



## 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](#)

Teacher Appraisal: Missed Opportunities for Learning

Claire Esther Louise Sinnema

A thesis submitted in partial fulfilment of the requirements  
for the degree of Doctor of Education,  
The University of Auckland, 2005

**The University of Auckland  
Thesis Consent Form**

This thesis may be consulted for the purpose of research or private study provided that due acknowledgement is made where appropriate and that the author's permission is obtained before any material from the thesis is published.

I agree that the University of Auckland Library may make a copy of this thesis for supply to the collection of another prescribed library on request from that Library; and

1. I agree that this thesis may be photocopied for supply to any person in accordance with the provisions of Section 56 of the Copyright Act 1994

Or

~~2. This thesis may not be photocopied other than to supply a copy for the collection of another prescribed library~~

(Strike out 1 or 2)

Signed:.....

Date:.....

## ABSTRACT

The improvement of teaching is now recognised to be a key to the sustained improvement of student achievement. Teacher appraisal, a central element of performance management, is intended to improve the quality of teaching (and therefore learning). This thesis examines whether teacher appraisal is achieving this purpose, through three empirical studies.

This thesis begins with a discussion of the context of teacher appraisal in relation to education reform of the last two decades. A normative model of teacher appraisal, one that focuses on data-based inquiry into student learning, is outlined, and the emerging education context in relation to that model is examined. Three studies, which explored schools' appraisal documents, appraisal discussions and teachers' appraisal goals, are presented.

The first study, a document analysis, investigated the emphasis on student learning in the statements of purpose in 17 primary schools' appraisal policies, and the emphasis on student learning in the performance indicators developed to assess teachers against the professional standards. Results showed that while improved student learning was an intended purpose of schools' appraisal policies, the indicators that were used to evaluate teachers seldom focused directly on student learning.

The second study examined whether data-based inquiry into student learning occurs in appraisal discussions, and the reasons for the level of reported inquiry. Eleven practitioners (four appraisers and seven appraisees from three schools) were interviewed about their most recent appraisal discussion to investigate whether student learning was a focus of the appraisal discussions and whether student achievement data were referred to in those discussions. Findings showed that appraisal discussions typically focused on teacher practices without exploring connections between those practices and the impact they have on student learning.

Only one of the eleven teachers described a discussion that included talk specifically about student learning. Furthermore, there were no reports of reference to student learning data in appraisal discussions. There was, however, a particularly positive response to the suggestion that appraisal should, and could, focus on student learning and on data relating to this learning. This study also highlighted the considerable influence of appraisal goals in determining the content and scope of appraisals.

In the third study, 68 teachers from eight schools responded to a questionnaire about their appraisal goals to establish the extent to which teachers' appraisal goals focus on data-based inquiry into student learning. Less than five per cent of goals were found to refer to student outcomes. The vast majority focused on teacher processes and behaviours. Goals also tended to be vague, rather than specific, and were not explicitly challenging.

Findings across the studies suggest that only limited attention is given, in critical elements of teacher appraisal, to student learning. The final chapter of the thesis explores two key strategies for closing the gap between current appraisal practices and the 'appraisal for learning' approach argued for here – an approach that focuses on teacher learning about student learning and that is based on evidence. The first strategy involves alignment within appraisal elements, and to initiatives beyond appraisal. The second strategy is capacity building to ensure that personal, interpersonal and organisational capacities necessary for 'appraisal for learning' are developed. This research shows how teacher appraisal policy and practice has been a missed opportunity to improve teaching effectiveness and how it could be reshaped in ways that maximise student learning.

## DEDICATION

A number of people deserve acknowledgement and thanks for their contribution to this research. I am particularly grateful to the teachers and principals who agreed, so willingly, to participate in the studies, and for the generosity with which they gave their time, and shared their experiences and ideas. My two supervisors, Professor Viviane Robinson and Doctor Judy Parr, deserve enormous thanks for their advice, support and encouragement. I have plenty of evidence to support a claim that their teaching helped me to learn. Gratitude also goes to the people in schools and educational organisations I have worked in during the course of my studies, for the assistance they have given me in pursuing my doctoral work. I am grateful, also, to professional colleagues and friends who have given support in many ways – reading material, engaging in dialogue, sharing ideas, and celebrating the milestones in the course of this EdD.

A most important acknowledgement goes to the team at home, since this work would simply not have been feasible without them. Bayard and Mum have been superb in their effort and energy in looking after two special young children. They have made my many ‘thesis writing escapes’ possible. Charlotte and George deserve special mention, too, for their remarkable understanding of what it means for Mummy to be “doing her thesis” and their willingness to draw on the reverse of so many discarded drafts! To the rest of my family and friends, those who have helped in various ways, I am sincerely grateful.

## TABLE OF CONTENTS

<b>LIST OF TABLES .....</b>	<b>VII</b>
<b>LIST OF FIGURES .....</b>	<b>VIII</b>
<b>CHAPTER 1 .....</b>	<b>1</b>
TEACHER APPRAISAL – AN ALTERNATIVE TO LARGE-SCALE EDUCATION REFORM AS A MEANS TO IMPROVING TEACHING AND LEARNING .....	1
<i>Education Reform as a Response to Achievement Issues</i> .....	1
The Nature of Education Reforms of the 1980s and 1990s .....	2
New Zealand’s Education Reform .....	3
The Effect of Education Reform on Performance Management in Schools .....	4
<i>The Persistence of Achievement Issues Despite Education Reforms</i> .....	6
<i>Time for Attention to the Source of Variation – Teaching</i> .....	7
Teacher Appraisal as a Site for Making a Difference .....	8
Recent shifts in teacher appraisal and evaluation .....	9
<b>CHAPTER 2 .....</b>	<b>12</b>
DATA-BASED INQUIRY INTO STUDENT LEARNING IN TEACHER APPRAISAL .....	12
<i>Part 1: The Case for Data-based inquiry into student learning in teacher appraisal</i> .....	12
Appraisal as a Site for Inquiry .....	12
Student Learning as the Focus for Appraisal Inquiry .....	14
Educational, professional and ethical arguments for focusing on student learning in teacher appraisal .....	14
The Imperative for Student Learning Data to Inform Appraisal Inquiry .....	18
Possible Objections to Data-Based Inquiry into Student Learning in Appraisal .....	23
Apprehensions about accountability .....	23
Concern about the recognition of the multiple and complex influences on student learning .....	23
Perceptions of data-based inquiry into student learning as a threatening process .....	26
Inquiry Cultures, Critical Dialogue and Learning Communities .....	27
<i>Part 2: The Professional Context of Appraisal in New Zealand Schools</i> .....	29
The Existing Appraisal Policy Context .....	29
Appraisal process elements .....	29
The focus of appraisal .....	32
Principles to underpin appraisal .....	36
The Promise of the Emerging Professional Context and National Policy .....	38
National Education Priorities .....	38
Schools’ Planning and Reporting Policy .....	39
Iterative Best Evidence Synthesis Programme (BES) .....	40
National Assessment Strategy .....	41
<i>Part 3: Methodology</i> .....	45
<i>Appraisal in Practice: The Studies Outlined</i> .....	45
<i>A Mixed Methods Approach</i> .....	46
Pragmatic Knowledge Claims .....	47
Causal relations and teacher appraisal .....	48
The Research Design: The Strategy and Methods – Integrating Qualitative and Quantitative Data .....	49
<b>CHAPTER 3 .....</b>	<b>57</b>
THE EMPHASIS ON STUDENT LEARNING IN APPRAISAL DOCUMENTATION .....	57
<i>Introduction</i> .....	57
School Appraisal Policies .....	59
School Indicators for the Interim Professional Standards .....	59
<i>Method</i> .....	61

The Research Setting .....	61
Procedures .....	62
Data Sources .....	63
Analysis - Coding of Policy Documents .....	65
Inter-rater reliability .....	67
Analysis - Coding of Performance Indicators.....	68
Inter-rater reliability .....	70
<i>Results</i> .....	71
References to Student Learning in Policy Intention Statements.....	71
Performance Indicators About Student Learning .....	73
Lack of Alignment Between Intentions and Professional Standards' Indicators .....	76
<i>Discussion</i> .....	78
'Best of Intentions' in School Policies .....	78
Indicators Reflecting Assumptions about Impacts Practices Have on Learning.....	79
Excessive Indicators as a Cause of Policy Implementation Slippage .....	80
Lack of Alignment Between the Policies and the Indicators .....	83
Bureaucratic and Professional Approaches to Appraisal .....	85
Indicators and Bureaucratic Versus Professional Approaches to Appraisal.....	87
Threat-Free Indicators.....	88
The Potential for Student Learning Focused Indicators .....	91
Organisational culture.....	91
Professional standards' indicators and school culture .....	92
Culture of inquiry.....	94
Indicators as a potential positive influence on school culture.....	95
Limitations of the Study .....	96
<i>Conclusion</i> .....	97
<b>CHAPTER 4</b> .....	<b>99</b>
APPRAISAL DISCUSSIONS – WHAT IS INQUIRED INTO DURING TEACHER APPRAISAL? .....	99
<i>Introduction</i> .....	99
<i>Method</i> .....	100
The Research Setting .....	100
Data Sources .....	102
Procedures .....	103
Analysis - Coding of Interview Transcripts .....	106
Inter-rater reliability .....	112
<i>Results</i> .....	112
Appraisal Discussion Topics.....	112
Appraisal Discussion Characteristics.....	116
Use Of Data About Student Learning .....	117
Summative Versus Formative Appraisal Discussions .....	117
Response to the Idea of Focusing On Student Learning during an Appraisal Discussion .....	118
Teachers' receptiveness - "What a good idea!" .....	118
Teachers' receptiveness to examining data for appraisal discussion purposes .....	120
Provisos.....	121
A solitary negative response .....	122
<i>Discussion</i> .....	123
Explaining the Lack of Focus on Student Learning in Teacher Appraisal Discussions .....	123
Undemanding appraisal goals.....	124
Assumed connections between practices and student learning outcomes .....	124
Perceptions of appraisal as an ineffectual process.....	125
Document templates as barriers to student learning focused appraisal discussions .....	126
Student learning discussed in contexts other than appraisal .....	127
Alignment.....	128
Perceptions of appraisal discussions as a summative phenomenon.....	128
Appraisal is Threatening .....	130



Student learning as ‘undiscussable’ .....	131
Discussing the ‘discussable’ and avoiding threat.....	131
An absence of data about student learning in appraisal discussions.....	134
The ‘discussable’ – positive messages in appraisal discussions.....	134
A Proposed Approach for Teacher Appraisal: Data-Based Inquiry into Student Learning.....	136
The predominant positive response.....	136
Examining the lone voice of disapproval.....	138
Conclusion.....	140
<b>CHAPTER 5 .....</b>	<b>141</b>
APPRAISAL GOALS – WHAT GOALS ARE USED IN TEACHER APPRAISAL? .....	141
Introduction.....	141
Method.....	144
Procedures .....	145
Data Sources .....	146
Analysis of Appraisal Goals .....	147
Inter-rater reliability .....	152
Results .....	153
Appraisal Goals’ Connections to Data-Based Inquiry into Student Learning .....	153
Student outcome focused teacher appraisal goals.....	154
Student process and teacher process focused goals .....	155
Inquiry .....	155
Data.....	155
Goal specificity and challenge.....	156
School Level Variable.....	157
Orientations of Appraisal Goals.....	158
Learning Areas and Information Communication Technology in Teacher Appraisal Goals.....	160
Discussion .....	161
Lack of Focus on Student Learning .....	162
Lack of Specificity and Challenge .....	163
Possible Explanations.....	166
Superficial Compliance.....	167
Norms of professional autonomy and supportive collegiality.....	167
The Risks .....	169
Conclusion .....	170
<b>CHAPTER 6 .....</b>	<b>172</b>
APPRAISAL FOR LEARNING.....	172
<i>The contribution of this research</i> .....	173
Implications for further research.....	174
Implications for practice.....	175
<i>Policy Issues</i> .....	176
Closing the Gap between Current Practice and the Potential for ‘Appraisal for Learning’ .....	176
Closing the Gap Through Alignment Within Appraisal .....	177
Closing the Gap Through Alignment Beyond Appraisal .....	178
Aligning to school planning and reporting policy and practices .....	180
Aligning to the national assessment strategy .....	181
Aligning to effective professional development approaches .....	183
Closing the Gap through Building Capacity for ‘Appraisal for Learning’ .....	183
Implications for leadership.....	186
<b>APPENDICES .....</b>	<b>188</b>
<b>REFERENCES .....</b>	<b>195</b>

## LIST OF TABLES

Table 1 Reflective Teaching Traditions .....	17
Table 2 Aspects of Teachers' Performance Which Should be Appraised.....	33
Table 3 Principles That Should Underpin Appraisal Policies and Processes .....	37
Table 4 Assessment Information and the National Assessment Strategy .....	44
Table 5 Characteristics of Participating Schools.....	62
Table 6 Policy Intention Statement Coding Rules.....	67
Table 7 Performance Indicators Coding Rules.....	70
Table 8 School Appraisal Policy Intention Statements .....	72
Table 9 Percentage of Performance Indicators by Category.....	75
Table 10 Cultural Norms and Assumptions Reflected in the Findings about Indicators.....	93
Table 11 Characteristics of Participating Schools.....	101
Table 12 Interview Participants' Appraisal Roles .....	102
Table 13 Appraisal Discussion Topics Coding Rules .....	109
Table 14 Appraisal Discussion Characteristics Coding Rules.....	110
Table 15 Types of Responses to the Idea of Data-Based Inquiry into Student Learning in Teacher Appraisal .....	111
Table 16 Appraisal Observation Template Headings.....	126
Table 17 Defensive and Productive Reasoning.....	133
Table 18 Characteristics of Questionnaire Respondents .....	146
Table 19 Coding Rules - Appraisal Goals in Relation to Data-based Inquiry into Student Learning.....	148
Table 20 Coding Rules - Student Outcome Goal Specificity and Challenge .....	149
Table 21 Coding Rules - Appraisal Goals and Their Focus on ICT and Learning Areas .....	151
Table 22 Coding Rules - Orientation of Appraisal Goals.....	152
Table 23 Percentage of Appraisal Goals Focused On Student Learning and Based On Data .....	154
Table 24 Specificity of References to Learners, Learning and Extent of Improvement in Student Outcome Focused Teacher Appraisal Goals.....	156
Table 25 Specificity and Challenge in Student Outcome Focused Teacher Appraisal Goals .....	157
Table 26 Number of Student Outcome Focused Appraisal Goals by School .....	158
Table 27 Orientation of Teachers' Appraisal Goals.....	158
Table 28 Appraisal Goals Focus on Curriculum Areas and ICT.....	160
Table 29 Connecting Teacher Appraisal to the Outcomes Focus in School Planning and Reporting Policy .....	181
Table 30 Integrating the Appraisal Context into the Assessment Strategy Framework.....	182

## LIST OF FIGURES

Figure 1 Elements of a Typical Appraisal Cycle.....	8
Figure 2 Performance Management in New Zealand Schools.....	31
Figure 3 Causality in Teacher Appraisal.....	48
Figure 4 Causality between Teacher Appraisal Elements .....	49
Figure 5 Deductive and Inductive Codes .....	51
Figure 6 The Mixed Methods Approach.....	53
Figure 7 Research Design Type - Multistrand Conversion Mixed Model .....	55
Figure 8 Appraisal Policy and Practice Alignment.....	58
Figure 9 Types of Intention Statements in Schools' Appraisal Policies.....	73
Figure 10 Total Number of Professional Standards' Indicators.....	74
Figure 11 Comparison of Percentages of Intention Statements and Professional Standards' Indicators Relating to Student Learning .....	77
Figure 12 Example of a Typical Appraisal cycle.....	80
Figure 13 Number of Minutes per Indicator Based on Four Hours of Discussion Time.....	81
Figure 14 Levels of Appraisal Decision Making.....	85
Figure 15 Teachers' Own Inquiry into Student Learning .....	96
Figure 16 Appraiser's Inquiry into Appraisee's Inquiry into Student Learning .....	96
Figure 17 Appraiser and Appraisee Joint Inquiry into Student Learning, and into Each Other's Inquiry ...	96
Figure 18 Themes for Teacher Appraisal Discussion Data Analysis .....	107
Figure 19 Teacher Appraisal Discussion Topics.....	115
Figure 20 Response to Suggestion of Data-based Inquiry into Student Learning in Teacher Appraisal Discussions.....	118
Figure 21 Percentages of Goals Focused on Learning Areas and ICT .....	161
Figure 22 Appraisal Goals' Minimising Positive Impact on Student and Teacher Learning .....	169
Figure 23 Appraisal Goals' Maximising Potential to Impact on Teacher and Student Learning.....	169
Figure 24 Appraisal for Learning Alignment Model.....	178
Figure 25 Aligning 'Appraisal for Learning' to Other Initiatives .....	179

## CHAPTER 1

### **Teacher Appraisal – An Alternative to Large-Scale Education Reform as a Means to Improving Teaching and Learning**

#### *Education Reform as a Response to Achievement Issues*

There has been substantial education reform over the last twenty years in education systems all over the world (Hopkins, 2001; Hopkins & Levin, 2000; Levin, 1998; Muijs, Harris, Chapman, Stoll, & Russ, 2004; Novlan, 1998). Like other countries which have undergone education reform, New Zealand's reform efforts were driven by the pursuit of economic stability and social well-being (Fancy, 2004). In the late 1980s there was considerable doubt regarding the education sector's capacity to address the issues that had brought New Zealand to a state of financial crisis. A "Task Force to Review Educational Administration" was established in 1987, and its report ("The Picot Report" (Picot, 1988)) was critical of the administration and operation of New Zealand schools. It claimed a lack of responsiveness to learners and the learning context of the time, declining achievement standards, increasing cost and complexity of management systems, and a lack of accountability for educational outcomes (Novlan, 1998). The Picot Report emphasised that the time had come for radical change. Most of the suggestions for educational reform in the report were included in the subsequent White Paper "Tomorrow's Schools: The reformation of educational administration in New Zealand" (Lange, 1988), and were reflected in subsequent education legislation ("The Education Act," 1989).

In other countries, too, attention turned during the 1980s to reviewing and reforming education systems. Education systems in England, Scotland, Canada, Australia and the United States were all, like New Zealand, the subject of substantial reform at around this time. The rationale, typically, was to improve educational outcomes and the capacity of the system to contribute to wider societal improvements. In Australia, for example, a number of reviews of education similar to those in New Zealand took

place. In New South Wales, for instance, a number of reports were commissioned in the late 1980s to review aspects of the education system (Riordan & Weller, 2000): “The Report of the Committee of Review of NSW Schools (The Carrick Report)” (New South Wales Government, 1989a), “School-Centred Education – Building a More Responsive State-School System (The Scott Review)” (New South Wales Government, 1989b)) and a White Paper entitled “Excellence and Equity – New South Wales Curriculum Reform” (New South Wales Ministry of Education and Youth Affairs, 1989). These reports highlighted the need to improve education and had an important role in shaping subsequent legislation – The Education Reform Act (1990) (New South Wales Parliament, 1990).

In the United States pressure for education reform intensified following the publication of two key reports – “A nation prepared: Teachers for the 21st century” (Carnegie Forum on Education and the Economy, 1986) and “A nation at risk: The imperative for educational reform” (National Commission for Excellence in Education, 1983). Wang, Haertael, & Walberg (1993) explain the impact of these reports:

.... the public became aware of the mediocre performance of the nation’s schools and the inadequacies of the teaching corps. The poor performance of the United States’ students, when compared to their international counterparts.... generated concern among business and civic leaders that the nation’s graduates would be ill-equipped to deal with the demands of a global economy (Wang et al., 1993).

Consequently, the first federally initiated programme of reform in the United States, “America 2000”, was adopted and implemented across 50 states.

#### *The Nature of Education Reforms of the 1980s and 1990s*

While the detail of initiatives varies across the many countries that have undergone education reform, a number of elements are, according to Hopkins and Levin (2000), common across nations. Reform is typically focused on curriculum, accountability, governance, market forces and the status of teachers, as Hopkins and Levin outline:

Curriculum: governments have instituted more restrictive curriculum requirements including increased emphasis on science, technology, and so-called basic skills such as literacy. Traditional subject divisions have been reinforced in many cases.

Accountability: governments have increased testing of students and have made the results public, and in some cases put in place extensive external inspection of schools.

Governance: while governments have centralized curriculum and assessment they have also decentralized many decisions from intermediate bodies such as school districts or local authorities to individual schools, and have given parents an increased role in school governance, all of which has put new pressures on professional staff

Market forces: governments have tried to introduce market elements to schooling through increasing the opportunity, or requirement, for parents to choose schools (or, in some cases, for schools to choose parents and students).

Status of teachers: in a number of countries the status of teachers and their organisations has been attacked directly through unilateral changes by governments to the status of unions or to collective bargaining arrangements. (Hopkins & Levin, 2000, p. 18-19)

#### *New Zealand's Education Reform*

These elements of reform (above) were all evident, to varying degrees, in New Zealand's education reform in the 1980s and 1990s. The nature of that reform in New Zealand is summarised briefly here:

Curriculum was reformed through the introduction of a new curriculum framework in 1993 (Ministry of Education, 1993). New curriculum statements were also introduced in all of the essential learning areas – Mathematics, English, Technology, Science, Social Studies, Health and Physical Education and The Arts (Ministry of Education, 1992, 1994, 1995, 1997b, 1997c, 1999, 2000).

