schools in New Zealand and internationally. Annan (2007) argues that sustainability of school reform would have a “better chance” if all those involved in schooling improvement agreed on and got buy-in for a set of evidence-informed standardised practices and then worked to ensure the integrity of their implementation.

Lastly, the question raised by Thompson and Wiliam (2007) is relevant in terms of this study and directions for new research. What is the relationship of depth of teacher content knowledge and formative assessment practices in relation to supporting continuous improvement in student achievement gains?

**Implications for policy and practice**

Typically, participants confused inquiry learning for students and inquiry as a teacher action for professional learning. Certainly most participants in the study had rarely experienced any deliberate conversations about this or any other such connections as they moved from project to project. Coherence appeared necessary, but not sufficient, as a school-based condition for sustainability. Where inquiry and coherence were strongly evidenced across school and classroom practices, with leaders working in concert, sustained outcomes were maximised. That is, student achievement gains for new cohorts of students were improved when deliberate actions were taken to deepen content and pedagogical knowledge, when outcomes of any fresh approaches to problems of underachievement were carefully monitored, and when coherent effective teaching routines were established across the school.

This study has defined and described coherence more explicitly in a school context, extending responsibility for alignment beyond school leadership to all those in the system, including professional development leaders, teachers, and students. More empirical research across learning domains is needed to establish the importance of this condition for sustainability and therefore for policy priority. Inquiry and coherence working in concert appeared to shift the participating schools from a maintenance to an improvement paradigm as they worked to sustain professional learning initiatives.
These findings afford a platform for dialogue with other researchers in this field who are concerned with crafting sustainability of improved outcomes for students; and they may help educators to compare and contrast their experiences and provide a useful first bridge across reform initiatives in New Zealand and internationally. The conceptual model for sustainability promoted in this thesis and its implications for design of professional learning in and across schools provides the starting point for these conversations and for further critique in new empirical studies about sustainability.
Appendix A

EXAMPLE OF PARTICIPANT INFORMATION AND CONSENT FORMS

Faculty of Education
Telephone: 373 7599 ext 88998

Participant Information Sheet

Research project title: The Literacy Professional Development Project
Principal Investigators: Associate Professor Judy Parr, Professor Helen Timperley, Pam O’Connell

For: Principal, literacy leaders, numeracy leader and teachers

This research is being carried out by Pam O’Connell, PhD candidate from the Faculty of Education, University of Auckland, under the guidance of Professor Helen Timperley and Associate Professor Judy Parr. The University of Auckland have been contracted to research the implementation of the professional development in the Literacy Project, which is funded by the Ministry of Education. You are invited to participate in the third phase of this study. You are under no obligation to do so but we would appreciate your input. We are particularly interested in exploring the conditions associated with sustained benefit from the professional development, including the nature of in-school conditions that are associated with enhanced student literacy achievement. We invite you to document and discuss your experiences in 2008 in the area of literacy and numeracy. This involves:

- An audiotaped interview with the principal focusing on the role they have taken in supporting sustainability (30 mins);
- An audiotaped interview with the principal and literacy leader(s) about their writing data in 2006 – 2007, the staffing changes that have occurred over 2005 – 2008 (45 mins);
- An audiotaped interview with the literacy leader and numeracy leader about approaches to professional learning (30 mins);
- An audiotaped literacy and numeracy lesson by one teacher who participated in the LPDP and one teacher new to the school, followed by interviews with 3 students in each of the two classes and a post lesson interview with the teachers (a morning in each classroom);
- Collection of documents related to teacher goal setting and professional learning;
- Student achievement data for 2008 in the area you focussed on (writing or reading comprehension) using asTTle or STAR and your interpretation of the trends in these data. Note that these data are in aggregated form for each year group and will not contain identifying information like names.

The source of the information you provide is confidential to the researchers and facilitators. Neither you nor your school will be identified in reports. The audiotapes will be transcribed by a typist who does not know you and you will be identified by a number only. Individual consent to participate to being interviewed or observed and to have that interview or observation recorded is also required from all participants. You can ask that the tape recorder be turned off at any time. You are able to review interview transcripts, make changes and ask that the tapes be returned. You may withdraw the school from participation without giving a reason, or withdraw your data at any time prior to the
completion of data collection by 31 December, 2008. Originals of your responses to the questionnaires and the transcriptions from the tapes are kept securely and shredded at the conclusion of the project.