Accountability measures, while not focused on the testing of students and public displays of results (as was the case in other countries), were increased in 1989 when 'Tomorrow's Schools' (Lange, 1988) was implemented. The Education Review Office

was established “as an independent government organisation to ensure effectiveness, equal opportunity for employment and education, financial management and property aspects. Schools would be appraised on the basis of meeting their chartered objectives” (Novlan, 1998, p. 12). More recently, increased accountability for schools has been introduced through a schools’ planning and reporting policy (Ministry of Education, 2002b, 2003b, 2003d, 2003e, 2003f, 2003g). This policy, referred to as ‘planning for better student outcomes’ requires School Boards to set priorities for student outcomes, specify how target relating to these priorities will be achieved and evaluate and report against them.

Market forces have, as a result of the reforms, been increasingly influential. In 1990 a policy of school choice was introduced in which funding followed students to their chosen school, and schools competed with each other for students, since enrolment numbers determined, largely, the allocation of funding to a school (Gordon, 2003). The needs of schools for goods and services have also, increasingly, been met through privatised support services, resource providers, publishers, advisors, and researchers from private organisations – further evidence of market forces at play.

The New Zealand education reforms brought a fundamental change to the governance of schools. Like in other countries, this involved decentralising decision making to individual schools, and increasing the role of parents and communities in running schools. Tomorrow’s Schools brought about the restructuring of the government Education Department and placed responsibility for all aspects of governance (including curriculum, personnel, financial and property matters) with Boards of Trustees. The personnel aspect of this responsibility was broadened in 1997 to encompass the appraisal of teachers – the context for the research reported in this thesis.

*The Effect of Education Reform on Performance Management in Schools*

The requirement that Boards of Trustees develop and implement performance management policies, and conduct performance appraisal of their teachers (Ministry

of Education, 1997a, 1998a, 1998b, 1998c), was a delayed but direct consequence of legislation in New Zealand in the late 1980s ("The Education Act," 1989; Lange, 1988; "The State Sector Act," 1988). That legislation requires Boards to operate personnel policies that comply with good employer principles and that ensures teachers provide an education that fully meets their students' needs. These requirements will be discussed more fully in the following chapters. In summary, Boards are required to implement policies and practices in regard to teachers' performance expectations, performance appraisal, reward systems, professional development and, where necessary, disciplinary/competency procedures. This move to focus central elements of education reform, evaluation and accountability on teachers was not unique to New Zealand. Similar moves occurred in many western countries that sought to address achievement issues, in part, by improving teaching and learning through performance management schemes. Teacher appraisal, or teacher evaluation as it is also called, was a key feature of those schemes.

In Australia, the eighties and nineties saw various appraisal schemes devised and implemented. According to Kleinhenz and Ingvarson, though, few of these made a difference to classroom practice, even though they were aimed at improving teaching (2004). Rather, "the privacy of practice prevailed, giving teachers (unearned) autonomy, but isolating them from the professional development lifeblood of formative assessment and feedback" (Kleinhenz & Ingvarson, 2004, p. 34). Similarly, appraisal schemes in the nineties across the United Kingdom sought to address poor performance in international educational achievement surveys. The schemes were, though, "poorly conceptualized, variably implemented and based upon individual teachers choosing the area of their practice that they wanted their appraisal to be on in order to facilitate development" (Reynolds, Muijs, & Treharne, 2003, p. 93).

In the United States, likewise, teacher evaluation became "a center piece of educational accountability and reform... in the 1980s" (Ellett & Teddlie, 2003, p. 106). This led to teacher evaluation systems for the purpose of licensure, based on state-mandated policies for assessing teachers on-the-job. Between 1982 and 1983, twenty



states put in place their first requirements for teachers' performance evaluation (Furtwengler, 1995). A fundamental flaw of those teacher evaluation systems, according to Ellett, was their predominant focus on teacher behaviour and teacher performance as opposed to a focus on student learning (Ellett, 1997).

### *The Persistence of Achievement Issues Despite Education Reforms*

Despite the large-scale reform efforts and performance management initiatives that have taken place in the last two decades across many nations, low levels of student achievement, and wide achievement disparity persists. According to Levin and Wiens, "the results [of reform over the last 20 years] in terms of increased student achievement have been generally disappointing" (2003, p. 659). Hopkins and Levin described the situation as one of breathtaking irony, since dramatic increases in education reform, intended to improve outcomes for students, have actually had insufficient impact on student achievement levels (2000). Some years earlier, Wang et al. had already suggested that "although some of the new programs and practices of the 1980s were supported by research findings, few of these innovations have shown replicable long-term impact on students" (Wang et al., 1993, p. 251). Reform and improvement models were, as Schmoker claims, "hugely popular, but patently discredited.... The record is clear that these failed" (Schmoker, 2004, p. 424).

They failed too in New Zealand, it could be argued, in terms of generating achievement gains and a high achievement, high equity education system. The expected student achievement gains, keenly anticipated following education reforms, have not been realised. New Zealand educators have, more recently, turned their attention to the growing evidence of disparities in student achievement.

The failure of education reform to impact positively on New Zealand students' achievement is evident in findings from the Programme for International Student Assessment (PISA) (Ministry of Education, 2001b). The PISA study revealed that

New Zealand has one of the widest gaps, of any of the 32 participating nations, between its highest achieving students and its lowest achievers. The Ministry of Education states, “while our system succeeds with most children, it struggles to succeed with our growing diversity of students. Children who are Maori, Pasifika, have special education needs, disabilities or who are gifted do not, on average, do as well as they could” (Ministry of Education, 2003a). Similar concerns were revealed in the Progress in International Reading Literacy Study (PIRLS). The range of scores for New Zealand students was wider than the range for most of the 35 other countries that participated in the study. Closing this achievement gap is one of the major challenges facing educators today, and demands a focus on raising achievement and reducing disparity (Ministry of Education, 2004b).

In order for student achievement gains to occur, educators’ attention needs to shift toward factors that are known to make a difference for learners. There is substantial research evidence to show that a major source of variation in student achievement lies in classroom teaching, rather than school level variables. Classroom teaching and learning should, then, be the focus of efforts to improve student achievement.

### *Time for Attention to the Source of Variation – Teaching*

There has been increasing evidence in recent years that points to the significance of the impact teachers have on student learning, as opposed to impacts other education system elements have (Creemers, 1994; Hattie, 1999b; Ministry of Education, 2004a; Reynolds et al., 2003; Rowan, Correnti, & Miller, 2002). Student achievement attributed to between-school variation, according to Reynolds et al. (2003), is dwarfed by achievement attributed to within-school variation. Greater variation in student performance between teachers/classes than between schools is also reported by the Ministry of Education in the document “Impact of Teaching and Schools on Variance in Outcomes” (2004a). These findings make clear the fundamental role teaching plays in influencing student learning and supports the view of Hopkins and Levin (2000) that education reforms have focused on the wrong variables – variables too

distant from students’ classroom learning experiences. Educators now have strong evidence to support understanding about the clear link between quality teaching variables and student learning, and as Kleinhenz and Ingvarson explain, “it is a short step from this understanding to recognising the importance of teacher evaluation in schools” (2004, p. 1).

*Teacher Appraisal as a Site for Making a Difference*

Appraisal is a process through which Boards of Trustees ensure that teachers are meeting their teaching, school-wide and management responsibilities (Ministry of Education, 1997a). A key purpose of appraisal is to improve teaching and learning, through a series of appraisal elements. A typical appraisal cycle, based on the mandatory requirements prescribed by the Secretary for Education, is outlined in Figure 1.

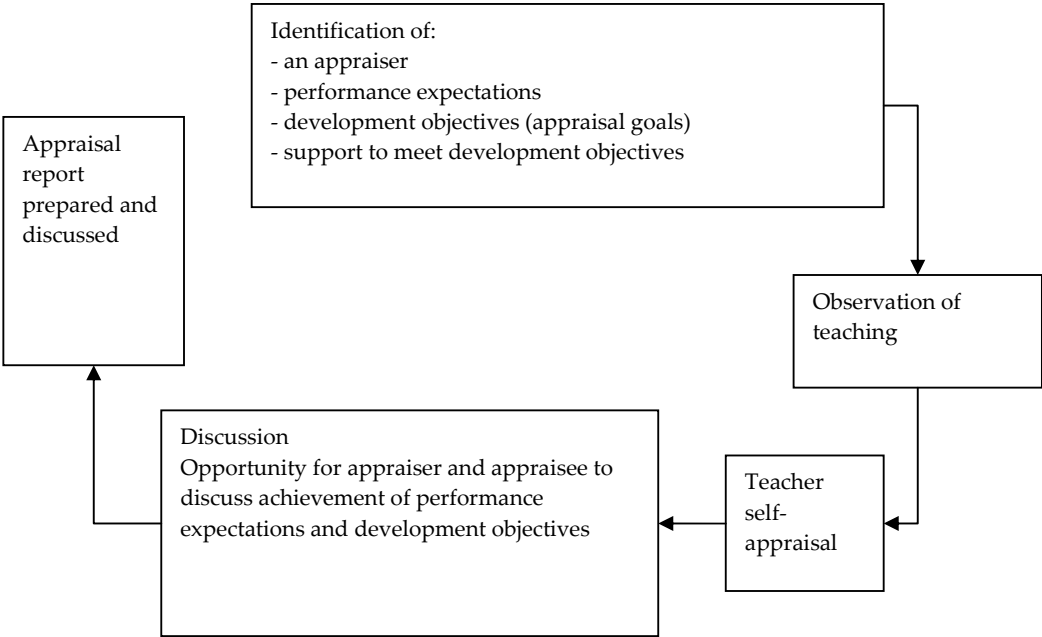


Figure 1 Elements of a Typical Appraisal Cycle

These elements of teacher appraisal, if carried out in a particular way, arguably have greater potential to improve teachers’ practice than large-scale system-level reforms.

Teachers are potentially quite likely to think about and improve their practice if they have a development objective meeting with their senior teacher, an appraiser observing in their classroom or a discussion after school with their appraiser. They are less likely to think about and improve their practice in a significant way in response to a new large-scale reform initiative that is being introduced across the country. Since teaching makes more difference to student achievement than school level initiatives, opportunities to focus on teaching and, most importantly, the impact of teaching on learners, should be utilised to the fullest. Appraisal is an ideal context for that focus.

There has, though, been wide-spread dissatisfaction with traditional appraisal (or teacher evaluation) approaches, and claims that appraisal has had limited impact on teaching and learning (Davis, Ellett, & Annunziata, 2002; Ellett & Teddlie, 2003; Gunter, 2002; Kleinhenz & Ingvarson, 2004). In response to that criticism, appraisal systems around the world have undergone changes and revisions over the last decade.

*Recent shifts in teacher appraisal and evaluation*

Appraisal policy changes in the United Kingdom have included linking appraisal to policies aimed at raising pupil achievement (Jennings & Lomas, 2003) through requirements that appraisals include “classroom observation; an assessment of the results achieved by pupils in a teacher’s care; and, an annual performance review linked to targets for enhanced pupil performance” (Department for Education and Employment, 1998, p. 49). These moves have been perceived as ‘harder-edged’ and more rigorous since they shift appraisal from a process centred on professional development, to one centred on individual teacher accountability (Bennett, 1999) and have an increased focus on student learning. This increased accountability is particularly evident in the performance-related pay aspect of performance management introduced in England, in 2000, and Wales, in 2001. Teachers with fewer than six years’ experience are paid under an automatic increment system, and beyond those years have to apply to cross a pay ‘threshold’. This application involves “a portfolio of evidence put together by the individual teacher and involving

analysis of pupils' results, data drawn from classroom observation, commitment to professional standards and its impact on classroom performance" (Farrell & Morris, 2004). The literature points clearly to the implementation of this system being perceived as problematic, but there is evidence of a mixed response to the principle of performance-related pay and its potential to improve teaching and learning (Bennett, 1999; Farrell & Morris, 2004; Haynes, Wragg, Wragg, & Chamberlin, 2003; Jennings & Lomas, 2003).

In a study of teacher evaluation policies and practices in Australia, Ingvarson and Kleinharz (Australian Council for Educational Research, 2004 ) found little coherence or consistency across the states. Approaches ranged from "brief and perfunctory generic criteria... to much more elaborately developed standards" (p. 7). That research also concluded that while most teacher evaluation policies and practices in Australia are at an "embryonic stage of development" there is a "burgeoning interest in the development of professional teaching standards and their application in teacher evaluation and professional learning".

In the United States most teacher evaluation systems are "designed to measure minimally essential teaching skills, with little concern for student learning" (Ellett & Teddlie, 2003). The reliance on particular teacher behaviours, practices and method in frameworks for teacher evaluation is not surprising, given the prolific period of process-product research that preceded the first state-wide systems of teacher evaluation in the United States in the 1980s. The extensive body of literature drawn from process-product research in education had established products or outcomes of those behaviours, practices and methods that, for a time, provided a rationale for focusing primarily on those teacher processes.

There is, though, some evidence in the United States of moves towards teacher evaluation systems that are designed to focus not just on teaching, but on connections between teaching and learning. Ellett and Teddlie outline examples of this – the Professional Assessment and Comprehensive Evaluation System (PACES) in the

Miami-Dade County Public Schools system in Florida and the National Board of Professional Teaching Standards (NBPTS). PACES is a learner-centred process designed to improve teaching and learning. The approach represents what Ellett and Teddlie refer to as “a new generation of learner-centered, classroom based evaluation systems” (2003, p. 108).

It seems, then, from the UK, Australian and US experience, that there is a trend for teacher evaluation and appraisal to become increasingly rigorous, learner-centred, and professional learning oriented. In New Zealand, however, there have been few changes in appraisal policy and, consequently, probably few changes in appraisal practice, since the 1990s. The following chapter explores a possibility for a ‘new-generation’ of appraisal in New Zealand – appraisal that inquires into, and focuses on, student learning and is based on evidence.

## CHAPTER 2

### **Data-Based Inquiry into Student Learning in Teacher Appraisal**

In the initial section of this chapter I develop an argument about why teacher appraisal should focus on data-based inquiry into student learning. There are three parts to this argument. Firstly, I explain why appraisal can and should be an inquiry process. Secondly, I contend that since the fundamental purpose of teaching is to make an impact on students, those impacts should be the focus of inquiry during appraisal. There are professional and ethical grounds for expecting this of appraisers and appraisees. Thirdly, I argue that inquiry should be data-based, since suppositions about the impacts of a teacher's practice are not sufficient for productive inquiry. Despite the range and complexity of factors that influence student learning, and despite issues related to the possibility of teachers feeling threatened, data about student learning should be at the heart of inquiry during teacher appraisal. In the second part of this chapter I consider how the current policy context of appraisal in New Zealand primary schools enables data-based inquiry into student learning. Finally, in part three of the chapter, I outline the methodology for a series of studies that evaluated the extent to which appraisal does focus on data-based inquiry into student learning, and explain associated issues and constraints.

#### ***Part 1: The Case for Data-based inquiry into student learning in teacher appraisal***

##### *Appraisal as a Site for Inquiry*

There is little that is contentious in proposing that inquiry should be a strong element of teacher appraisal. Teachers and school managers are familiar with the arguments put forward by proponents of reflective teaching and inquiry. The arguments are

convincing and it is difficult to challenge the emphasis on inquiry advocated by a range of writers.

Stoll, Fink, and Earl (2003) see inquiry as a tool for enhancing a school's capacity for learning (Stoll, Fink, & Earl, p. 49). When teachers inquire, they learn, and not always the things they expected or wanted to know. Stoll, Fink, & Earl emphasize the need to reserve judgment and be open to new ideas, to gather the necessary information, and to question and challenge beliefs and perceptions. Senge (2000) also claims a connection between schools that learn and schools that use inquiry, which he defines as "thoughtful, reflective and informed deliberation about one's practice" (p. 285). In Cardno's "Triple I Approach" (2001) inquiry is used to check the thinking of others once information has been given and reasoning and evaluations illustrated. It is an important means of promoting productive rather than defensive reasoning. Cardno says that inquiry questions should seek information and check and test one's own and others' views. They should not be controlling of others' responses, and should be used to check assumptions about facts and emotional responses. The argument of this thesis builds on an already established acceptance of the benefits of inquiry.

What is perhaps more provocative than advocating inquiry generally is the suggestion that inquiry be used in the appraisal component of performance management. While inquiry is utilised in a number of schools, it is not typically associated with appraisal practices. The principles of inquiry mentioned above – principles of learning, openness to new ideas, challenging beliefs, deliberation and productive reasoning – would be useful to the appraisal process. The required elements of observation, self-appraisal and interviews (Ministry of Education, 1997a, 1998c) each provide an ideal opportunity for reflection and inquiry. There is the potential, then, to undertake inquiry not only informally as part of teachers' regular activities but also more formally during teacher appraisal.



*Student Learning as the Focus for Appraisal Inquiry*

Appraisal that simply includes inquiry is, however, not sufficient. What the inquiry focuses on also needs to be considered. To be productive and valuable, the focus should be on that which is most important and fundamental to the teacher appraisal process. I propose that it is the impact of teaching on students' learning that is most important and fundamental. It is that impact that should be the concern of at least a substantial portion of the appraisal inquiry. Darling-Hammond suggests that the promotion of such inquiry is the key to real improvement in teaching. Truly professional assessment enables teachers to "inquire productively into the effects of their teaching on student learning, thus developing the commitment and the disposition to evaluate themselves in light of how well they can help students succeed" (Darling-Hammond, 1998, p. 472).

The Ministry of Education's guide "Teacher Performance Management" includes a reference that suggests teachers should "reflect on [their] own teaching approaches and techniques and take action to improve" (Ministry of Education, 1998c). This falls short of the type of inquiry proposed above. I would posit that the reflection should go further than considering approaches and techniques, and should include evidence-based reflection on the impact of those approaches and techniques on the achievement and learning of students.

*Educational, professional and ethical arguments for focusing on student learning in teacher appraisal*

The case for student learning focused inquiry in teacher appraisal is compelling for three main reasons. The first reason is that the promotion of student learning is at the heart of all education. Each profession has a telos, or overriding purpose, that the members of that profession must be committed to (Soltis, 1986). In education, that commitment is to a purpose of promoting learning. This purpose should, therefore, be the focus of school processes, including appraisal. Shulman describes this as the pedagogical imperative which requires teachers to examine the impact of their work on students and the consequences (Shulman, 2002). Clark's ideas about the primacy of education and the need to make children's education a foundation of accounts of teaching also support this view (Clark, 2001).

The professional nature of teaching and the rationale this provides for appraising teachers is well established (Grootenboer, 2000), and is the second reason to encourage student learning focused inquiry in appraisal. There are two particular expectations of teachers as professionals that relate to the student learning focus argued for here. Firstly, there is an expectation about the way they perceive the learners they teach. It is assumed that students and their learning will be a teacher's foremost concern. Indeed, client centeredness (Hoyle, 1982) and altruism (A. Hargreaves & Goodson, 1996) are defining characteristics of a profession. Secondly, there is an expectation that teachers themselves learn. Having a commitment to learning and to improvement of practice are also characteristics of professionals (Hoyle & John, 1995; Sachs, 1997; Shulman, 1998; Sockett, 1993; Soder, 1990; Stoll et al., 2003). Shulman (1998) explains that one of the attributes of a profession is the need to learn from experience as theory and practice interact, and Stoll et al. (2003) describe a new form of professionalism which requires of teachers an inquiry-oriented approach to their work. Similarly, Sachs claims that "continuous learning and the improvement of our practice should be at the core of teacher professionalism" (Sachs, 1997, p. 268).

Inquiry, reflection and learning, though, are not sufficient. It is the nature of the inquiry, reflection and learning that is most important. Combined, the two professional obligations of teacher learning and concern for students compel educators to take a professional approach to appraisal by focusing on the impact of their teaching on students.

There are ethical considerations, too, that suggest it would be unacceptable to examine the quality of teaching without taking into account the impact that teaching has had on students. These provide a third argument for the idea that student learning should be a focus of inquiry in teacher appraisal. We have a compulsory education system, and parents are legally required to send their children to school. Requiring parents to 'surrender' (Soder, 1990) their children to our system is justified

by an expectation that their learning is at the core of that system. This view is aligned with the ethical principle of beneficence, one of four principles outlined by Lovat, in which professionals are expected to do that which is good for their client (Lovat, 1997). Parents, and society as a whole, can expect not only that the obvious ethical requirements of keeping children safe, and free from physical and mental harm, are upheld, but also that their learning is the ultimate focus of the education system that serves them. Shapiro & Stefkovich describe this as 'the ethic of the profession' – a fourth paradigm beyond the ethics of justice, critique and care (Shapiro & Stefkovich, 2001). The draft 'Ethics Framework for New Zealand Registered Teachers' includes an obligation that teachers will strive to 'effect learning for the children, and other learners they teach' (Hall, 2003, appendix D).

This framework, however, falls short of making explicit that inquiry into that effect is necessary, and equally as important as the effect itself. An equivalent set of draft ethical principles from the Australian state of New South Wales does achieve this, stating that teachers ought to critique the impact that their teaching and professional values have on students (Brock, 1998, 'A Proposed Code of Ethical Principles' section). This thesis suggests that this ethical principle be applied to schools' appraisal policies and practices, so that teachers are ethically obliged to make inquiry into the learning of students an essential element of appraisal.

The suggestion of student learning focused inquiry demands that teachers reflect on matters beyond those emphasized in the traditions of reflective teaching outlined by Zeichner and Liston (Zeichner & Liston, 1996 p. 55-61). These traditions of reflective teaching do not necessarily require reflection about the connection between teaching and the learning gains of students. Table 1 summarises the views of Zeichner and Liston in terms of the emphases and focus of reflective teaching traditions:

Table 1 Reflective Teaching Traditions

	Emphasis of the tradition:	Reflection is mostly about:
Academic	Subject matter and the representation and translation of subject matter knowledge to promote student understanding	The content of teaching
Social Efficiency	Scientific study of teaching as a basis for teaching expertise	The match of teaching to external research about teaching approaches
Developmentalist	Natural development of the learner as the basis for deciding about readiness for tasks	Teachers' own students
Reconstructionist	The institutional, cultural and political contexts that instruction is embedded in.	Teachers' own practices and the social conditions in which those practices occur
Generic	That reflecting on actions that they are deliberate and intentional actions will make them better	Teaching in general without concern for how teachers reflect, what it is about, or the extent to which the social and institutional contexts are considered

Student outcomes are not the explicit focus in any of these traditions of reflective teaching. While each tradition may include worthwhile considerations, the value of the reflection is diminished when the ideas are not connected to outcomes that students have experienced.

Inquiry into student learning calls for the kind of attitude towards responsibility for consequences described by Dewey (1933). According to Dewey, teachers are responsible not only for the personal and social and political consequences of their teaching, but also for the academic consequences. This necessitates thinking about the intellectual development that results from one's teaching, a notion supported by Ghaye & Ghaye (1998) in the 'Reflection-on improvement' focus of reflective conversations. This focus includes reflection on the improvement of both teaching practices and students' learning.

Schön's notion of 'reflection in action', as outlined below, is a useful basis for inquiry in teacher appraisal:

... the practitioner allows himself to experience surprise, puzzlement, or confusion in a situation which he finds uncertain or unique. He reflects on the phenomena before him, and on the prior understandings which have been implicit in his behavior. He carries out an experiment which serves to generate both a new understanding of the phenomena and a change in the situation.... He does not separate thinking from doing.... Because his experimenting is a kind of action, implementation is built into his inquiry. (Schön, 1983, p. 69)

The appraisal model suggested here demands that the central phenomenon to be reflected on is student learning. The surprises, puzzlements and confusions that Schön values should be related to the learning (or lack of learning) of one's students, and the change that is generated through the reflection should be change in terms of learning gains (both teachers' and students').

It is possible that inquiry into student learning may be included in the appraisal process, but be overshadowed by a prevalence of inquiry into other matters. These could be of a compliance nature, where the completion, rather than quality, of teachers' documents is appraised. There are a myriad of teaching behaviours and practices that could be inquired into, as opposed to the outcomes of those practices. Similarly, there could be a focus on professional attributes such as communication skills, cooperation with colleagues, and contribution to wider school activities. While important, these attributes do not necessarily bring about student learning. Where there is little consideration of student learning, appraisal processes are unlikely to actually impact on teaching practice and student outcomes. Since time for such inquiry is limited, the focus should be on the core of schools' purpose – the learning of its students. This should be emphasized in the entire appraisal system, including appraisal rhetoric, the associated documentation, and appraisal practices such as observations and interviews.

*The Imperative for Student Learning Data to Inform Appraisal Inquiry*

Undertaking student learning focused inquiry in appraisal could mistakenly be considered enough to satisfy the approach argued for in this thesis. There is a further

proposition about the nature of the inquiry, which is particularly significant. For inquiry into student learning to be most productive, it needs to be based on data. Inquiry into data about student learning in teacher appraisal will significantly enhance the rigour of the process, and ensure that appraisers and appraisees are not left merely to assume or speculate about the impact of a teacher's practice. The use of data as a basis for discussing and examining connections between teaching and learning is crucial, and has a number of desirable outcomes, as outlined by Schmoker:

data can help us confront what we may wish to avoid and what is difficult to perceive, trace, or gauge; data can substantiate theories, inform decision, impel action, marshal support, thwart misperceptions and unwarranted optimism, maintain focus and goal-orientation, and capture and sustain collective energy and momentum. Data help us answer the primary question "what do we do next?" (Schmoker, 1996, p. 42)

As well as the benefits for teachers with regard to the use of data claimed by Schmoker, there are significant and sustainable benefits, most importantly, for students. In a study about the sustainability of a professional development approach, student achievement was found to be highest in schools where teachers analysed their students' achievement data and were members of data-based professional learning communities (Timperley, 2004; Timperley & Wiseman, 2002). If teacher appraisal is to facilitate learning gains for students, then principles of successful professional development, such as the analysis of student achievement data outlined above, will need to be harnessed in appraisal systems. Appraisal discussions should not only inquire into the impact of teaching on student learning, but also engage appraisers and appraisees in the analysis of student achievement data.

In schools that function as professional learning communities, teachers "examine together how well students are doing: i.e., they study student work and assessment data, they relate this to how they are teaching (i.e., to instructional practice), and they make continuous refinements individually and with each other (i.e., as a professional community)" (Fullan & Mascal, 2000, p. 60). This is in contrast to professional

communities that have individualistic or balkanized cultures in which teachers either leave each other alone or disagree and do not tend to engage in a process to solve differences.

The potential for teacher learning that results from the use of data-based inquiry provides the overriding rationale for its use. Warren-Little emphasizes the relationship between schools that make inquiry into student learning a “cornerstone of professional development” and the likelihood of teacher learning (Warren-Little, 1999, p. 235). This leads to what Dimmock (2000) refers to as “informed practice”. Informed practice draws on data or evidence, with consideration for the context in which the data lies, in a manner distinct from “best practice” approaches to teaching. In a best practice approach attention is not directed towards evidence of student learning, but towards the practices used. Inherent in a practice, or style-based, approach is an unquestioned acceptance that particular practices will impact on student learning, but the actual impact in any particular context is left untested. Appraisal should provide practitioners with an opportunity to test and challenge those assumptions through the use of data. Data, according to Senge, improves inquiry and enables people’s reasoning to be drawn out. He suggests that practitioners need to ask, “What data from my practice lead me to believe that this is the best way? What other alternatives might I consider?” (Senge, 2000, p. 285).

To maximise teacher learning data should be used in both confirming and disconfirming ways. Teachers need to be open to “differing – and even unexpected and surprising – evidence and interpretations” (Reagan, Case, & Brubacher, 2000, p. 41). Stoll et al. agree, and describe the importance of “enquiry mindedness” for teachers in their study:

[teachers were] adept not only at collecting data, but much more importantly, at thinking about it, interpreting it in context and using it wisely, so they were no longer in awe of data, nor were they looking for data to confirm their prejudices or endorse their practices. Instead, they were actively searching for understanding, struggling

to describe the complexity of their work and using systematic inquiry procedures to stand back and think about their school. (Stoll et al., 2003, p. 145)

Using data in this way requires teachers to be skilled in selecting, interpreting and analysing data.

A further reason to employ data as a basis for appraisal inquiry is to utilize the wide array of data that is collected in schools. While clearly not a sufficient reason on its own, to make productive use of data that is gathered for other purposes would be productive. It has already been argued that school leaders should use achievement data to improve teaching programmes as part of school-based curriculum review (Robinson, Phillips, & Timperley, 2002). Teacher appraisals would increase the contribution data make to teacher learning and to improved practice.

Management of the challenges and tensions involved in appraisal could be improved through the use of data. Appraisal discussions present participants with challenges and demands that are perceived as more easily avoided than confronted. According to Cardno and Piggot-Irvine (1997), participants may be challenged both intrapersonally (due to the emotions and beliefs involved) and interpersonally (due to the personal and professional interactions involved). Stake (1989) and Townsend (1995) also highlight the threatening nature of appraisal, and explain it in terms of the tension caused by the dual roles of improvement and accountability. Data have the potential to make the “undiscussable” (Piggot-Irvine, 2001b) discussable, since speculation, opinions and bias are less likely to surface when inquiry in an appraisal discussion is based on data.

Such an inquiry-minded approach to student data has implications for school assessment systems and for the assessment literacy required of those involved. The type of data, the curriculum areas it relates to and the manner in which it is aggregated would all need to take into account that it is likely to be used during appraisal. There is, then, a level of assessment literacy required of those involved,



but the benefits are clear. Such “coupling” of assessment practice and teacher professional development provide opportunities for creating teacher learning about student learning (Sykes, 1999).

The type of data that would be useful is wide in scope, including internally and externally generated data, data aggregated at the school, syndicate or classroom level, and individual samples of student work. Each of these data types is valid, depending on the purpose it is intended to support. Sykes views the use of multiple sources of evidence in teacher professional development, to which appraisal is clearly linked, as important in strengthening the connection between teacher and student learning (1999). Stoll et al. also highlight the importance of inquiry with an evidence base, but warn that “as schools are increasingly expected to use evidence (not only their own, but from elsewhere) to inform their practice, it is important to evaluate the quality of evidence and use it wisely in making decisions (Stoll et al., 2003, p. 146). Quality evidence, as Warren-Little points out, need not be costly to be powerful. She recommends that individuals and groups of teachers systematically study student work with the purpose of figuring out how the work has resulted from the practices and choices of teaching (Warren-Little, 1999).

To summarise, I am arguing for appraisal that is inquiry based, that the inquiry is focused on student learning and that data are used to provide evidence of connections between teachers’ practice and students’ learning (or shifts in achievement over time).

The idea of data-based inquiry into student learning is prevalent in the literature on school level reform and change (Earl & Fullan, 2003; Earl & Katz, 2002; Fullan, 1992, 2001, 2002; Hopkins, 2001; Robinson et al., 2002) and teacher professional development (Joyce & Showers, 2002; Timperley, Phillips, & Wiseman, 2002a, 2002c). It has not, however, been extended to inform the research and literature relating to teacher appraisal. I now turn to possible reasons for this situation.

*Possible Objections to Data-Based Inquiry into Student Learning in Appraisal*

There are a number of possible objections that could be made to this argument that may explain the absence of notions of data-based inquiry into student learning in the teacher appraisal literature. These objections include the extent to which teachers are accountable for students' results, the extent to which influences on student learning other than teaching are given emphasis, and the likelihood of perceived threat in such an approach. These possible objections will be addressed in turn.

*Apprehensions about accountability*

My argument is that appraisal should encourage teachers to think about not only the means of their teaching, but also the outcomes they have achieved (Senge, 2000). For some, the emphasis on student outcomes is provocative and challenging. Focusing on inquiry into student learning should not be confused, however, with holding teachers accountable for their students' results. The accountability is not for the results themselves, but for ensuring three things. Firstly, those appraising others should be accountable for promoting inquiry into student learning. Secondly, those being appraised need to be capable of inquiring into the effects of their teaching on their students. Teachers need to have the capability, in terms of skills, attitudes and dispositions, which will enable them to make links between data about their students' learning, and their past and future practice. Thirdly, those being appraised should be able to provide evidence that they have initiated their own inquiry into the learning of their students and that this has led to reflection about their own teaching. This ethical imperative is consistent with the non-consequentialist view of ethics (Reagan et al., 2000) in which the consequences are not the issue, but the universal moral and ethical principles that guide behaviour. In this case, student learning focused inquiry is of utmost importance, and should therefore be the guiding principle for appraisal behaviour.

*Concern about the recognition of the multiple and complex influences on student learning*

Another possible objection to a focus on student learning and achievement is that there are too many other factors that influence student learning to expect teachers to focus on their own practice. The importance that is given to concentrating on the

impact of teaching on students does not discount the wide range and complexity of other influences on student learning. It is acknowledged that it is difficult to establish a precise causal connection between a teacher's practice and the learning that results for their students (Brereton, 2001; Muijs & Reynolds, 2001; Sykes, 1999; West-Burnham & O'Neill, 2001).

To address this complexity, and the "untidy, ill defined, and difficult-to evaluate" nature of the classroom setting, Grootenboer makes a distinction between "bureaucratic managerial" and "professional" perspectives of teaching that underpin appraisal (2000, p. 122). These suggest that a professional perspective is required to address the complexity of the teaching task, rather than a bureaucratic perspective which seeks to evaluate definable phenomena that are measurable and quantifiable. A professional perspective that values formative appraisal for professional growth is far better placed to address complexity, than summative appraisal that aims to manage and judge.

The fact that teaching is complex and that causal connections between teaching and learning are difficult to establish does not, however, mean that causal connections do not exist. There is actually significant research evidence, and widespread recognition, that effective teaching does impact on learning. According to Muijs and Reynolds, British school effectiveness research has increasingly shown that "the influence of the teacher and of the learning level considerably exceeds that of the school" (2001, p. viii). This is in contrast to assumptions about differences in achievement being due to between-school variations.

The Ministry of Education also reported, in a paper on the impact of teaching and schools on variance in outcomes, that quality teaching is the largest single education system influence on student achievement (Ministry of Education, 2004a). According to this analysis of more than 30 local and international studies, between 16 and 60 per cent of variance in student achievement can be attributed to teaching and only up to 20.9 per cent attributed to schools. These findings, whilst self-evident to most

teachers, should alert educators to the importance of inquiring into teaching and the impacts that teaching has on students' learning. It should also prompt attention toward the potential for appraisal (since it focuses on teaching rather than the school level of the education system) to positively influence the quality of teaching and, in turn, to influence educational outcomes for students.

Hattie's work also signals the relative significance of what teachers know, do and care about in influencing student achievement. He acknowledges that there are a myriad of influences that impact on student learning. These include schools, different teachers, subjects, school administration systems, students' ages, as well as gender, prior ability, quality of instruction and teaching style moderators (Hattie, 2002). Hattie's meta-analysis assesses the relative magnitude of the effect of a number of influences that are categorised according to teacher process, teacher methods, student influences, home influences, social influences and school policy influences. His work reveals that the most powerful factors of all those that occur in schools are, in fact, those that are within a teacher's control.

While influences beyond the teacher's control are important, such as students prior cognitive ability and home factors, most of the influences that have a bigger than average impact are effected by the teacher. These are, in summary, teachers' use of critical innovations, their provision of feedback about constructions students have made from information, and their use of specific and challenging goals with students (1999a; Hattie, 1999b). Teachers can, through the use of these and other approaches, have a positive impact on student learning relative to other school-based influences. Teachers should then inquire into how their own practice actually fulfils that potential to make a difference to their students' learning.

While focusing on the impact teaching has on students is of principal importance, the other factors that also influence student outcomes could also be acknowledged and considered as part of the inquiry Brereton suggests:

Many in the profession, and outside, firmly believe in the underlying principle that the output of school teachers, that is, the education of their pupils, is multidimensional, not easy to measure and dependent upon a plethora of external and imperceptible influences. Teachers will resent being held professionally accountable for targets that are at the mercy of so many variables beyond their control.... the best way to manage these external influences was not to pretend otherwise, but to recognise this inevitability 'up-front' and to identify what potential outcomes these influences may have. (Brereton, 2001, p. 33)

Influences beyond the control of a teacher provide an important context for discussion of the influences within their control. Any consideration of external influences should not, however, overshadow the significance of the impact a teacher has, and the necessity to inquire into that impact. While a teacher does not have control over all of the influences on students' learning, they may have control over the impact that those influences have. Darling-Hammond has emphasised that "students' right to learn is directly tied to their teachers' opportunities to learn what they need to know to teach well" (1996, p. 6). Most importantly, teachers need to know how to ensure that diverse learners, regardless of the other influences on their learning, have their right to learn met. In turn, teachers need opportunities to learn, and that learning can be situated in appraisal.

*Perceptions of data-based inquiry into student learning as a threatening process*

Some may also object to this approach on the grounds that data-based inquiry into student learning is threatening for teachers. Those managing the appraisal process need to be mindful that performance appraisal is inherently a threatening endeavour (Cardno, 2001 ). Misunderstanding of the subtle differences between accountability for results and accountability for inquiry could intensify the perception of threat concerning the appraisal approach proposed here. Focusing on the impact of teachers' practice on their students may indeed present appraisers with dilemmas, particularly where this involves surfacing data that reflects negatively on the teacher. A typical response would be to avoid that threat by abdicating responsibility for inquiring into the connections between teachers' practice and students' learning.

...without explicitly inquiring into the moral obligations inherent in our work, we insulate ourselves from personal responsibility for any negative consequences that students may suffer from our decisions. We are, thus, shielded from the burden to take action. (Senge, 2000, p. 278)

Cardno and Piggot-Irvine also describe threat avoidance as a response to dilemmas involved in appraisal. They suggest that appraisers typically emphasise or suppress either the formative or the evaluative function of appraisal (Cardno & Piggot-Irvine, 1997). Defensive responses are used in an effort to protect those involved. Avoidance may well protect the emotional concerns of an appraiser, but good intentions fail to promote learning for either the teacher or his or her students. Besides avoidance, other typical strategies are “soft-sell (being nice), and hard-sell (being nasty)” (Cardno, 2001, p. 148). Neither of these is productive in terms of improving practice and could be replaced by a strategy based on evidence, rather than one focused on being “nice” or “nasty”.

There are a number of ways that the threatening nature of appraisal could be addressed, and productive (rather than defensive) responses promoted in teacher appraisal. In particular, the features of learning communities (Sykes, 1999), inquiry cultures (Goodlad, Soder, & Sirotnik, 1990; Reagan et al., 2000; Warren-Little, 1999; Zeichner & Liston, 1996) and critical dialogue (Robinson, 1993) are well suited to student learning focused teacher appraisal.

#### *Inquiry Cultures, Critical Dialogue and Learning Communities*

Where a culture of inquiry exists the inquiry process will include characteristics of intellectual curiosity, motivation, and openness to challenge. Reagan et al. suggest that a culture of inquiry requires “ongoing commitment to valuing curiosity, mutual respect and support among teachers and between teachers and administrators, a willingness to try new ideas and practices, and the ability to remain open to the unforeseen and unexpected” (Reagan et al., 2000, p. 43). Warren-Little goes further, suggesting that not only should a culture of inquiry be embraced, but that schools’

resources should be organized in a way that places inquiry into student learning at the heart of professional development for teachers (Warren-Little, 1999).

Critical dialogue involves strategies that could contribute to a culture of inquiry. Using this model, appraisers say what they think, why they think it, and then check with the appraisee (Robinson, 1993, p. 55). Following these three rules supports two of the guiding values of the model – increasing valid information for all and enhancing freedom of informed choice. The third value is about enhancing commitment and responsibility (Robinson, 1993, p. 140). These values, together with strategies of openness, public testing and bilateral control, serve to limit the threat of appraisal. An appraisal in which the appraiser merely states their thinking is likely to generate a sense of threat. Alternatively, an appraisal in which the appraiser gives reasons for their thoughts, based on data, and then checks their understanding with the appraisee is not likely to generate a sense of threat. Critical dialogue minimises threat by leading to better problem solving, quality information, and commitment from those involved. It involves taking shared rather than unilateral responsibility for recognizing and managing emotional threat.

An inquiry culture is likely to exist in a professional learning community. The characteristics of such communities are particularly conducive to a culture of inquiry, and student learning focused appraisal. They include:

...the presence of shared norms and values cultivated through regular reflective dialogue; the deprivatization of practice, referring to open scrutiny of individual teachers' practice through dialogue, observation and feedback, and examination of student work; a school wide focus on student learning; and high levels of collaboration around such tasks as curriculum development, coordination of instruction, and assessment of student learning. (Sykes, 1999, p. 156)

In professional learning communities with contexts that enable the kind of appraisal inquiry argued for here, the connection between teacher and student learning is strengthened (Sykes, 1999). The notion of teachers themselves as learners (Hodkinson & Hodkinson, 2005; James, 2005) in the appraisal context, is crucial to this appraisal approach.

The nature of the professional communities in which appraisal takes place, and the nature of appraisal itself, is clearly and purposely influenced by the policy context in which it occurs. In the next section I will discuss the policy context directly relating to teacher appraisal, in relation to the argument for data-based inquiry into student learning. It will also consider this approach to appraisal in relation to the broader educational policy context in New Zealand, including the Ministry of Education's National Education Priorities, Planning and Reporting Policy, National Assessment Strategy and the Iterative Best Evidence Synthesis Programme.

## ***Part 2: The Professional Context of Appraisal in New Zealand Schools***

### *The Existing Appraisal Policy Context*

In the previous section I have argued for an approach to teacher appraisal that is focused on student learning and that draws on data to inform appraisal inquiry. In this section I will introduce key features of the appraisal policy context and examine the extent to which they facilitate this type of approach to appraisal. I will also consider how the wider policy context of the New Zealand education system fosters this type of approach to appraisal.

Three key features of the appraisal policy context will be introduced and examined here. They are the policy requirements for

1. appraisal process elements
2. the focus of appraisal
3. principles to underpin appraisal

### *Appraisal process elements*

At least once every twelve months Boards of Trustees are responsible for appraising the performance of each teacher in their school according to their policy. The



elements of appraisal, as can be seen in Figure 2, are an important part of a wider process known as performance management.