The final report on this study will be made available to you so that you can examine the overall findings about how schools worked to sustain the gains made in the Literacy Professional Development Project.

The researchers would be pleased to give more information about the project on request. If you have any questions to ask or concerns you wish to discuss please contact the appropriate person from the following:

Pam O’Connell
Professional Development Manager
Learning Media Ltd
P O Box 3293
Wellington
Tel: 04-471 5538
Email: pam.oconnell@learningmedia.co.nz

Associate Professor Judy Parr and Professor Helen Timperley (Principal investigators)
Faculty of Education
University of Auckland
Private Bag 92019 Auckland
Tel: 09-623 8899 ext 88998 and 87401.
Email jm.parr@auckland.ac.nz  Email h.timperley@auckland.ac.nz

Frances Langdon
Head, School of Teaching, Learning & Development
Faculty of Education
University of Auckland
Private Bag 92601, Symonds St, Auckland
Tel: 09-623 8899

The Chairperson
University of Auckland Human Participants Ethics Committee
Private Bag 92019
Auckland
Tel: 09-3737599 ext 87830

Approved by the University of Auckland Human Participants Ethics Committee
On 13/05/04 for a period of six years until 13/05/10 (Reference no 2004/059).
Consent Form

THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

Research project title: The Literacy Professional Development Project

Principal Investigators: Associate Professor Judy Parr and Professor Helen Timperley, Pam O’Connell

For: Principals, literacy leaders, teachers

This project has been explained to me and I have had the opportunity to ask questions and have them answered to my satisfaction. I give my consent for the project to be conducted in the school. I understand that I can withdraw from this research project at any time up to the end of data collection by 31 December, 2008.

I also understand that the source of the data obtained is confidential to the researchers and that the name of the school or staff will not be used in any report without express written consent. I understand that participants are able to review and edit transcripts and to request the return of audiotapes and that the transcriptions are kept securely on University premises and shredded at the conclusion of the project.

Name: ____________________________
Position: __________________________
School: ____________________________
Signature __________________________
Date: ____________________________

Consent to audiotape interview, lesson or meeting
I agree / do not agree to the audiotaping of interviews / classroom lesson / meeting (delete where not applicable)
Signature __________________________
Date: ____________________________

I wish/ do not wish to review transcripts of interviews.
I request/ do not request that the original audiotapes be returned (delete where not not applicable)

Approved by the University of Auckland Human Participants Ethics Committee
On 13/05/04 for a period of six years until 13/05/10 (Reference no 2004/059).
### APPENDIX B

#### COHERENCE AND INQUIRY SCORES FOR 16 SCHOOLS, 2006

<table>
<thead>
<tr>
<th>Question</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Principal Questionnaire</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. No new teachers</td>
<td></td>
<td></td>
<td>C</td>
<td>I</td>
<td>C</td>
<td>I</td>
<td>C</td>
<td>I</td>
</tr>
<tr>
<td>2. Nos still remain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Leader turnover</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Leader turnover</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Leaders still remain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6b.7 Reasons for continuing or shifting PD focus</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>8. Decision to move onto PL focus</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>9.a,b,c Funding/resource</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>10/11. Principal Descriptors of sustainability</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Leadership Interview</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Leaders views of key messages LPDP</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2. Leaders views of indicators of sustainability</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3. Sig different changes?</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4. Leaders views of connections</td>
<td>0.00</td>
<td>0.00</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5. PL focus for Lit</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>6. purpose for data gathering</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>7. Data gathering process</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>8. How data analysed</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>9. How classroom t’s work with data</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>10. Think differently about data</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>11. Observations</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>12. profess readings</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Question</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>13. PD structures useful li</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>14. transfer of processes</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>15. leaders goals</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>16. induction</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>17. difficult practice</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>18. difficult teachers</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>19. Lit plan</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>20. Eval effectiveness</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>21. Beliefs</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>22. s/w expectations</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>23. extra comments</td>
<td>215</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Teacher Questionnaire**

<table>
<thead>
<tr>
<th>Question</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Teachers view of key messages of LPDP</td>
<td>0.91</td>
<td>0.27</td>
<td>1.00</td>
<td>0.00</td>
<td>1.06</td>
<td>0.47</td>
<td>1.00</td>
<td>0.83</td>
</tr>
<tr>
<td>2. Looks like in classroom</td>
<td>0.82</td>
<td>0.09</td>
<td>1.00</td>
<td>0.00</td>
<td>1.12</td>
<td>0.18</td>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td>3. changes in practices 06</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.53</td>
<td>0.00</td>
<td>1.75</td>
</tr>
<tr>
<td>4. Expectations</td>
<td>0.00</td>
<td>1.27</td>
<td>0.00</td>
<td>3.00</td>
<td>0.00</td>
<td>1.88</td>
<td>0.17</td>
<td>1.50</td>
</tr>
<tr>
<td>5. Judging effectiveness of tcg</td>
<td>0.18</td>
<td>1.36</td>
<td>0.00</td>
<td>3.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>2.50</td>
</tr>
<tr>
<td>6. purpose for data gathering</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.53</td>
<td>0.00</td>
<td>1.75</td>
</tr>
<tr>
<td>7. Purpose of s/w data</td>
<td>0.00</td>
<td>0.82</td>
<td>0.00</td>
<td>3.00</td>
<td>0.00</td>
<td>1.88</td>
<td>0.17</td>
<td>1.50</td>
</tr>
<tr>
<td>8. Lit focus for ownPL</td>
<td>0.00</td>
<td>1.27</td>
<td>0.00</td>
<td>3.00</td>
<td>0.00</td>
<td>1.88</td>
<td>0.17</td>
<td>1.50</td>
</tr>
<tr>
<td>9. Beliefs</td>
<td>0.00</td>
<td>1.27</td>
<td>0.00</td>
<td>3.00</td>
<td>0.00</td>
<td>1.88</td>
<td>0.17</td>
<td>1.50</td>
</tr>
</tbody>
</table>

**Interviews LL and NT**

<table>
<thead>
<tr>
<th>Question</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Decision about effectiveness</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2. Planning</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>2.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3. Purpose for class data</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>4. Purpose for s/w data</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5. Key messages LPDP</td>
<td>0.00</td>
<td>3.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>6. LL and NT views on connections</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>2.00</td>
</tr>
<tr>
<td>7. a/b/c/d. Goals for PL</td>
<td>0.00</td>
<td>5.00</td>
<td>0.00</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Question</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>8. Beliefs</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>2.00</td>
</tr>
<tr>
<td>9. Expectations</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>School Document</td>
<td>3.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>School Meeting</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Classroom lesson</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Student interviews</td>
<td>0.00</td>
<td>1.44</td>
<td>0.00</td>
<td>0.22</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.33</td>
</tr>
<tr>
<td>Graph Co-ordinates</td>
<td>5.91</td>
<td>38.43</td>
<td>6.00</td>
<td>29.34</td>
<td>8.18</td>
<td>17.28</td>
<td>6.50</td>
<td>10.16</td>
</tr>
<tr>
<td>Question</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Principal Questionnaire</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. No new teachers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Nos still remain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Leader turnover</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Leader turnover</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Leaders still remain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6b.7 Reasons for continuing or shifting PD focus</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>8. Decision to move onto PL focus</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>9.a,b,c Funding/resource</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>10/11. Principal Descriptors of sustainability</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Leadership Interview</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Leaders views of key messages LPDP</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2. Leaders views of indicators of sustainability</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>3. Sig different changes?</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4. Leaders views of connections</td>
<td>0.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5. PL focus for Lit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. purpose for data gathering</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>7. Data gathering process</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. How data analysed</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>9. How classroom t's work with data</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>10. Think differently about data</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>11. Observations</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>12. profess readings</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>13. PD structures useful li</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>14. transfer of processes</td>
<td>1.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>15. leaders goals</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>16. induction</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Question</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>17. difficult practice</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>18. difficult teachers</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>19. Lit plan</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>20. Eval effectiveness</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>21. Beliefs</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>22. s/w expectations</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>23. extra comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Teacher Questionnaire**