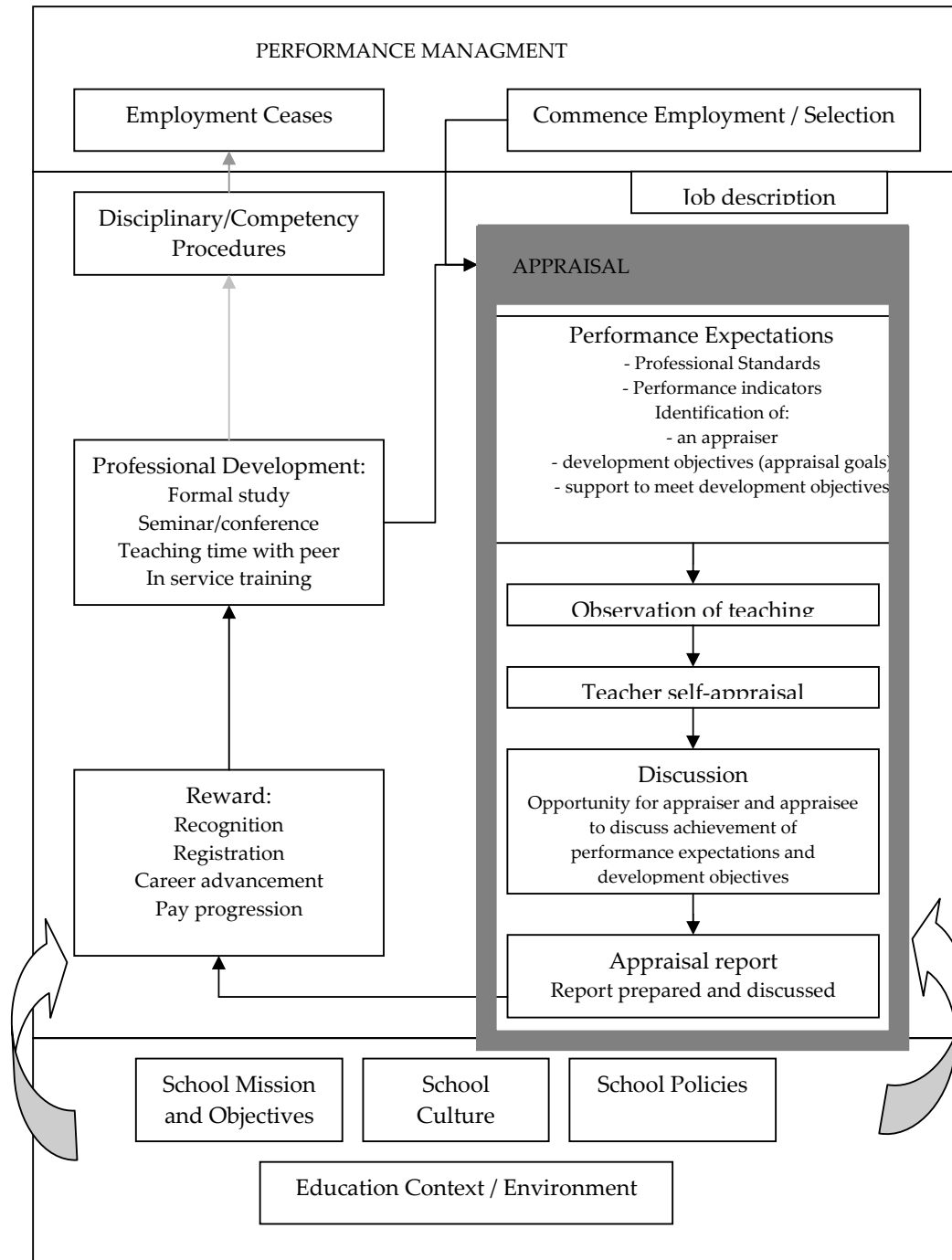


Figure 2 Performance Management in New Zealand Schools

The scope of performance management requirements undoubtedly makes the process a time consuming and demanding one. The rationale for enduring the challenges it

presents lies in two significant potential benefits. The first is the potential for performance management to lead to a positive impact on student learning. The second is the potential opportunity it presents for promoting teacher inquiry into student learning. The required elements of the appraisal process, outlined below, are well placed to support both teacher inquiry into student learning, and the improvement of student learning:

Identification of an appraiser, in consultation with the teacher concerned

Development of a written statement of performance expectations, in consultation with each teacher

Identification and written specification of one or more development objectives to be achieved during the period for which the performance expectations apply

For each development objective, the identification and written specification of the assistance or support to be provided

Observation of teaching

Self-appraisal by the teacher

Opportunity for the teacher to discuss their achievement of the performance expectations and the development objective(s) with their appraiser

An appraisal report prepared and discussed in consultation with the teacher

(Ministry of Education, 1997a, p. 5)

At a number of stages of the process – those in which appraisers and appraisees develop objectives, observe teaching, engage in self-appraisal, discuss with each other the achievements and a write a report – attention could be focused on data about student learning. The guidelines as they stand, though, do not suggest this possibility and certainly do not imply this as a desirable approach.

#### *The focus of appraisal*

Appraisal guidelines provided to schools by the Ministry of Education also outline what should be appraised. In this outline of aspects of teachers' performance that

should be appraised (see Table 2), student outcomes or achievement is conspicuous by its absence. Aspects of planning, preparation, teaching techniques, classroom management and environment are disconnected from their relationship to student learning and their impact on student achievement.

Table 2 Aspects of Teachers' Performance Which Should be Appraised

Key responsibilities/performance areas	Examples
Teaching responsibilities	planning and preparation teaching techniques classroom management classroom environment curriculum knowledge student assessment
School-wide responsibilities	contribution to curriculum leadership school-wide planning school goals the effective operation of the school as a whole pastoral activities and student counselling community relationships
Management responsibilities	planning decision-making reporting professional leadership resource management

(Ministry of Education, 1997a, p. 5)

The examples provided promote a task-driven, rather than outcome-driven, approach to appraisal in which the completion of particular tasks is likely to be focused on rather than the achievement of the overall purpose of the tasks – to improve student learning. A teacher may complete planning and prepare the class programme to an apparently high standard, use a range of seemingly impressive teaching techniques,

manage the class in an efficient and orderly manner, create a glamorous class environment, have strong knowledge of curriculum and assess students. It is possible, though, that the same teacher's students may achieve little or no gains in their learning. To limit inquiry into just these tasks is to rely too greatly on an assumption that those practices will in fact lead to positive impacts on student learning. The nature of the teaching task, individual students and the classroom context is far too complex to justify making such assumptions. Rather, aspects of a teachers' performance that should be appraised should focus on student outcomes, and connections between teachers' practices and those outcomes.

The risks of making assumptions about the impact particular practices or approaches have on student outcomes have been highlighted at a system-wide level by Alton-Lee (2004, p. 1). She alludes to the Education Review Office finding that a learning styles approach is used widely in New Zealand schools despite findings by Higgins that such an approach is linked to less effective instructional experiences for Maori and Pasifika learners (2001) and despite international evidence about the lack of positive (and presence of negative) impacts of this approach. Inappropriate assumptions about the application of a particular learning style approach to a particular group of students, and failure to inquire into the impact of that practice, can lead to undesirable outcomes.

One year after the initial guidelines were given to schools (Ministry of Education, 1997a), the focus of appraisal was clarified further through the introduction of professional standards (Ministry of Education, 1998a, 1998b, 1998c). In 1999 professional standards were linked to the Primary Teachers' Collective Employment Agreement and it became required to assess these as part of each teacher's appraisal. The reference text "Teacher Performance Management" (Ministry of Education, 1998c) was developed in response to requests from schools for a practical reference to integrate the Interim Professional Standards into performance management systems. The objective of performance management in schools, according to this text, is to "improve learning outcomes for students by improving the quality of teaching and

leadership” (Ministry of Education, 1998c, p.3). The text suggests students’ performance results and assessment records in a list of possible methods of assessing teachers. However, none of the other procedural guidelines or examples explicitly suggest a focus on student learning.

None of the seven professional standards dimensions (professional knowledge, teaching techniques, motivation of students, classroom management, communication, support for & cooperation with colleagues and contribution to wider school activities) explicitly relates to the outcomes of students’ learning. Across those dimensions there are 24 indicators for a fully registered teacher. Only one of those, “demonstrate a range of effective teaching techniques” has the potential to require a data-based inquiry into student learning approach to appraisal. Those working with this indicator, though, may or may not have an understanding of “effectiveness” that is grounded in student achievement.

A key resource used by schools to guide their appraisal systems, “Teacher Performance Management” (Ministry of Education, 1998c), also outlines 78 sample performance indicators. It is stressed that the examples are guidelines only, which schools may or may not use in developing their own indicators. Only three of these 78 example indicators for a fully registered teacher go any way towards promoting the data-based inquiry approach advocated here. They are “adapts own teaching approaches and techniques to maximise students’ learning opportunities and achievement”, “reflects on own teaching approaches & techniques & takes action to improve” and “uses assessment results to improve the teaching, learning and assessment cycle” (Ministry of Education, 1998c, p.28-30). Each of these represents just part of the proposed appraisal approach. Only when considered together, which is unlikely since they are spread throughout the 78 examples, could they really support appraisal inquiry that focuses on student learning and utilises data as a basis for inquiry.

The professional standards and sample indicators do not strongly support a data-based inquiry into student learning approach to appraisal mainly because those that do align with this idea are few in number – only one of the 24 standards, and three of the 78 sample indicators. It is unlikely that appraisal practices in schools will make inquiry into student learning and data a prominent aspect of appraisal when the national policy itself does not emphasise the importance of such an approach. Schools themselves may choose to develop appraisal systems that focus on student learning, and data about that learning. Widespread acceptance of such an approach, though, is dependent on the professional and policy context in which appraisal is situated.

*Principles to underpin appraisal*

A number of Ministry of Education published documents that guide schools' implementation of performance management, whilst not directly encouraging student learning to be focused on or data to be used, do reflect a concern for student learning. Ministry of Education Guidelines state that the primary purpose of appraisal is about "improving the quality of teaching (and therefore learning)" (Ministry of Education, 1997a) and "maximizing students' learning outcomes" (Ministry of Education, 1998c). These reflect a theme in the wider appraisal literature about the importance of appraisal for improving student learning (Bailey, 1997; Battersby, 1991; Fullan, ; Irons, 1993; Stewart, 1997; Townsend, 1995).

The original, and most current, guidelines for implementing appraisal outline seven key principles, presented in Table 3, which should underpin appraisal policies and processes:

Table 3 Principles That Should Underpin Appraisal Policies and Processes

Principles	Concept	Purpose
Have a professional development orientation	Professional orientation	To foster positive improvement and professional growth and to improve teaching and learning
Are appropriate to individual teachers, the school and the wider community	Flexibility	To meet individual and school needs
Are developed in a consultative manner with teachers	Consultation	To create teacher ownership of the process and outcomes
Are open and transparent	Transparency	To ensure no hidden agendas
Are part of an integrated performance management system operating within the school	Integration	To achieve linkage to other key performance management elements
Are timely and helpful to the individual teacher	Timeliness	To integrate the appraisal process with individual teacher needs and other school events
Give consideration to matters of confidentiality, including the provisions of the Privacy Act and the Official Information Act	Confidentiality	To give confidence to, and respect the rights of, participants

(Ministry of Education, 1997a, p. 4-8)

Student learning is referred to in the purpose of the “professional orientation” principle, which suggests that appraisal is “to foster positive improvement and professional growth and to improve teaching and learning”. The principle itself, though, does not explicitly suggest that data about student learning should be central to the process.

In this section I have examined guidelines provided to schools that shape the elements of appraisal processes, the focus of appraisals, and the principles and purposes that underpin appraisal. It is clear from this examination that the current framework for appraisal in New Zealand schools does allow schools to use data-based inquiry into the impact teachers’ practices have on learning outcomes for



students. Inquiry into these impacts is, though, by no means required or even promoted. A number of more recent policy initiatives, while not specifically about appraisal, do suggest an emerging professional context in this country in which a focus on student learning based on the analysis of achievement data is valued.

*The Promise of the Emerging Professional Context and National Policy*

In recent years an increasing number of policy initiatives and publications have been introduced that support the notion of data-based inquiry into student learning. Although these policies are not explicitly about appraisal, they are connected, and the evidence based inquiry direction they take would be appropriate to apply to performance management in schools. Amongst them are the National Education Priorities, recently revised Planning and Reporting legislation, the 'Iterative Best Evidence Synthesis' programme, as well as the national assessment strategy.

*National Education Priorities*

The National Education Priorities document highlights, as well as goals that are priorities for the sector, a number of approaches that are needed to improve the state of education. One of those is "focusing on outcomes" (Ministry of Education, 2003b, p.9). The strategies suggested by the Ministry of Education to realise this focus on outcomes, in regard to the primary sector, are outlined below:

Making learning outcomes central to all debates about education and focusing on the impact of what we do on learning

Requiring all providers to be much more deliberate, explicit and strategic about what improvements in student learning they will achieve and the basis from which they are making these judgements

Increasing the systematic use of information to continuously improve institutional performance

Developing leadership capacity for outcomes-focused management across the education system, especially in the schooling sector

Implementing a planning and reporting regime in the schooling sector

(Ministry of Education, 2003b, p.9)

In Chapter 6, these strategies are adapted to reflect the potential that teacher appraisal processes have to draw on these strategies and encompass a focus on student learning and the use of data.

*Schools' Planning and Reporting Policy*

The Education Standards Act ("Education Standards Act," 2001) made a number of amendments to the process of self-review, strategic planning and reporting, that schools are required to undertake. Since 2003 school charters have been required to include an annually updated section that sets out the following five aspects of the schools' plan – aims, directions, objectives, priorities, and targets. Each of these must relate to intended student outcomes. It also requires that an annual report is submitted to the Ministry of Education that includes "analysis of any variance between the schools performance and the relevant aims, objectives, directions, priorities, or targets set out in the school charter" ("Education Standards Act," 2001). Performance management and appraisal processes in schools are not directly part of this policy. There are, however, a number of approaches taken in the new planning processes for schools that could enrich and provide focus for teacher appraisal.

The Ministry of Education emphasise that responsiveness to evidence about student learning outcomes is important in good teaching and school management. This sentiment is evident in many of the statements in the "Planning for Better Student Outcomes" series of newsletters provided to schools during 2003 to support them in meeting the new requirements:

The first questions for evidence-based planning ask what the evidence is telling us: What are students learning well? What are they not learning well? Which students are learning well, and which students are not learning well? These questions can be answered by data analysis. The second question asks why some students are not achieving. The answer to this question does not come from an analysis of the student outcomes data but from a reflective look at teaching practice. (Ministry of Education, 2003g, p.5)

The approach to appraisal proposed here also stresses the importance of inquiring into evidence of impact on student learning and could serve two purposes in relation to the quote above. Firstly, it could provide a context in which some evidence-based reflection could take place, and secondly it could provide a means for appraisers to establish if teachers have the capability to engage in evidence-based reflection for themselves. I would argue that for learning (both students' and teachers') to be genuinely prioritised, and for planning to really be strategic, appraisal processes should be in alignment with school review policy and also focus on evidence of impact on students' learning outcomes. This would also help schools "to adopt a culture of continuous improvement" which is one purpose of the planning and reporting requirements (Ministry of Education, 2003d, p.1). For continuous improvement to become part of a schools' culture, I would argue it should be a focus not only in schools' planning and self-review processes, but also in all aspects of their performance management and other processes.

*Iterative Best Evidence Synthesis Programme (BES)*

The focus on evidence into student learning evident in the national education priorities, as well as in the planning and reporting requirements, correspond with key purposes of the Ministry of Education's Iterative Best Evidence Synthesis (BES) programme. The BES programme aims to deepen understandings of what works in education through the synthesis of research literature and evidence. It is a "knowledge building strategy to assist in the project of strengthening the accessibility and use of rigorous evidence-based research in education... Its purpose is to systematically identify, evaluate, analyze, synthesise, and make accessible, relevant evidence linked to a range of learner outcomes" (Alton-Lee, 2004, p. 1).

While the Best Evidence Syntheses are not intended to relate directly to appraisal practices in schools, there are a number of features of the BES programme that could potentially be paralleled in appraisal. The most obvious is the notion of making accessible evidence about what works. A primary purpose of the BES programme is "to illuminate educational influences that can make a bigger difference for desirable learning outcomes" (Alton-Lee, 2004, p. 2). Appraisers and appraisees could use the

appraisal process as a means of eliciting and illuminating aspects of their own practice that have made the most (and also the least) difference for their own students.

Best Evidence Syntheses draw on “analyses of patterns and exceptions in the findings, analyses of comparative magnitude of impact of various influences, and consideration of explanatory coherence to interrogate... evidence” (Alton-Lee, 2004, p. 2). Similarly, appraisal could potentially encourage teachers to ask questions that reflect those concerns at an individual teacher and classroom level. Important questions could include – What trends are there for my students as a result of my practice?, What did I do that had the most and least impact on my students?, and How good is my explanation about reasons for my students’ learning and lack of learning? Educators engaging in appraisal could also ensure that reference is made to multiple sources of evidence, and that a wide range of outcomes are considered, as do Best Evidence Syntheses.

*National Assessment Strategy*

An approach to appraisal that focuses on data-based inquiry, as advocated here, requires of participants at least a reasonable degree of assessment literacy.

Fullan defines assessment literacy in terms of:

- the capacity to examine student data and make sense of it
- the ability to make changes in teaching and schools derived from that data
- commitment to engaging in external assessment discussions, that is the active seeking of external standards against which to test performance

(Fullan, 1999)

Since the national assessment strategy began, a number of national initiatives to provide schools with external assessment information, and to support them in gathering their own information, have been developed or continued. These include, amongst others, the National Education Monitoring Project, School Entry

Assessment, Assessment Resource Banks, Project AsTTle, and the National Exemplars. They are briefly outlined below.

The National Education Monitoring Project's (NEMP) purpose is to provide quality information about the achievement of New Zealand students. The NEMP monitors student achievement of years 4 and 8 students and its reports reflect "the richness, scope and comprehensiveness of curriculum by giving a detailed, task by task picture of what students know, what they can do, and how they go about it" (2003).

In 1997 School Entry Assessment (SEA) was introduced into New Zealand primary schools. This standardised assessment package enables teachers to assess the nature and extent of their new entrant students' knowledge and skills in literacy, numeracy and oral language. More than half of the 576 schools and a majority of the 572 teachers who participated in a 2001 study reported using it in that year (Ministry of Education, 2003h).

Data are also made easier for teachers to gather through thousands of tasks available in the Assessment Resource Banks. These are an online collection addressing levels two to five of the English, mathematics and science curriculum statements. They are intended to help teachers assess progress within strands or achievement objectives, and to establish the relative performance of their students against the 'typical' performance of national samples of students at given year levels (New Zealand Council for Educational Research, 2005).

Assessment Tools for Teaching and Learning – He Pūnaha Aromatawai mō te Whakaako me te Ako (asTTle) provides tools for assessing literacy and numeracy, in English and te reo Māori. The asTTle tools enable teachers to "track the progress and achievement of both individual students and groups of students against national standards" (Ministry of Education, 2002a).

In addition to the assessment initiatives already outlined are the National Exemplars. These annotated examples of student work are intended to illustrate learning, achievement and quality at levels one to five in each curriculum area. Since discussion is an important component of teacher appraisal, and one of the three key purposes of the exemplars is to provide “a basis for discussing important qualities, aspects... [and] indicators of learning”, exemplars could be used in connection with appraisal. In fact, the action research project of Poskitt et al. revealed that teachers do not recognise the “basis for discussion” purpose of exemplars, nor do they use or anticipate using them for this purpose (Poskitt, Brown, Maw, & Taylor, 2003, p. 4-8). Use of exemplars as a basis for discussion could prompt teachers to use their own students’ work as a basis for discussion, and as evidence of their learning.

The strategies, tools, processes and resources outlined above suggest that schools already have access to ample assessment data that could be utilised as a basis for inquiry in appraisal. Nowhere in the assessment strategy summary tables (see Table 4) or other related documentation is a connection made between assessment evidence and performance management or appraisal. This appears to be a significant omission and a weakness of the wider policy context.

Table 4 Assessment Information and the National Assessment Strategy

	The purpose of gathering assessment information	Assessment information is needed by	Tools and processes	Use
Classroom level (students and teachers)	<p>Provide the most appropriate learning opportunities for students</p> <p>Provide feedback to students and identify their next learning steps</p> <p>Develop partnerships with parents</p> <p>Modify teaching programmes</p> <p>Ensure continuity of education for individual students</p>	<p>Students</p> <p>Parents</p> <p>Other teachers</p> <p>The school management team</p> <p>The principal</p>	<p>Exemplars</p> <p>Assessment for Better Learning development programmes</p> <p>Online development programmes</p>	<p>Focus:</p> <p>Feedback and guidance to learners</p> <p>Changes to teaching programmes</p> <p>Reliable, clear and valid information for parents and the next teacher</p>
School level	<p>Evaluate the success of the school's curriculum and teaching programmes</p> <p>Inform strategic planning and school development</p> <p>Improve the achievement of individual students and subgroups</p>	<p>The Board of Trustees</p> <p>The staff</p> <p>The school community</p>	<p>School review</p> <p>ERO reports</p> <p>Aggregated assessment information</p> <p>Professional development in assessment</p>	<p>To inform:</p> <p>Strategic planning</p> <p>Professional development</p> <p>Curriculum development</p> <p>Review of school policies, procedures and programmes</p>

Adapted from (Ministry of Education, 2001a, p. 8-9) and (Chamberlain, 2000, p. 25)

Chamberlain (2000, p. 26) asserts that “the gathering, analysis, reporting, and, most importantly, use of assessment information at classroom, school and system level is critical. Despite such suggestions it seems that the assessment strategies are not intended or recommended to be used to support teacher learning in the sense that teachers appraisal development objectives, observations and discussions might draw on the data they produce.

### *Part 3: Methodology*

The wider policy context outlined above allows for, though does not particularly emphasise or foster, appraisal processes that focus on evidence-based inquiry into student learning. The sequence of three studies reported in the following chapters sought to explore the extent to which schools have such a focus in their policies, professional standards indicators, appraisal discussions and development objectives. The main focus of the following three chapters, and the studies they are based on, are outlined below.

#### *Appraisal in Practice: The Studies Outlined*

Chapter 3 reports on a study of key appraisal related documentation – schools’ appraisal policies and documents outlining performance indicators for the Interim Professional Standards’ (Ministry of Education, 1998c, p. 19-21). I investigated the emphasis on student learning in the statements of purpose in appraisal policies, and the emphasis on student learning in the performance indicators that schools develop to assess teachers against the professional standards. The focus question was “to what extent is student learning a focus of inquiry in primary teacher appraisal?” In the following study, reported in Chapter 4, the research focus shifted from permanent records to reports of appraisal practices. This study examined the extent of data-based inquiry into student learning during appraisal discussions, and the reasons for



the level of reported inquiry. Eleven practitioners (both appraisers and appraisees) were interviewed about their most recent appraisal discussion to investigate two main questions – “is student learning a focus of inquiry during teacher appraisal discussions?” and “is student learning data referred to during appraisal discussions?” Participants in this study were also asked to respond to the suggestion that appraisal interviews should be focused on student learning and based on data about students’ learning. The results of the appraisal discussion study demonstrated the considerable influence appraisal goals play in determining the content and scope of appraisals. For this reason, a study about the nature of teachers’ appraisal goals was carried out, and is reported in Chapter 5. A questionnaire was completed by 68 teachers from eight schools about their appraisal goals. The focus of this study was the question “to what extent do teachers’ appraisal goals (or development objectives) focus on data-based inquiry into student learning in teacher appraisal?” The final chapter discusses the implication of these three studies for implementing appraisal systems that are focused on evidence-based discussion of student learning.

### *A Mixed Methods Approach*

The series of studies outlined above represent a mixed methods approach to research. Mixed methods research is referred to by Tashakkori & Teddlie (2003) as “the third methodological movement”, a movement that shifts from a previously held incompatibility theses (referring to perceived incompatibility between quantitative and qualitative methods). This methodological movement is one that is evolving and developing, so terms and definitions relating to the use of mixed methods are not all commonly-held or agreed upon. Two key definitions of mixed methods research are used here to explain why I use the term ‘mixed methods approach’ in relation to this research. The first is taken from Tashakkori & Teddlie:

...mixed method research studies use qualitative and quantitative data collection and analysis techniques in either parallel or sequential phases. This mixing occurs in the methods section of a study...  
(Tashakkori & Teddlie, 2003a, p. 11)

The methods in this research are mixed since both qualitative and quantitative underpin the studies, and all three data sources were submitted to both qualitative and quantitative analysis (through quantizing). The quantitative question (relating to the relative extent of focus on student learning as compared to other aspects of appraisal foci) is important since findings in relation to that question strengthen the rationale for examining the qualitative questions (about the presence or absence of a focus on student leaning and data and the issues associated with that).

The second definition of mixed methods approaches is from Creswell:

...[in a mixed methods approach] the researcher tends to base knowledge claims on pragmatic grounds (e.g., consequence-oriented, problem-centred, and pluralistic). It employs strategies of inquiry that involve collecting data either simultaneously or sequentially to best understand research problems. The data collection also involves gathering both numeric information (e.g., on instruments) as well as text information (e.g., on interviews) so that the final database represents both quantitative and qualitative information. (Creswell, 2003, p. 20)

The knowledge claims, strategies of inquiry and research methods (as outlined by Creswell) of this series of studies, will be discussed here, though the methods associated with the individual studies will be outlined in more detail in their respective chapters.

#### *Pragmatic Knowledge Claims*

The research methodology is underpinned by what Creswell calls “pragmatic knowledge claims” (2003, p. 11). Research based on pragmatic knowledge claims, typical in mixed methods approaches, is concerned with problems of practice in the real world and the consequences of actions. In this case, teacher appraisal is presented as the real-world problematic context. The problem to be examined is about what the focus of appraisal should be, and the extent to which appraisals reflect that focus. The normative model outlined in the first part of this chapter

provides a framework against which to examine the problems of appraisal practice. These studies draw on both quantitative and qualitative assumptions about how, firstly, to evaluate the extent of focus given to student learning and data and, secondly, to explain the extent of that focus.

*Causal relations and teacher appraisal*

This research is concerned with causal relations at two levels. At the broadest level, the research seeks to emphasise causality between the nature of teacher appraisal (in particular its focus on student learning and data), and the teacher and student learning that results from appraisal.

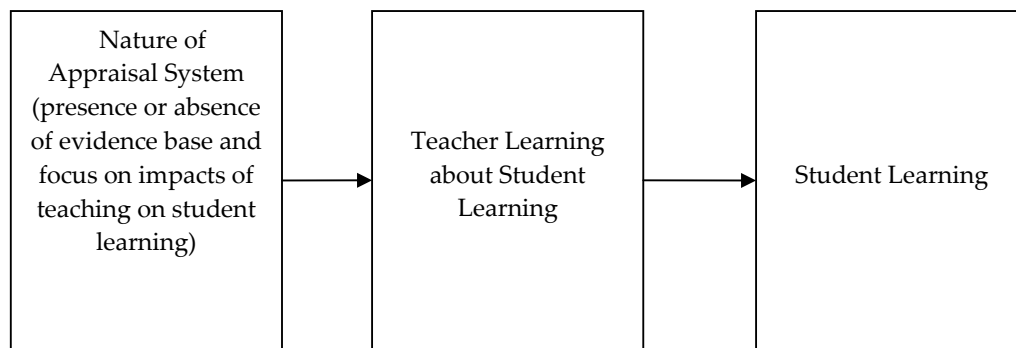


Figure 3 Causality in Teacher Appraisal

A second level of causality that emerged from the research findings indicated likely causal relations between elements of an appraisal process. That is, complex causality of several key elements of the appraisal process – causality between the content of appraisal policies, performance indicators, appraisal goals and the discussions practitioners engage in. The arrows in the figures above and below indicate likely direction of influence.

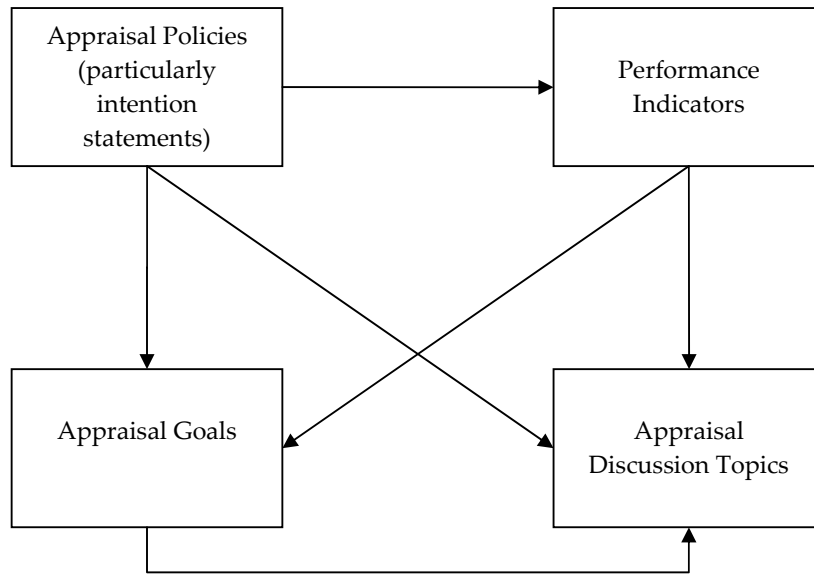


Figure 4 Causality between Teacher Appraisal Elements

These interpretations of causal relations reflect a typical position taken by pragmatists regarding causality (Teddlie, 2005). The causal relations are “transitory, contextually bound, and hard to identify” (p. 214) since there are complex interactions between the appraisal elements (as represented in Figure 4). Researcher values, represented within the normative theory developed at the beginning of this chapter, are also central to the interpretation of the causal relations.

*The Research Design: The Strategy and Methods – Integrating Qualitative and Quantitative Data*

Both qualitative and quantitative data were collected in this research. For example, the semi-structured interview schedules (Appendix A and Appendix B), which included mainly open questions about the topics of appraisal discussions, also included some closed-questions, such as “did you refer to data about student learning during your appraisal?” Closed questions were also used to seek confirmation about whether student learning was or was not discussed in an appraisal.

While there was an emphasis on the collection of qualitative data from research participants (textual documents and interview transcripts), in many instances that data were immediately transformed into quantitative data and used as a basis for further analysis. Others refer to this process as “quantitizing” (Miles & Huberman, 1994; Tashakkori & Teddlie, 2003a, p. 9). The qualitative aspect of that “data transformation” (Creswell, 2003, p. 212) involved examining which policy statements, performance indicators, descriptions, discussion points and appraisal goals related to student learning or data. The quantitative aspect involved the counting of references to student learning and ‘other’ categories to enable percentages to be calculated and comparisons made. For example, policy statements were coded in relation to their inclusion of reference to student learning and counted to establish what percentage of statements related to student learning, and to compare across schools the focus given to student learning (see Chapter 3).

The development of codes for the analysis of data across these studies occurred both inductively and deductively. The theoretical position outlined in Chapter 2 provided a basis for deductive codes and demonstrates an emphasis on what Onwuegbuzie and Teddlie refer to as “confirmatory thematic (a priori) analysis”. These codes related primarily to the themes of student learning focus, and reference to data. They are represented by the square boxes in Figure 5. There was, though, also significant attention to “exploratory (a posterior) analysis” (Onwuegbuzie & Teddlie, 2003, p. 364). The exploratory analysis, and the inductive codes that emerged, were important in providing an insight into the nature of appraisal inquiry where it was established that the inquiry did not focus on student learning or refer to student learning data. Inductive codes are represented by the ovals in Figure 5.

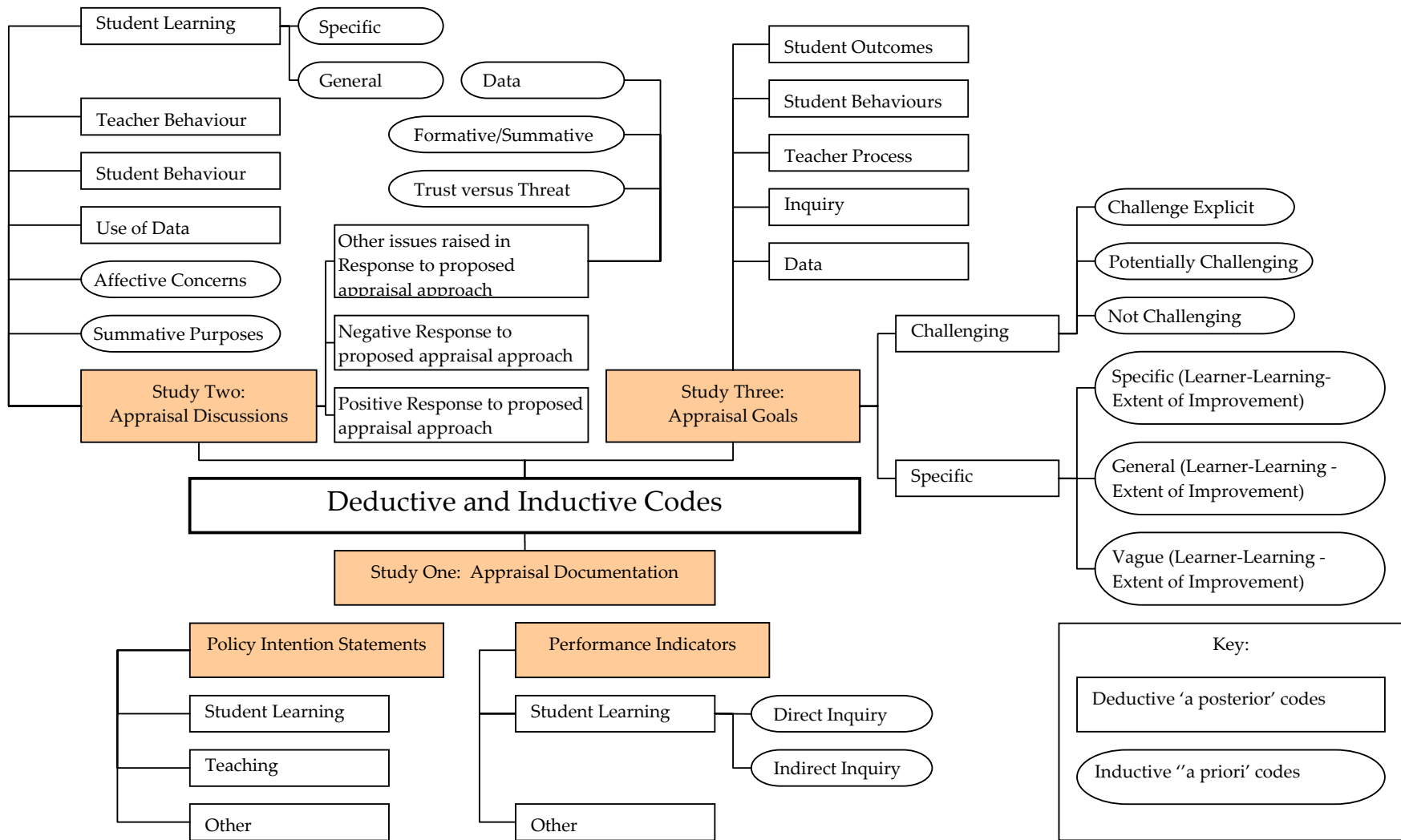


Figure 5 Deductive and Inductive Codes

In some instances inductive codes were developed as a result of close analysis of data coded to the deductive codes established at the outset. Where important distinctions were evident in data in the existing codes these were refined. For example, in study one important distinctions were apparent between indicators initially coded in the 'student learning' category. Close analysis revealed a distinction between indicators directly focused on student learning, or indirectly through a practice that was assumed to impact on student learning. Hence the deductive code 'student learning' was elaborated into the two inductive codes 'direct inquiry' and 'indirect inquiry'. This elaboration of deductive codes also occurred in the six categories of challenge in the appraisal goals study. The emergent themes supported the development of the theory about the significance of focusing explicitly on student learning in appraisal. In other instances the inductive codes did not elaborate deductive codes, but represented new themes not signalled by the prior theoretical position – for example, the 'affective' and 'summative purposes' codes in the appraisal discussions study.

The combinations of qualitative and quantitative methods for data collection and analysis are outlined in the following 'Mixed Methods Approach' figure.

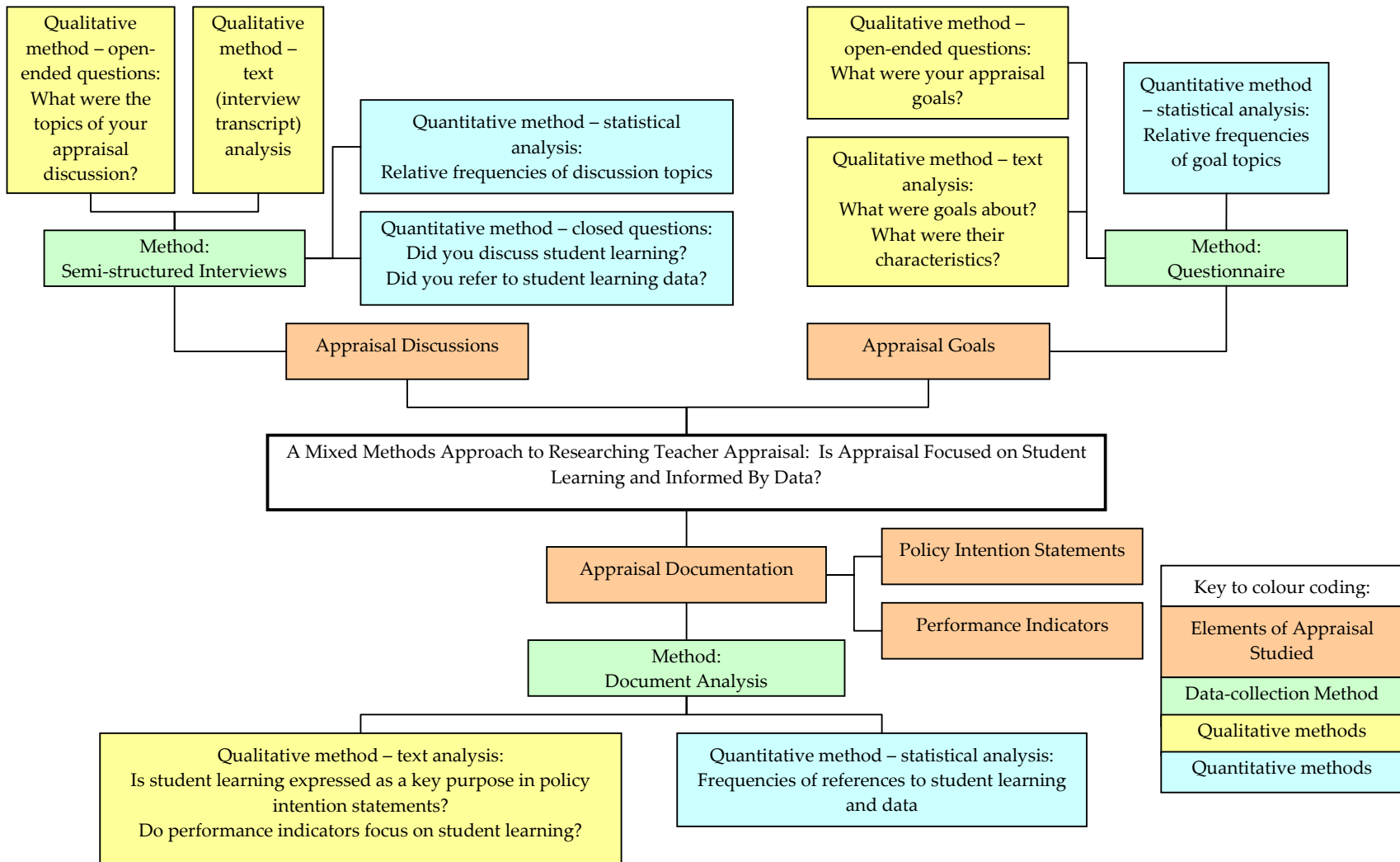


Figure 6 The Mixed Methods Approach



Mixed data types were necessary since quantitative data alone would not have provided access to the lived experiences of appraisers and appraisees that are crucial to understanding, in a rich and contextualised way, the focus of the dialogue in appraisal discussions. Similarly, qualitative analysis alone would not have provided an insight into the relative frequency with which educators focus on student learning, as opposed to other matters of teaching practice. Establishing the relative frequency of focus was important for determining the extent to which current appraisal practices in the sample schools approximate the normative model of appraisal argued for at the start of this chapter. For these reasons, a mixed methods approach was appropriate.

The research outlined above represents what Tashakkori and Teddlie refer to as a multistrand design. They have developed a typology of mixed methods research designs that encompass the many diverse designs in the literature. In particular, the design type that they call “multistrand conversion mixed model” captures the key elements of this research design (2003b p. 686-687). Figure 7 shows how this model has been adapted to this research.

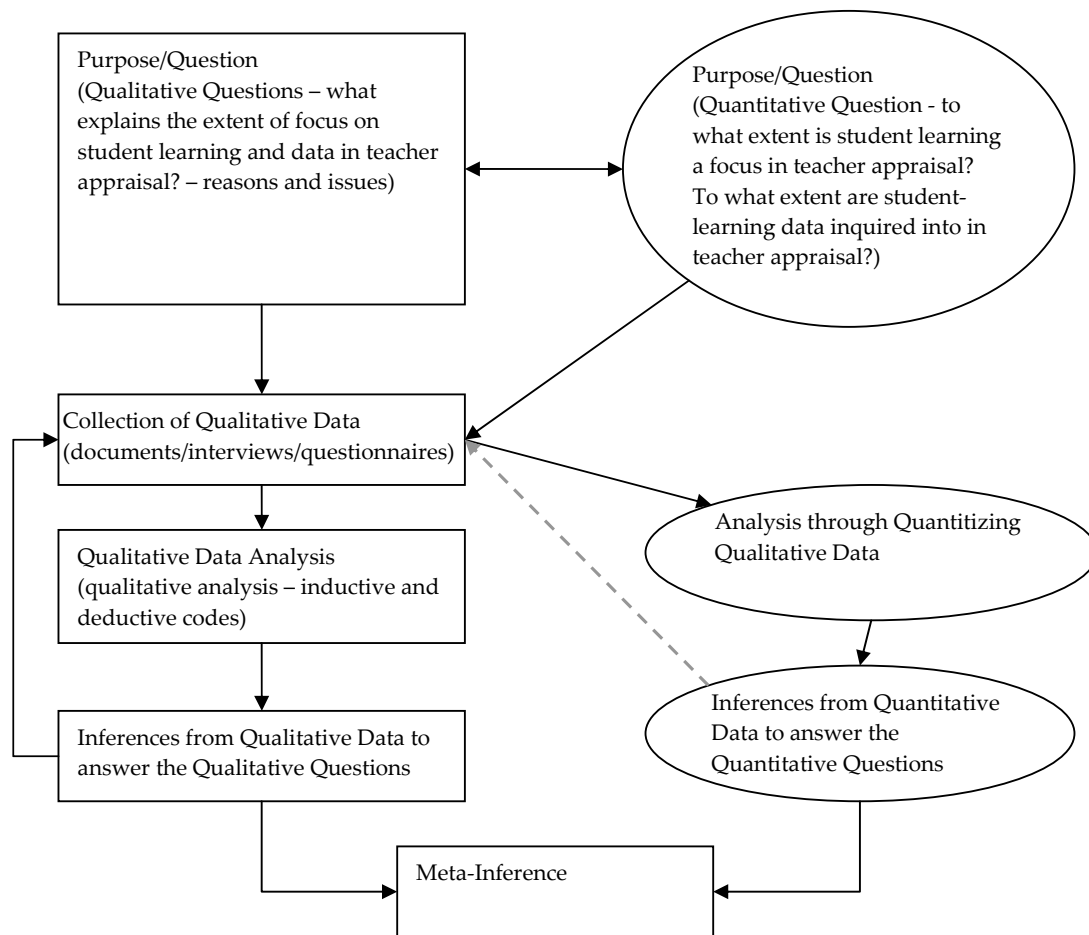


Figure 7 Research Design Type - Multistrand Conversion Mixed Model

Based on Tashakkori and Teddlie, 2003, p. 689

Tashakkori and Teddlie explain the multistrand conversion mixed model as one in which “multiple approach questions are asked. One type of data is collected and analysed and is then transformed to another data type (qualitized/quantitized) and analysed accordingly. Two types of inferences are made on the basis of each set of results and are pulled together at the end to generate meta-inferences” (Tashakkori & Teddlie, 2003, p. 689). The grey dashed arrow represents a variation to Tashakkori and Teddlie’s representation of the multistrand conversion mixed model design. It shows the sequence in which the quantitative analysis/inferences revealed a need for subsequent data collection (a questionnaire about teachers’ appraisal goals).