<table>
<thead>
<tr>
<th>1. Teachers view of key messages of LPDP</th>
<th>1.13</th>
<th>0.75</th>
<th>0.50</th>
<th>0.00</th>
<th>0.75</th>
<th>0.17</th>
<th>1.20</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. looks like in classroom</td>
<td>1.25</td>
<td>0.13</td>
<td>1.00</td>
<td>0.00</td>
<td>0.75</td>
<td>0.00</td>
<td>1.30</td>
<td>0.30</td>
</tr>
<tr>
<td>3. changes in practices 06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Expectations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Judging effectiveness of tcg</td>
<td>0.00</td>
<td>2.13</td>
<td>0.00</td>
<td>0.00</td>
<td>0.17</td>
<td>2.00</td>
<td>0.60</td>
<td>3.20</td>
</tr>
<tr>
<td>6. purpose for data gathering</td>
<td>0.00</td>
<td>2.38</td>
<td>0.00</td>
<td>2.50</td>
<td>0.00</td>
<td>1.33</td>
<td>0.00</td>
<td>2.10</td>
</tr>
<tr>
<td>7. Purpose of s/w data</td>
<td>0.25</td>
<td>2.25</td>
<td>0.00</td>
<td>1.50</td>
<td>0.00</td>
<td>2.50</td>
<td>0.00</td>
<td>3.00</td>
</tr>
<tr>
<td>8. Lit focus for ownPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Beliefs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Interviews LL and NT**

| 1. Decision about effectiveness                                     | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 1.00 | 2.00 | 0.00 |
| 2. Planning                                                          | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| 3. Purpose for class data                                           | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| 4. Purpose for s/w data                                             | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 5. Key messages LPDP                                                | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| 6. LL and NT views on connections                                   | 0.00 | 2.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 |
| 7. a/b/c/d. goals for PL                                             | 1.00 | 0.00 | 1.00 | 1.00 | 0.00 | 2.00 | 0.00 | 2.00 |
| 8. Beliefs                                                           | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| 9. Expectations                                                     | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 |

**School Document**

| 1.00 | 1.00 | 1.00 | 0.00 | 2.00 | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |

**School Meeting**

| 2.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 2.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

**Classroom lesson**

<p>| 0.00 | 2.00 | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 2.00 | 1.00 |</p>
<table>
<thead>
<tr>
<th>Question</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student interviews</td>
<td>C</td>
<td>I</td>
<td>C</td>
<td>I</td>
<td>C</td>
<td>I</td>
<td>C</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.99</td>
<td>0.00</td>
<td>0.00</td>
<td>0.66</td>
<td>0.00</td>
<td>0.33</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.11</td>
<td>0.00</td>
<td>0.11</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graph Co-ordinates</td>
<td>8.63</td>
<td>22.63</td>
<td>11.50</td>
<td>14.00</td>
<td>9.67</td>
<td>18.66</td>
<td>14.10</td>
<td>26.93</td>
</tr>
<tr>
<td></td>
<td>4.00</td>
<td>17.50</td>
<td>7.73</td>
<td>17.17</td>
<td>7.60</td>
<td>11.31</td>
<td>7.30</td>
<td>16.05</td>
</tr>
</tbody>
</table>
APPENDIX C
RATES OF PROGRESS FOR STUDENTS’ LITERACY ACHIEVEMENT IN 13 SCHOOLS 2004-07

Figure 3. School L achievement data in reading with successive cohorts of students, 2004–07.

Figure 4. School P achievement data in reading with successive cohorts of students, 2004–07.
Figure 15. School A achievement data in writing with successive cohorts of students, 2004–07.

Figure 16. School B achievement data in reading with successive cohorts of students, 2004–07.
Figure 17. School C achievement data in writing with successive cohorts of students, 2004–07.

Figure 18. School D achievement data in reading with successive cohorts of students, 2004–07.
Figure 19. School E achievement data in writing with successive cohorts of students, 2004–07.

Figure 20. School G achievement data in reading with successive cohorts of students, 2004–07.
Appendix C

**Figure 21.** School H achievement data in reading with successive cohorts of students, 2004–07.

**Figure 22.** School I achievement data in writing with successive cohorts of students, 2004–07.
Appendix C

School J average stanine level for two cohorts showing progress over two years of schooling.

Figure 23. School J achievement data in reading with successive cohorts of students, 2004–07.

School K average stanine level for two cohorts showing progress over one to two years of schooling.

Figure 24. School K achievement data in reading with successive cohorts of students, 2004–07.
Figure 25. School N achievement data in reading with successive cohorts of students, 2004–07.