The mixed methods approach to this research served to enhance the descriptive, interpretive and theoretical validity of the research (Johnson & Turner, 2003). For

example, to increase descriptive validity qualitative data (quotes from participants) are reported to ensure that descriptions of the participants' accounts of their appraisal discussions are accurate. To increase interpretive validity a semi-structured approach to interviews was taken. The researcher checked her conclusions about the absence or presence of a student learning focus with participants during interviews, and also gave them an opportunity to confirm that the transcripts reflected their views accurately. The focus on these types of validity (descriptive, interpretive and theoretical) also reflects the dominance of qualitative rather than quantitative concerns in these studies. Findings (reported in the following chapters), regarding the extent to which student learning, and data about student learning is a focus in teacher appraisal, were confirmed across data types and methods, as well as across participants.

## CHAPTER 3

### The Emphasis on Student Learning in Appraisal Documentation

#### *Introduction*

School administration reforms in the 1990s led to requirements that schools develop and implement appraisal policies as part of their performance management system. Since 1999 it has been mandatory for the Interim Professional Standards for Primary School Deputy/Assistant Principals and Primary Teachers to be incorporated into those policies and procedures. National policies and material provided to schools to support their implementation do not make certain that outcomes of student learning, or the impact of teachers' practice, are a focus of appraisal inquiry (as outlined in Chapter 2). While there is no particular emphasis in national policy relating to appraisal on adopting this focus, it is not precluded as a possibility that schools might embrace. This chapter reports on the investigation of the emphasis on student learning in teacher appraisal in 17 primary schools through an analysis of their appraisal policies and supporting documents.

Emphasis on student learning in schools' policies and Professional Standards indicators is useful and, arguably, necessary, in order for schools to engage in the kind of appraisal inquiry proposed in Chapter 2. Schools' policies and indicators provide an important context for all other elements of the appraisal process including the setting of development objectives, observations, self-appraisal, and discussions or appraisal interviews. National policy statements also call for alignment between the requirements set out at national policy level and schools' policies and practices (see Figure 8). The requirements state that schools' performance management policies and practices should take into account "principles that underpin the policies and processes boards have in place for the appraisal of teacher performance; features of the process which is followed in appraisal; and aspects of teachers' performance

which should be appraised” (Ministry of Education, 1997a, p.4). The annual appraisal process for teachers is required to be “completed in accordance with the policy” (Ministry of Education, 1997a, p.5) and schools must make “annual assessments against all relevant Interim Professional Standards” (Ministry of Education, 1998c, p. 4). Figure 8 illustrates this requirement for alignment both between national and school levels of policymaking, and between elements of schools’ appraisal systems internally.

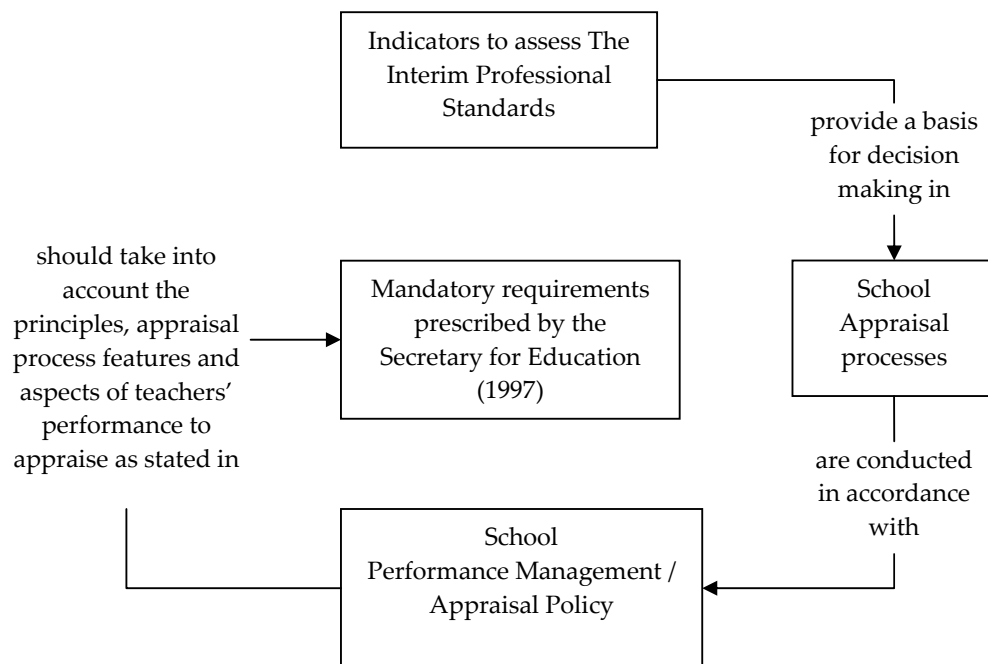


Figure 8 Appraisal Policy and Practice Alignment

It is clear, then, that school policies and indicators are significant aspects of the appraisal process in New Zealand primary schools. They have the potential to direct educators’ attention towards the ultimate purpose of teaching and professional obligation of teachers: improvement gains in students’ learning. That potential is most likely to be realised if all aspects of appraisal policy and practice (since they should be aligned) reflect a focus on student learning. The study of schools’ appraisal policies and their indicators for the Interim Professional Standards was to establish the extent to which they focus on student learning, and also to establish how well aligned elements of appraisal systems are.

### *School Appraisal Policies*

National Administration Guideline Two states that “each Board of Trustees, with the principal and teaching staff, is required to: (i) develop a strategic plan which documents how they are giving effect to the National Education Guidelines through their policies, plans and programmes, including those for curriculum, assessment and staff professional development” (Ministry of Education, 2003c). In accordance with this, schools are to develop appraisal policies as prescribed by the Secretary for Education in a series of guidelines published in 1997. The first of these, “Performance Management Systems: A Series of Guidelines of Performance Management Systems”, focused on performance appraisal (Ministry of Education, 1997a). That document sets out principles for appraisal policies and processes, features of the procedures to take place in appraisal and the aspects of teachers’ performance that should be appraised. It is a key document in guiding schools in the scope and content of their appraisal policies. These policies were the first of the two document types analysed in this study. They are considered part of the context or environment that makes schools’ appraisal systems unique (Ministry of Education, 1998c) and were relevant to this study because they include statements of intent that reflect the extent of focus on student learning.

### *School Indicators for the Interim Professional Standards*

The second document type analysed in this study was those outlining performance indicators that schools are strongly recommended and encouraged to develop in relation to the Interim Professional Standards. In a resource published in 1997, sample indicators were outlined to help schools to meet their obligation to assess teachers against the seven dimensions of the professional standards – professional knowledge, teaching techniques, motivation of students, classroom management, communication, support for and cooperation with colleagues and contributions to wider school activities (Ministry of Education, 1998c, p. 26). Indicators are statements of what teachers should have done to have met the professional standard (either as a beginning, fully registered or experienced teacher) relevant to them. Following are examples taken from the Ministry of Education indicator samples in

relation to one of the seven dimensions outlined above, teaching techniques, for a fully registered teacher:

Use appropriate teaching objectives, programmes, learning activities and assessment

Demonstrate a range of effective teaching techniques

Demonstrate flexibility and responsiveness

Impart subject content effectively

Use appropriate technology and resources

Reflect on teaching with a view to improvement

(Ministry of Education, 1998c, p. 26)

Where schools have followed the recommendation to develop indicators, and most have, they are generally similar to the examples above, though in many cases are more specific. Indicators provide a framework for the appraisal of teachers. They become a key tool used by schools to meet the objectives of performance management systems in schools, the first of which is “to improve learning outcomes for students by improving the quality of teaching and leadership” (Ministry of Education, 1998c, p. 3). Schools’ performance indicators documents were significant because they are a key text referred to during appraisal. They often guide observations and interview procedures, providing the framework for inquiry. Although these documents do not necessarily reveal actual appraisal practices, they provide a valuable insight into the kind of inquiry that takes place during teacher appraisal.

## *Method*

This section will outline the research setting, sources of data, and procedures for gathering and analysing data in relation to the question – to what extent is student learning a focus of inquiry in primary teacher appraisal?

### *The Research Setting*

Seventeen schools provided documentation and participated in the study. They were located in four Auckland regions and included contributing, full primary and intermediate schools of varying sizes. They also represented a range of roll sizes and deciles as shown in Table 5.



Table 5 Characteristics of Participating Schools

School Descriptors		Number of Schools
Local Authority	Auckland City	7
	North Shore City	3
	Waitakere City	3
	Manukau City	4
School Type	Contributing	12
	Full Primary	3
	Intermediate	2
School Roll	200-400	5
	400-600	4
	600-800	6
	800-1000	2
Decile	Low (Decile 1-3)	4
	Mid (Decile 4-7)	7
	High (Decile 8-10)	5

*Procedures*

All of the primary schools belonging to the University of Auckland Schools Consortium were invited to participate in this study. The Consortium comprised 46 primary schools that cooperate with the University in providing mentors and practicum experience in schools for teacher education students. The author has been involved with these schools for the past two years through her role as lecturer in University of Auckland pre-service teacher education programmes. Schools were sent an information package including an introduction, information sheet, project summary table and consent form. Principals or mentors in 12 schools, with whom the author had an established professional relationship, were approached by phone to explain the research. The remaining 34 schools were contacted only by mail. Schools were asked to provide school policy and supporting documents relating to performance appraisal and professional development. For the purposes of this study, only documents relating to appraisal were analysed.

Nineteen schools returned consent forms accepting the invitation to participate in the study. These comprised 11 of the 12 schools contacted personally, and eight of the 34 contacted only by mail. Three schools made contact declining to be involved. After follow up calls and visits to those that agreed to participate, documentation was provided by 17 schools. The response rate from those contacted personally was 75%, while for the others it was 24%. A code was used for each school to provide the researcher with a reference to the original data source. For instance, [S3] indicates school number three.

The sampling approach for this study was purposeful (Patton, 2002) which means that there is a focus on “selecting information-rich cases whose study will illuminate the questions under study” (p. 230). In this first of a series of studies it was important to ensure the participation of schools that represented maximum variation in terms of primary school characteristics. Patton describes this as a “maximum variation” approach to purposeful (sometimes referred to as purposive or judgment) sampling. Approaching the consortium schools was intended to ensure that the initial sampling frame comprised schools with a range of characteristics in terms of school size, location, decile and type. This strategy was intended to ensure that schools with a variation in characteristics were potential participants, since characteristics such as school size and, hence, complexity of senior management structure could affect the way that appraisals were implemented in the school. Characteristics of the 17 schools which did agree to participate, and went on to provide documentation, were analysed to ensure that maximum variation was evident in the sample of participants. As Table 5 shows, there was variation in those characteristics, so all of those who volunteered were selected, and it was not necessary to invite additional schools to participate.

#### *Data Sources*

Schools provided a range of performance management related documents including such things as job descriptions, performance agreements, performance management

checklists, guidelines on setting annual objectives, documentation checklists, interview questions, self-appraisal questions, and classroom observation formats. As previously mentioned, this study focused on the two document types that were provided by most schools – policies and performance indicators.

Policies provided by schools were about “appraisal”, “staff appraisal”, “teacher appraisal”, and “performance appraisal and performance management” (with explicit references to appraisal). Despite these variations they were classified collectively as “policies” for the purposes of this study, given that they refer to an equivalent process across schools.

Most schools’ policies included at least one key section that stated the purposes or intent of the appraisal or performance management process. These sections had various titles across the schools including “aims” and “goals” or most commonly, as in the example below from school three [from here schools will be referred to with the abbreviation ‘S3’ according to their numerical reference], “rationale” or “purposes”:

Rationale: performance management involves the development and implementation of policies and procedures to ensure that the teachers and staff provide education and services which effectively meet the needs of their students consistent with the goals and objectives in each school’s charter.

Purpose: Performance appraisal is a continuous and systematic process intended to help individual teachers with their professional development and career planning. Encouraging work performance is critical to the effectiveness of the school and leads to greater work satisfaction for staff. It includes identifying desired work performance and objectives; developing and evaluating work performance; identifying needs and removing constraints; giving praise, feedback and encouragement and recognising achievement.  
[S3]

This study focused on the intention statement section/s of schools’ policies, such as those above (regardless of whether they were referred to as rationale, purpose, aims

or goals) because they were common to most policies, and they provided an opportunity to discover whether there was an intention to focus on student learning.

Policies were provided by all but one of the schools in the final sample, while 15 schools provided performance indicator documents. Only the indicators pertaining to fully registered teachers were used, as this group represents the core of the teaching body in New Zealand. The first five dimensions of the professional standards were analysed – professional knowledge, teaching techniques, motivation of students, classroom management and communication (when related to students). These dimensions were selected because they relate more explicitly to classroom practices and student learning, which are relevant to this study, unlike the dimensions concerning colleagues and wider school activities.

Performance indicators (for the professional standards) provided by schools were referred to in a number of ways, including “indicators”, “performance indicators”, “professional standards’ indicators”, and “performance management indicators”. Despite these variations they were classified collectively as “indicators” for the purposes of this study, given that they refer to the same aspect of the appraisal requirements. They will be referred to as indicators or performance indicators from here on.

#### *Analysis - Coding of Policy Documents*

References in policy intention statements were categorised and relative frequency of each type of statement was calculated. Three categories were used – “student learning”, “teaching” and “other”.

The “student learning” category included references to evaluating, responding to and impacting positively on learning. The “teaching” category was also included, in recognition that some intentions or performance indicators about teaching may be based on a definition of effective teaching that deems student learning as central. For example, those who subscribe to a view of effective teaching like Berliner’s,

“effectiveness means simply that teachers get most of their students to learn most of what they are supposed to learn” (1987, p. 94) may refer broadly to “teaching” without making explicit the importance of student learning having occurred for teaching to be deemed effective. Intention statements that comprised references to neither the student learning nor teaching categories were coded as “Other”. Table 6 outlines examples of how references in intention statements were coded to the three categories. The coding rules outlined in Table 6 are important to ensure “a self-conscious procedure with clear discussion rules” (Miles & Huberman, 1994, p. 12). In a number of instances statements reflected both the student learning and the teaching category and so were coded in both. An example of a double coded statement, “affirm and improve the quality of teaching and learning outcomes” can also be seen in Table 7.

Table 6 Policy Intention Statement Coding Rules

Category	References to	Examples of policy intention statements To...
Student Learning	Student learning	attest to the achievement of effective teaching and learning [S11]
	Student success	
	Student achievement	achieve improved outcomes for student [S6]
	Student outcomes	
	Learning	promote ongoing and improved teaching and learning [S6]
	Success	
Achievement	affirm and improve the quality of teaching and learning outcomes [S8]	
Teaching	Teaching	improve and extend professional skills [S13]
	Teaching effectiveness	promote the development of effective performance [S4]
	Professionalism	
	Work	support the establishment of effective teaching programmes [S7]
	Effectiveness	
	Standards	affirm and improve the quality of teaching and learning outcomes [S8]
	Excellence	
	Achievement	
	Skills/abilities	
	Performance	
	Work performance	
	Teaching performance	
	Other	Other than above
create a climate for two-way communication [S15]		

One school had an atypical policy format, with no specific intentions section, and a less easily definable number of intention statements. This school's data was not taken into account in the calculation of percentages.

*Inter-rater reliability*

For the studies reported here, and in Chapters 3 and 4, a second coder was asked to analyse a portion of the data to establish the extent of inter-coder agreement. In all

cases agreement was high and strong validity established. The second coder was an experienced teacher and teacher educator, involved in a range of educational contexts over the last 15 years. She had a thorough understanding of the educational appraisal context, and I trained her in the categories by working through a number of real examples from the studies involved and hypothetical examples of data.

For this study, the unit of analysis was a policy intention statement. The second coder was given a printout of the complete dataset of intention statements and asked to colour code each statement to indicate the category it related to.

Agreement for the first two categories was calculated by counting the number of statements she coded to each category that I had also coded to that category. For instance, in the student-learning category, I identified 19 intention statements while the second coder initially identified 14 intention statements. There was, then, initial reliability of 84 per cent in the student-learning category. The disagreements related to statements about students needs being met. I had coded these to the student-learning category while the second coder had not. After discussion we reached complete agreement (100 per cent inter-rater reliability) in relation to the student-learning category. With regard to the teaching category, the second coder agreed with 68 of the 76 statements coded to this category (89 per cent inter-rater reliability). Disagreements related to references to achieving school goals and personal achievements, which I did code to the teaching category, while the second coder did not.

The overall agreement for analysis of intention statements referring to student learning or teaching was 92 per cent since the second coder agreed with 87 of the 95 coding decisions made by the first coder.

#### *Analysis - Coding of Performance Indicators*

Indicators developed by schools, in relation to the five professional standards' dimensions analysed in this study, were also analysed and relative frequency of the

indicator types established. The three categories were “direct inquiry into student learning”, “indirect inquiry into student learning” and “other”. Examples of indicators coded to these categories are given in Table 7. Indicators coded to the first category, direct inquiry into student learning, demonstrated a focus on actual student learning or outcomes for students. These kinds of indicators suggested that appraisers promoted inquiry into student learning.

The second category, “indirect inquiry into student learning”, was for indicators that expressed an assumed causal relationship between certain teacher processes and student learning. These indicators focused on processes rather than outcomes and were stated in a way that suggested an assumption about the positive impact particular processes (teacher practices, strategies, approaches or behaviours) would have on students’ learning. An example of indirect inquiry into student learning is the indicator “demonstrates an attractive, busy and challenging physical environment that promotes student achievement and further learning across the curriculum” [S11]. This indicator would primarily initiate inquiry into a teacher’s classroom environment. A classroom environment being attractive, busy, or challenging, however, will not necessarily relate to student learning. Those things could, in fact, impact negatively on student learning for those students who prefer a plain, calm and quiet versus attractive and busy environment. Indicators such as this rely on assumptions and prevent the testing of those assumptions through inquiry into student learning directly.

Thirdly, indicators that were not related to student learning, either directly or indirectly, were coded as “other”. Details of coding rules, with examples, are outlined in Table 3.



Table 7 Performance Indicators Coding Rules

Category	Indicators	Examples
Direct Inquiry into Student Learning	about actual:	Demonstrates appropriate emphasis and successful learning in the areas of reading, writing and mathematics [S11]
	- Learning	
	- Success	Children are achieving success [S13]
	- Progress	
	- Outcomes	Achieves results
	- Achievement	
- Results		
Indirect Inquiry into Student Learning	that express an assumption that a particular practice will:	Adapts own teaching approaches and techniques to maximise students' learning opportunities and achievements [S14]
	- lead to student learning	Encourages pupil progress by providing feedback [S4]
	- cause student learning	
	- be related to student learning	Demonstrates an attractive, busy and challenging physical environment that promotes student achievement and further learning across the curriculum [11]
	- promote student learning	
Other	Neither of the first two categories	Work dated in exercise books [S1] Cloakroom area is tidy [S5] Professional dress standards [S8]

*Inter-rater reliability*

Coding of the performance indicators by the second coder revealed an extremely high degree of agreement for the judgements of the 503 indicators – more than 99% agreement. The three indicators for which there was disagreement were ones I had coded to the “other” category, which the second coder noticed should have been coded to the “indirect inquiry into student learning” category. They were “regular verbal and written feedback given to students to encourage progress” [S8], “indicates

a range and depth of learning experiences to meet the intended learning outcomes” [S11] and “learning centres that encourage and challenge independent learning are evident and updated in at least three curriculum areas” [S11].

This error affected the results by just half of a per cent for the two affected categories.

## *Results*

The first set of results reported below are those relating to the statements in the intention sections of teacher appraisal policies. They are followed by results from analysis of the performance indicators.

### *References to Student Learning in Policy Intention Statements*

There were 119 intention statements in total, 19 of which referred to student learning (see Table 8). Those 19 statements were from 12 of the 17 schools and reflected a range of intended impacts, the majority of which (11) were to improve learning or learning outcomes for students. Typical examples were “to improve the quality of student learning through classroom observation and analysis of classroom processes” [S13] and “to enable staff to improve personal effectiveness and thereby enhance the educational outcomes of pupils” [S5]. Purpose statements in policies such as “to encourage more effective teaching and learning” [S10] and “to facilitate better pupil learning by improving classroom practice” [S11] suggest an aspiration for a formative and developmental approach to appraisal.

Other intended impacts, more summative in nature than improvement focused, were for appraisal to “ensure”, “affirm”, “attest to” or “evaluate” learning. One example of a summative intention statement was to “attest to the achievement of effective teaching and learning” [S11].

Two schools referred to students in terms of their being provided with teaching – “providing students with effective teaching” [S7] and “to ensure teachers and staff

provide education and services which effectively meet the needs of their students” [S8]. While these are more explicitly about teaching than student learning, these statements do indicate a concern for students in the context of appraisal. A number of statements, such as “to improve the quality of teaching and learning” [S17] were coded to both the teaching and learning categories.

Table 8 School Appraisal Policy Intention Statements

Intention Statement Category	Statements		Schools	
	Number (n=119)	Percentage	Number (n=17)	Percentage
Reference to Student Learning	19	15%	12	70%
Reference to Teaching	76	70%	17	100%
Reference only to Other (than Student Learning or Teaching)	42	35%	12	70%

Analysis by school (see Figure 9) shows that the majority of policies, from 12 of the 17 schools, had at least one reference to student learning, and six schools had more than one reference to student learning. All schools made reference in their intentions to the ‘teaching’ category, including the five schools that did not make any references to student learning. These ‘teaching’ category notions of teaching, performance, professional standards, achievements or effectiveness could potentially encompass concern for student learning. For example, statements about leading to greater effectiveness, maintaining high standards of teaching and achieving improved performance are likely to encompass an intention to make an impact on student learning.

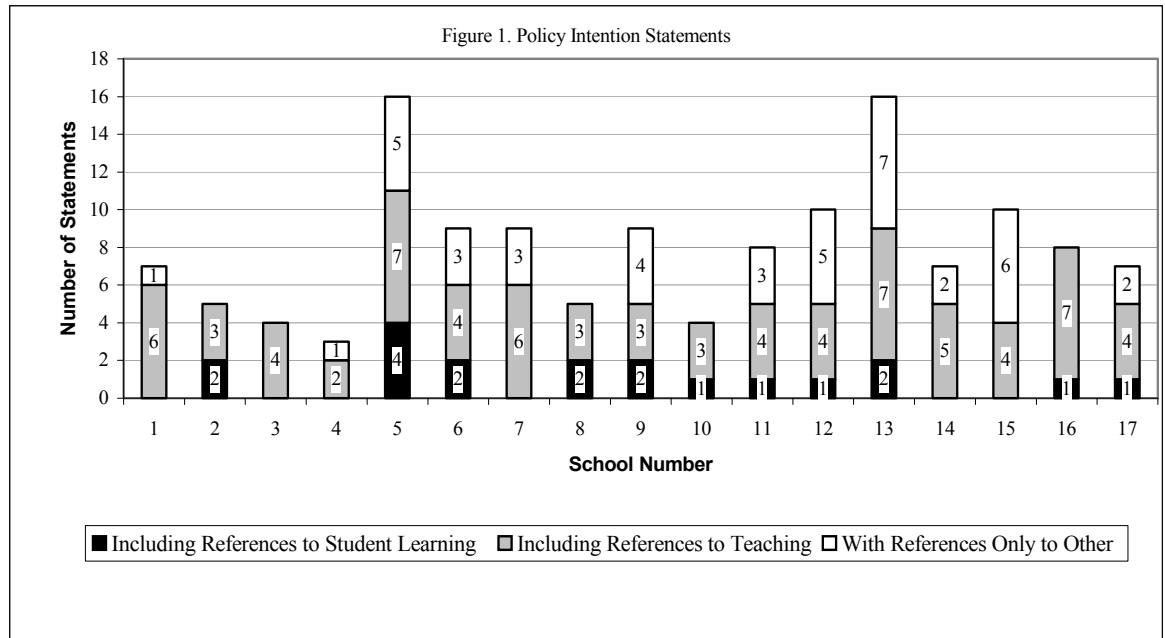


Figure 9 Types of Intention Statements in Schools' Appraisal Policies

*Performance Indicators About Student Learning*

Indicators for the professional standards, which play an important role in appraisal observations and discussions, did not reflect schools' explicit (albeit limited) intentions in their policies to improve student learning. This was the case despite more than half of the schools having between 40 and 110 indicators for all dimensions of the professional standards (see Figure 10). On average, schools had 46 indicators.

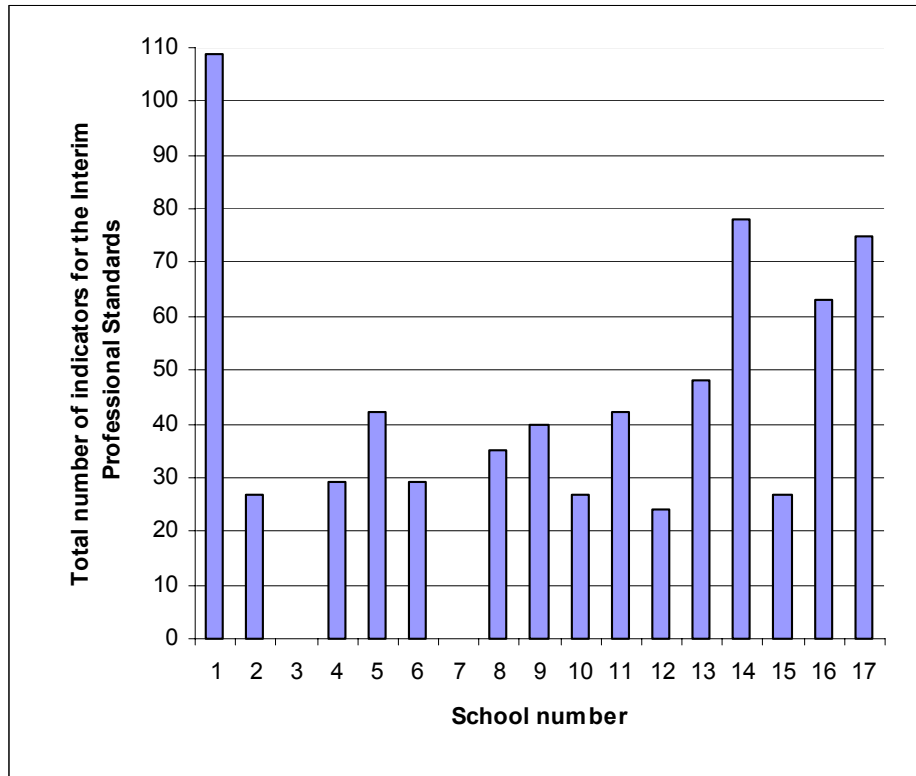


Figure 10 Total Number of Professional Standards' Indicators

The extent of focus on student learning in the performance indicators was extremely limited. Only six schools had indicators that directly encouraged inquiry into student learning, and one of those did not actually indicate an intention to do so. The indicator “children are achieving success” [S13] is an example of one that encouraged further inquiry into student learning. An indicator of this type raises questions, for instance, about which children are succeeding, in which curriculum areas, and to what extent. The six schools that did inquire directly into student learning did so for between 2.4% and 12.5% of their indicators. This is noticeably lower than the percentage range of student learning related references (for the 12 schools that had them) in intention statements, which was between 10% and 66.7%.

Table 9 demonstrates that only three percent of indicators, for the five professional standards dimensions analysed, were directly concerned with student learning. Even if direct and indirect indicators of student learning were combined, their total of six

percent does not suggest that appraisal provides teachers with enough opportunity to reflect on their practice in terms of outcomes for students. There are a significantly greater number of indicators encouraging inquiry into things other than student learning. As expected, from an analysis of the guidelines for incorporating the professional standards into performance management (Ministry of Education, 1998c), the 'other' indicators reflected the professional standards dimensions of professional knowledge, teaching techniques, motivation of students, classroom management and communication. The analysis also revealed a predominant focus on teacher processes related to each of these. For example, the indicator "checks to ensure that students understand the requirements of an instructional activity" [S1]) focuses on what the teacher does rather than on the outcome for students of demonstrating understanding of task requirements.

Table 9 Percentage of Performance Indicators by Category

	Indicators (n=503)	
	Total	%
Direct Inquiry Into Student Learning	15	3%
Indirect Inquiry Into Student Learning	16	3%
Inquiry into Other	472	93%

A similar number of indicators were about student learning directly as were about student learning indirectly. Only six of the fifteen schools (40%) had indicators directly about student learning. Eight of the fifteen schools (53%) had indicators that assumed an indirect or causal link between particular teaching behaviours, and progress, achievement, success or learning. Examples of these are listed below.

"encourage pupil progress by providing feedback, identifying difficulties and encouraging self evaluation" [S4]

“demonstrates an attractive, busy and challenging physical environment that promotes student achievement and further learning across the curriculum” [S11]

“adapts own teaching approaches and techniques to maximise students’ learning opportunities and achievements” [S14]

“focus on individual student’s needs so that they know specific goals and experience success” [S6]

While these indicators reflect a concern for student learning, their phrasing takes for granted that learning, success and achievement will result from particular teacher practices, such as providing feedback, providing a particular type of environment, adapting approaches and focusing on individuals. This finding is significant since it represents a considerable contrast to the appraisal approach proposed in Chapter 2 in which learning is not taken for granted and is explicitly attended to.

Two thirds of the schools that included indicators directly inquiring into student learning also had indicators that supported indirect inquiry. Even when schools had indicators both directly and indirectly about student learning, there were, in all instances, a greater number of indirect indicators. One third of all the schools, though, had no indicators inquiring into student learning, either directly or indirectly.

#### *Lack of Alignment Between Intentions and Professional Standards’ Indicators*

As Figure 11 shows, there was significant discrepancy between the two document types analysed in this study. What schools espoused as important to the purpose of appraisal, and what the performance indicators suggested they focused on were not aligned.

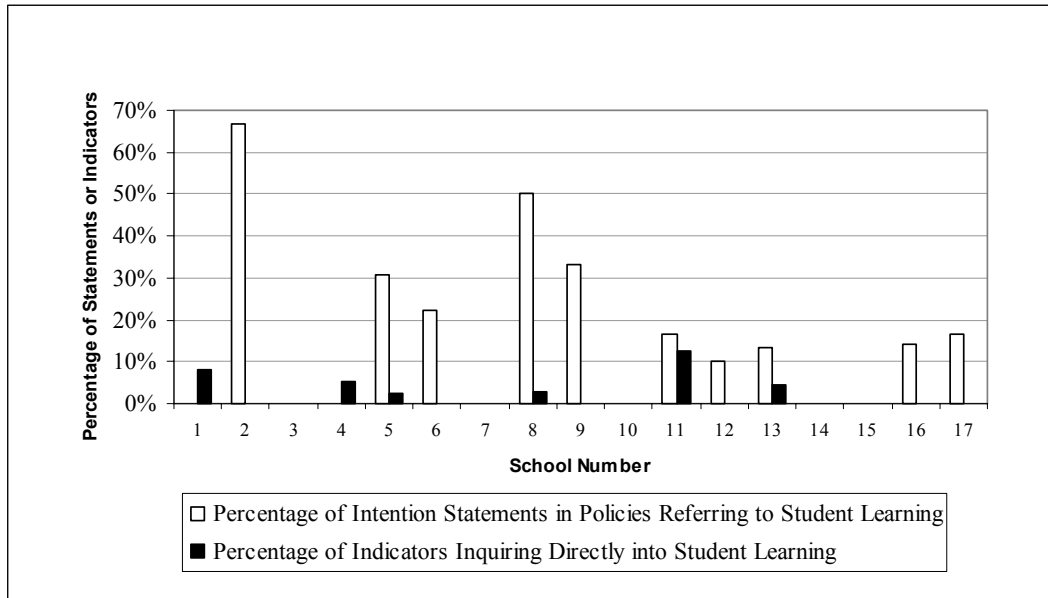


Figure 11 Comparison of Percentages of Intention Statements and Professional Standards' Indicators Relating to Student Learning

Seven schools that intended to evaluate or improve student learning through appraisal, had no performance indicators directly related to that purpose (S2, S6, S9, S10, S12, S14, S17). One of these, a low decile school with one of the smallest roll sizes in the sample, had the greatest emphasis on student learning in their policy intention statements, and yet no indicators to reflect those intentions. This school’s policy included the following:

“Performance appraisal will increase self-awareness of what is expected of one, and lead to greater effectiveness. This school recognizes the need to evaluate regularly the quality of teaching and learning at all levels, relative to agreed standards of teaching performance and to the school’s goals and objectives as expressed in the charter. To achieve this, a system of staff performance appraisal is carried out with the aim of improving the quality of teaching and learning outcomes achieved by the school and its staff by providing support and development opportunities that will enable them to achieve their personal and professional goals.” [S2]

While this school policy advocated the importance of learning, outcomes and achievement of goals, their 20 indicators were not about these at all. Rather, they



were primarily about teacher processes such as teachers' use of a variety of strategies and a range of resources. There was an overall lack of alignment between espoused intentions and the practices suggested by professional standards' indicators.

Four schools (S5, S8, S11, S13) had a greater degree of alignment between intentions and indicators. Their intentions included references to student learning, and they also had between one and four indicators to facilitate that intention being met. In contrast, one school did not have any student learning related policy intentions, but had six indicators that inquired into student learning, the most of the entire sample [S1]. Overwhelmingly, though, the findings reveal no relationship between schools' intention to focus on students as espoused in school appraisal policies, and the indicators that guide appraisers in their work.

### *Discussion*

In this section I will discuss what an analysis of schools' policy intention statements reveals about the schools endeavouring to improve student learning through appraisal processes. I will then consider the role indicators, developed for the Interim Professional Standards, play in promoting or limiting the extent to which policy objectives are met. These will be discussed in terms of organisational alignment, bureaucratic versus professional approaches to appraisal and organisational culture.

#### *'Best of Intentions' in School Policies*

The purposes of teacher appraisal outlined in schools' policy statements do reflect, in most cases, a focus on improving student learning. This is a critical aspect of the argument put forward in Chapter 2 and a pre-requisite for teacher appraisal that includes data-based inquiry into student learning. In all cases a purpose of improving teaching was evident, which could, potentially, be construed to include student learning, though not explicitly. School policy statements are then aligned to

the policy objective stated in the national guidelines “to improve learning outcomes for students by improving the quality of teaching and leadership” (Ministry of Education, 1998c, p.3). Having established that schools generally expect their appraisal systems to contribute to improvements in student learning, I analysed the indicators the schools developed for the professional standards to see whether the indicators were about student learning.

*Indicators Reflecting Assumptions about Impacts Practices Have on Learning*

Analysis of the schools’ indicators suggests that too great a reliance is placed on assumptions about the impact practices have on student learning. There were more indicators in total in this study that related to ‘indirect inquiry into student learning’ (16) than ‘direct inquiry into student learning’ (15). Also, eight schools had at least one indicator relating to the ‘indirect’ category as opposed to only six in the ‘direct’ category. My preference for indicators that directly inquire into student learning will be explained through the following example indicator – “adapts own teaching approaches and techniques to maximise students’ learning opportunities and achievements”. This indicator would primarily encourage inquiry into a teacher’s adaptation of their teaching approaches and techniques. Such adaptation, however, is not necessarily related to student learning. Frequent changes to teaching approaches could, in fact, impact negatively on student learning, by causing inconsistency, lack of routines and excessive change. These indicators rely on assumptions and prevent the testing of those assumptions through direct inquiry into student learning. Alternatively, the focus of the indicator could be placed on student learning, encouraging teachers to check that learning improvement actually results from the adaptation or tailoring of their practice to meet students’ needs.

A learner-centred focus, such as that evident in a recent teacher assessment approach (PACES) in the US, would encourage educators to be concerned “not simply [with] the quality of the teacher’s performance, or the demonstration of selected teacher behaviours that might be counted and recorded” (Ellett & Teddlie, 2003, p. 110). The indicators analysed in this study, though, tend towards the prescription of particular

teacher processes, which Ellett and Teddlie warn fails to challenge teachers to consider the “uniqueness of their particular teaching and learning environments to find the best ways to actively engage their students in learning tasks”, in the way that some emerging models of appraisal do (2003, p. 110).

*Excessive Indicators as a Cause of Policy Implementation Slippage*

Procedural guidelines in schools’ policy documents suggest that appraisers, on behalf of the Board of Trustees, typically attest that teachers have met the professional standards by assessing against the school’s indicators. The average number of indicators across the first seven professional standards dimensions in the schools studied was 46. It is problematic and difficult, I would contend, to give sufficient attention to anywhere near this number of indicators within the parameters of a typical appraisal cycle (such as the one in Figure 12).

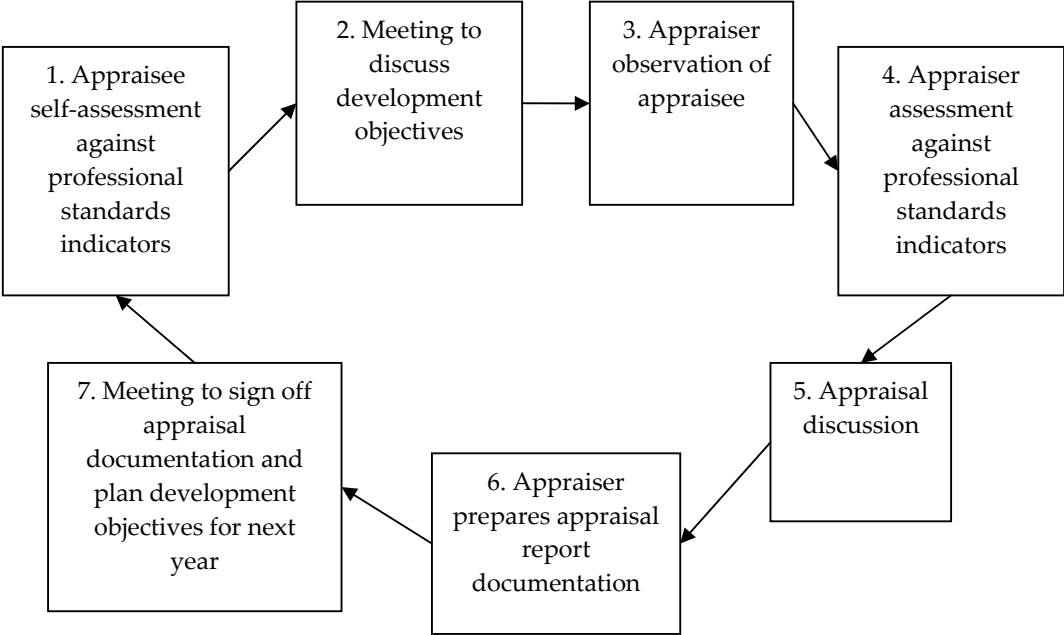


Figure 12 Example of a Typical Appraisal cycle

There are at least three possible consequences of having lengthy lists of indicators that are then used as a basis for appraisal. The first is that indicators may only be included if they are able to be assessed in a time-efficient manner through a checklist

approach. This approach favours surface indicators aimed at assessing compliance aspects of practice (such as completion of tasks) rather than the more time-consuming and complex affair of engaging in dialogue about impacts on student learning.

A second possible consequence is that indicators may be addressed in a cursory manner despite their representing important aspects of the professional standard of a teacher’s work. Figure 13 reveals the number of minutes available to engage with each indicator in a hypothetical context in which four hours of appraiser/appraisee contact time is available over an appraisal cycle. The basis for using four hours for this example came from reports by practitioners in a subsequent study, reported in Chapter 4, of appraisal discussions occurring between two and four times a year for no more than an hour each time. Four hours is, then, an optimistic scenario in which to consider the time available to engage with indicators.

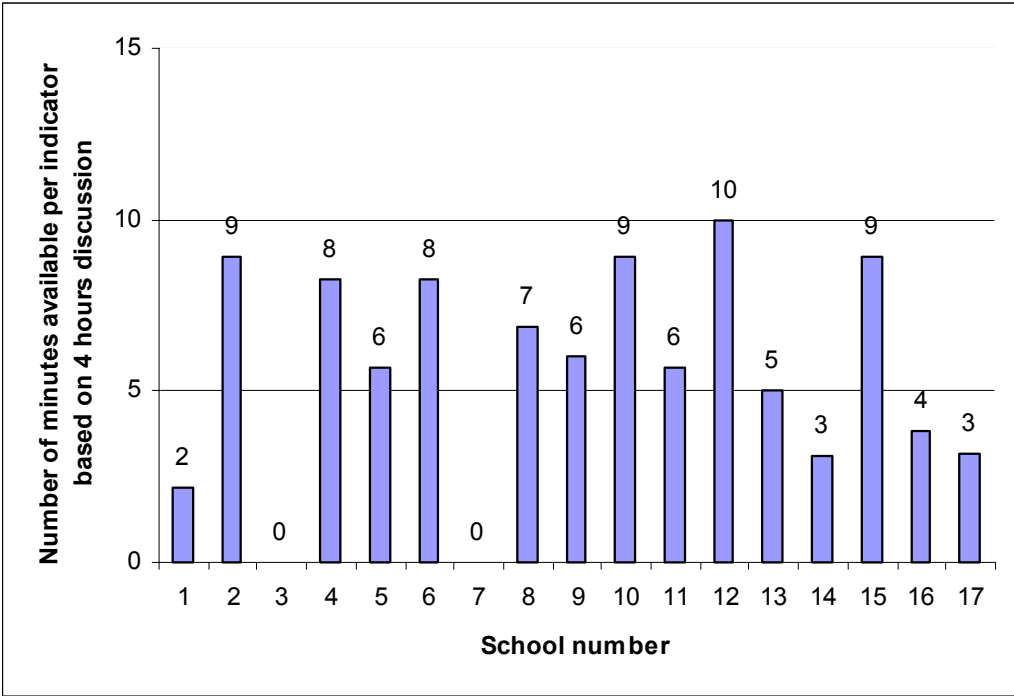


Figure 13 Number of Minutes per Indicator Based on Four Hours of Discussion Time

Figure 13 shows that opportunity for in-depth consideration of indicators is limited. In this hypothetical example (since schools may, in practice, have more or less than four hours available) a teacher and his/her appraiser meeting for four hours over the

course of an appraisal cycle could allocate an average of just five minutes time to attend to each indicator. It is likely, then, that a coverage approach will be taken in which only surface attention can be given to even the most important indicators included within the set.

Even those schools that do include indicators that encourage consideration of impacts on student learning, and six of the 15 schools did, can only give them minimal attention. Consider here one school's indicator – “demonstrates appropriate emphasis and successful learning in the areas of reading, writing and mathematics, consistent with the school's priority” [S11]. This indicator was just one of 42 in that particular school's list. Based on a hypothetical four-hour time frame for discussion over the appraisal cycle, each of this school's indicators would receive, on average, less than six minutes to consider each one. I question whether rich and revealing dialogue could occur, in that time frame, to address such an important matter as students' reading, writing and mathematics learning. While time allocations would not, in practice, be divided up in this way on a 'per-indicator' basis, the analysis does reveal the challenge of meaningfully addressing so many indicators in a time-constrained appraisal process.

A third, and most likely consequence, given the number of indicators, is that they will be afforded status in a hierarchy of importance. A prioritising approach such as this opens the process up to subjective decisions about the relative significance of particular indicators and which should be given most attention. Those indicators worthy of the least attention may ultimately receive the most attention in this approach. The most appropriate basis for establishing that hierarchy of indicators, I argue, would be the extent to which they make connections between student learning in relation to teachers' practice. The lists of indicators, provided by schools in this study, did not make any such hierarchy apparent. It was not clear that indicators like “evidence of progress in written work” [S1] should receive any more attention or priority than those along the lines of “work displayed neatly/attractively” [S1].

To avoid the potentially negative consequences outlined above, I would suggest a 'less is more' approach should be taken to indicators. When sets of indicators are extensive, numbering between 60 and 109 as in the case of a quarter of the schools in this study, the most important of those are likely to be obscured, given superficial attention, or missed entirely. Rather, indicators should deal with the goals and purposes most fundamental to the teaching profession, and the education system as a whole – that is, raising student achievement.

*Lack of Alignment Between the Policies and the Indicators*

The issue of the scope of indicators is coupled with a further issue to do with findings about the substance of the indicators. This study has revealed that intentions in policy documents are not necessarily reflected in practices signalled by professional standards' indicators. A closer match between indicators and policy intentions would strengthen organizational alignment. Performance alignment, one aspect of organizational alignment, represents "the degree of similarity between actual organizational behavior and the strategic ideal or vision" (Semler, 1997, p.2). The significance of alignment is that it leads to "high-performing organizations, able to flourish despite their complexity, with strongly linked strategies, structures, and cultures" (Semler, 1997). It would then be desirable for schools that espoused a purpose of appraisal to "promote ongoing and improved teaching and learning", as school six did, to have indicators that are aligned with that purpose. Listed below are a number of possibilities, aligned to appraisal purposes of improving student learning. Indicators could encourage appraisers and appraisees to consider:

- the extent of improvement in their students' learning (quantitative)
- the type of improvement in their students' learning (qualitative)
- the improvement of a particular group of students
- the improvement of a particular individual
- comparisons of improvement of students' improvement with other classes
- comparisons of students' improvement in different learning areas
- comparisons of students' improvement in different strands of a learning area
- lack of improvement in particular learning areas, strands or achievement objectives
- lack of improvement for particular classes, groups of students, or individuals

This approach to aligning schools' policy intentions and indicators reflects the kind of consistency advocated by Tosti (1994) who suggests that there should be consistency between the values that people (and organisations) espouse, and their day to day behaviours. Performance indicators are, mostly, worded as activities and behaviours teachers are expected to engage in. It is these activities and behaviours, Tosti suggests, that are the execution of the intent of the organisation. Tosti's point is that alignment matters. In organizational terms, he says, it matters for improvement in profits, job satisfaction, quality of service, and in long-term organizational health. In educational terms alignment is, then, likely to improve student learning, teacher satisfaction, quality of teaching and the 'health' of the school organisation. While these claims suggest substantial benefits to be gained through organizational alignment, alignment in itself is not sufficient to ensure quality.

This study has revealed a lack of focus given to student learning in schools' indicators that reflects alignment to the national policy documents in which there is also a lack of focus on student learning. In this case, then, alignment has been detrimental to the quality of appraisal systems in schools, if a quality appraisal system is deemed to be one with the kind of focus argued for in Chapter 2. There is a risk, then, of seeking alignment where that which is being aligned to, is deficient in some way. This highlights the need to consider the system wide context for appraisal, and not just to seek improvement at the teacher/classroom or school level. This is particularly important given the number of levels of decision making that occur between the national policy context and the actual implementation of teacher appraisal (see Figure 14).

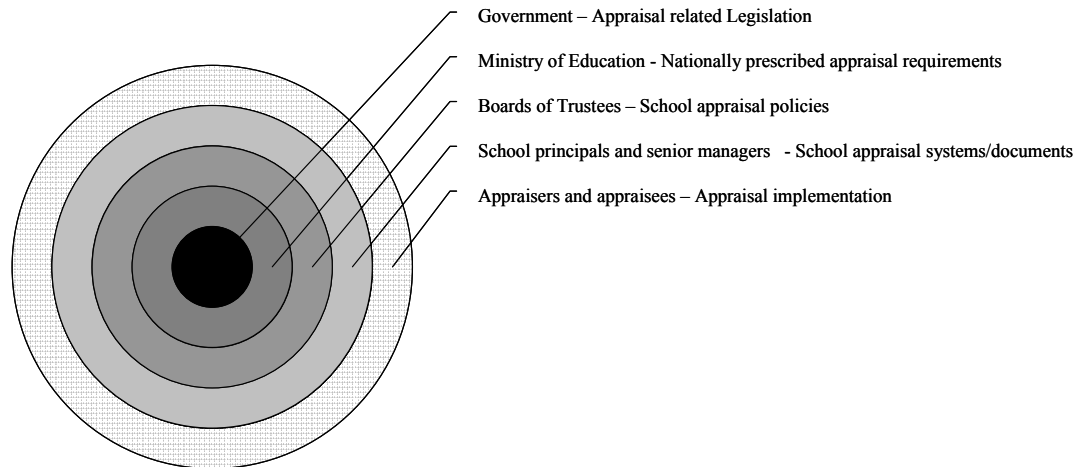


Figure 14 Levels of Appraisal Decision Making

Glass (1999) refers to changes and revisions to, and undermining of, policy as ‘policy mutation’. He highlights the significant role of individuals, and their perceptions, in policy implementation, and reminds that no guarantee can be given that intentions of a policy will be apparent in its implementation. This phenomenon would explain the potential for any given emphasis afforded to student learning in national or school level appraisal policy, to be shifted or weakened as the policy is processed through the levels of the system. In this study, alignment was apparent between the lack of focus on student learning in national appraisal policies and the lack of focus on student learning in performance indicators. Alignment was not apparent, though, with schools’ own policy documents, which did indicate a focus on student learning.

*Bureaucratic and Professional Approaches to Appraisal*

The results of the analysis of professional standards’ indicators in this study could be interpreted as an indication that appraisal is being driven by more bureaucratic/managerial approaches than professional/improvement-focused ones. The connection between performance management systems introduced in New Zealand in the late 1990s and the bureaucratisation of teaching has been well documented. Fitzgerald, Youngs and Grootenboer explain “there is an increasingly bureaucratic system of assessment, evaluation, attestation, review and appraisal that



surveys and controls the professional work of teachers in a variety of ways for a variety of purposes" (2003, p. 94). The increased accountability measures involved in teacher appraisal, particularly since the introduction of the professional standards, have led in particular to concerns about the dual purposes of appraisal (formative and summative). These dual purposes require educators simultaneously to "work in a collaborative, collegial and supportive way [but also] adopt a hierarchical stance to ensure ... an objective and performance-driven performance management system" (Fitzgerald et al., 2003, p. 94).

Indicators of the schools studied appear better suited to a summative than formative process since they focus less on the impact of teaching on students' learning outcomes than on matters other than student learning, such as teachers' practice. The formative aspect of appraisal centres on providing opportunities for teachers to learn. The summative aspect of appraisal centres around using the appraisal system as a tool to ascertain whether teachers have met criteria to be, for example, given promotions, gain increments on the salary scale, or given attestation that they have met professional standards. A formative approach to appraisal would be promoted through indicators concerning student learning, since they would provide the basis for considering the development and improvement of teachers' practice as indicated by evidence about students' learning. Indicators about matters other than student learning, as were predominantly found in this study, suggest a more summative approach to appraisal. These indicators, frequently about what teachers did rather than what their students learnt, tend towards evaluation of a summative nature since they promote inquiry that establishes whether teachers did or did not perform particular practices.

Critiques about conflicting demands and purposes of the two main purposes of appraisal, formative and summative, centre on polarising their characteristics. Negative aspects of bureaucratic, managerial, accountability-driven appraisal are contrasted with appraisal that is professional, formative, and collegial and development oriented. This is apparent in much of the New Zealand literature

(Cardno, 1998, 1999, 2001; Cardno & Piggot-Irvine, 1997; Fitzgerald, 2001; Fitzgerald et al., 2003; Grootenboer, 2000; Middlewood & Cardno, 2001; O'Neill, 1997; Piggot-Irvine, 2000; Townsend, 1995; Youngs & Grootenboer, 2003), and also in relation to Australia (Down, Chadbourne, & Hogan, 2000; Down, Hogan, & Chadbourne, 1997; Hogan, Down, & Chadbourne, 1998; O'Brien & Down, 2002) and England (Gunter, 2002). According to many, these approaches are mutually exclusive and their co-existence is problematic and undesirable. Both attesting to a teacher's competency in relation to performance standards and addressing developmental and improvement needs, can create anxiety and tension for the educators involved.

*Indicators and Bureaucratic Versus Professional Approaches to Appraisal*

The very existence of indicators inevitably reflects at least a degree of bureaucracy and managerialism in appraisal. They do, though, have the potential to serve as a useful formative tool that contributes to the improvement of teaching and learning. Framed appropriately, they can promote deep and authentic inquiry into connections between teachers' practices and their students' learning.

Most of the indicators analysed in this study represent tasks or matters to be complied with to establish competence. The vast majority of them were in the 'other' category of this study (rather than 'direct inquiry into student learning' or 'indirect inquiry into student learning'). One possible explanation for that focus might be the predominance of a clinical supervision model of teacher appraisal in New Zealand primary schools. Clinical supervision is characterised by a pre-conference, observation, post-conference sequence that is well suited to establishing if teachers have complied with indicators such as "furniture arranged appropriately" [S1]. Clinical supervision is, though, less suited as a model through which to improve the practice of teachers whose competence has already been established, and to pay deep attention to the impacts teachers have on student learning. The characteristics of clinical supervision, as outlined by Hannay and Telford below, may serve to limit the scope of indicators developed to assess through such a model of appraisal. They suggest that clinical supervision is:

...useful in the supervision of teachers new to the profession or teachers who exhibit specific weaknesses in their practice,...[and] allows a teacher and supervisor to focus on identifying and correcting particular areas of practice requiring attention. It also allows teachers to demonstrate that they are competent in their classroom practice. For teachers who are past their initial induction period and have proven their competence, continued use of clinical supervision is of minimal value. For these teachers, there is no need to continually demonstrate that they are competent and there are no critical weaknesses that require correction. Rather, supervision for experienced, competent teachers needs to focus on ongoing professional growth and refinement of practice. (Hannay & Telford, 2003, p. 122)

They go on to argue for a collaborative approach to teacher appraisal that fosters professional growth and development. Indicators to serve this purpose would need to be centred on improvement, and on developing expertise. The predominance of indicators in the 'other' category of this study (rather than 'direct inquiry into student learning' or 'indirect inquiry into student learning') highlights the competence focus of many indicators currently. Many of the indicators in the 'other' category, such as the furniture arrangement example given above, have very little potential to promote inquiry that improves practice. The likelihood of improved student learning resulting from teachers improving their 'furniture arranging' is also extremely minimal. Indicators, while meeting accountability obligations, must also encourage a professional approach in which impact on student learning is central. This may require schools to rethink not only the nature of the indicators they use, but also the model of appraisal within which those indicators are used.

#### *Threat-Free Indicators*

More than twice as many indicators in this study were coded to the 'other' category than to the 'student learning' category. One conceptualisation of the significance of this finding is that it may indicate a tendency for schools, probably not purposely, to develop indicators that are threat-free. It is well documented that many appraisal participants find aspects of teacher appraisal threatening (Cardno, 1994, 1998; Stake,

1989; Townsend, 1995). One response to such threat is to employ what Argyris terms an “organizational defense”. An organizational defense is explained as:

...a policy, practice, or action that prevents the participants (at any level of any organization) from experiencing embarrassment or threat, and at the same time, prevents them from discovering the causes of the embarrassment or threat. Organizational defenses are therefore anti-learning and overprotective. (Argyris, 1999, p. xiii)

Evidence, in this study, of a lack of inquiry into student learning in teacher appraisal, signals a significant degree of avoidance of student learning focused appraisal inquiry. I would suggest that the student learning focus is avoided, and not merely absent, since improvement of student learning is so explicitly stated at least once in most schools’ appraisal policies. For a focus on student learning to be present in a schools’ policy, but absent in their indicators, suggests avoidance that may be either intentional or unintentional. By not inquiring directly about student learning, appraisers and teachers prevent the unearthing of potentially embarrassing data about student learning and, hence, teaching practice. By disregarding student learning, in this context, appraisers avoid having the difficult and time-consuming task of giving negative feedback, and they avoid hurting teachers in the process. Avoidance also prevents the unearthing of pleasing data. Not having student learning focused indicators reduces the opportunity for inquiry into student learning that might reveal positive student outcomes, and lead to recognition of teachers’ effectiveness. Avoiding threat, then, limits learning opportunities for teachers that would arise through the surfacing of both negative and positive data about student learning. While avoidance addresses a desire to manage appraisal in a non-threatening way, it does not address the need for teachers’ practice and students’ learning to be improved through appraisal.

An organisational defence, of not having student-learning indicators, may not be intentional, but the anti-learning and protective consequences of the omission (described by Argyris above) remain. Consequences are highlighted here through

discussion of one school's indicators. Four out of 42 indicators were in a section headed 'pupil's work'. They were:

"marking of work is current"

"setting out as per syndicate guidelines"

"high standard of care and presentation"

"feedback is constructive" [S5]

The limited nature of these indicators, in a section signalled to be about pupil's work, is striking. The marking might be up to date, but what did the marking reveal about the learning, and what are the implications for the teacher's practice? Students' books may be impressive in their adherence to the use of lines, margins, and syndicate-wide prescriptions, but what, and how much, did students learn? Care and presentation may be of a high standard, but what about the level of success on important learning outcomes about things other than presentation? Feedback might be constructive, but was that feedback given in relation to the learning intentions, or was it about the presentation, surface features of writing, quantity and effort? This critique is intended to signal how well-intentioned indicators can steer the focus of appraisal to matters other than those that policies suggest should be prioritised. Other indicators from that school (not in the 'student work' category) were even less concerned with student learning. Indicators such as "register cover is completed", "cloakroom area is tidy", and "class timetable is on the wall" [S5] are apparently entirely driven by a compliance focus. They may be necessary tasks, but do not contribute to the improvement of student learning to an extent that deserves them to be a focus of time and energy in appraisal. They provide a means of simply assessing teachers, in a way that any potential anxiety or threat is avoided, and mandated requirements are met. Indicators that focus on inquiry into student learning may be perceived as difficult and complex, but they should be confronted rather than avoided since they are so fundamental to, and necessary for, the improvement of practice. This would lead to the kind of process described by

Hannay and Telford that is “authentic and focused on professional growth” (2003, p.121).

*The Potential for Student Learning Focused Indicators*

The findings in relation to the number and type of indicators being used in the schools in this study, suggest a predominance of bureaucratic concerns, over and above professional concerns that emphasise student learning. The case put forward in Chapter 2 for data-based inquiry into student learning is possible to implement, given appropriate school culture and, in particular, an inquiry culture, and by aligning indicators with that culture.

*Organisational culture*

Organisational culture (or school culture) is a concept relevant to teacher appraisal since the appraisal process involves many rituals and traditions, in which norms and values are manifested. Peterson and Deal define organisational culture as:

a powerful web of rituals and traditions, norms, and values that affects every corner of school life. School culture influences what people pay attention to (focus), how they identify with the school (commitment), how hard they work (motivation), and the degree to which they achieve their goals (productivity). (2002, p. 10)

The rituals, traditions, norms of values of the school culture, and particularly the culture around appraisal, will influence what teachers pay attention to during their appraisal and at other times. That culture will also influence how committed, motivated and productive they are in relation to the achievement of student learning goals. The approach to and focus of appraisal, as evident in policies and documents outlining indicators, also impacts over time on the nature of a school’s culture.

According to Stoll:

Culture describes how things are and acts as a screen or lens through which the world is viewed. In essence, it defines reality for those within a social organisation, giving them support and identity and creates a framework for organizational learning. (2000, p. 9)

She goes on to explain the significance of the things that are given emphasis in a school – “whether religion or spirituality, pupils’ learning, sporting achievement, or discipline are emphasised in assemblies, provides a lens on one facet of school culture”. It would follow, from this assembly example, that the central concern of teacher appraisal, be it student learning or other matters, also provides a lens on a facet of school culture (Stoll, 2000, p. 10).

*Professional standards’ indicators and school culture*

Indicators, such as those analysed in this study, are particularly influenced by school culture. These documents, since they are used year after year, also transmit and perpetuate particular practices that over time contribute to school culture. According to Evans, “there are three levels of school culture: artifacts and creations, values, and basic assumptions” (Evans, 1996, p. 42). Artefacts and creations are the most tangible. Values are more complex and basic assumptions are the most unconscious and implicit. Indicators for the professional standards, since they are generally documented, provide visible evidence of school culture and represent the first of the three levels outlined above. Their content also provides an insight into the implicit values and assumption that underlie the indicators.

Table 10 Cultural Norms and Assumptions Reflected in the Findings about Indicators

Findings from this study	Possible cultural norms and assumptions these findings reflect
Schools have extensive sets of indicators	We should try to assess everything a little bit
Extremely few indicators are focused on student learning	We accept that learning is a given. We don't need to ask whether students have learnt anything or not
When indicators do include reference to student learning, this is stated in a way that assumes a link between a teacher practice and the learning – the learning itself is not emphasised.	If we check what teachers are 'doing', and we believe that 'doing that' might lead to student learning, then we can avoid asking about the actual learning
Four times as many indicators about teaching than about student learning	Teaching must lead to learning
Twice as many indicators about things 'other' than student learning or teaching, than about student learning	We need to address the complex and wide-ranging aspects of a teacher's practice in appraisal, rather than focusing on student learning
Schools' indicators reflect similar characteristics to the example draft professional standards indicators (that also lacked focus on student learning)	Our appraisal practices should comply with, and be similar to the examples and guidelines outlined in official documents.

The intention of Table 10 is to highlight the significance of cultural norms and assumptions. Simply changing indicators (so that they reflect the qualities of indicators I have argued for) would not be sufficient. The influence of a school's culture means that the very culture of a school would need to shift to reflect a concern for and commitment to focusing on student learning. Such a shift would take time.

"The way we do things around here" is an alternative and simple way of understanding school culture (Bower, 1996) and would include 'the way we develop indicators'. Even what people talk about or avoid talking about is an element of school culture (Deal & Peterson, 1999, p. 2), and so the framework for that talk (indicators) becomes an important aspect of a school's culture. The norms and expectations that indicators outline may or may not include a focus on student learning, but the potential for them to do so certainly exists. There are a number of



ways 'things might be done' around schools, in organisational culture terms, that would be conducive to appraisal focused on inquiring into student learning, and using data to inform the inquiry. In particular, there are a number of approaches to and conceptualisations of 'inquiry' that would be helpful.

*Culture of inquiry*

Schools that have a 'culture of inquiry' will be more open to appraisal that values inquiry into data about student outcomes. A culture of inquiry, once embedded into a school's culture, will be apparent in many aspects of the appraisal process, including the policy statement, procedural guidelines and indicators for the professional standards. There are a number of conceptions of inquiry, with unique but related aspects, that would support the approach to appraisal argued for here.

Mills gives an account about a teacher study group in a professional development school that developed a culture of inquiry. She refers to an 'inquiry stance' being institutionalised over years. Members of the group described how that stance was promoted:

Inquiry became a way of living and learning together once we realized it was a stance we take toward knowledge, learning, learners, schooling, and society. As such, inquiry transcended time and space. It allowed us to create our own culture, to focus on the big things, the things we truly valued. (Mills, 2001, p. 27)

Extensive sets of indicators for the professional standards serve to steer teachers' attention away from the big things that teachers truly value. In Mills' study, it was the opportunity to focus on the bigger picture, including student achievement, which was so important in developing an inquiry stance. Cochran-Smith and Lytle (1999) also use the term 'inquiry stance' to describe an approach in which teachers are impelled to examine their own practice in depth, as well as the effects this practice has on the learners they teach. Where indicators are long documents with many items to assess in one appraisal cycle, and when there is relatively little focus on student learning, it becomes difficult to adopt an inquiry stance, and will result in teachers not seeing the woods for the trees.

Norms of continuous critical inquiry have been described in relation to school improvement (Boyd, 1992), but this notion is clearly applicable also to improvement of teaching through appraisal. Where norms of continuous critical inquiry operate, the strengths and weaknesses of a school, and teachers' practice, are acknowledged. This creates a willingness to deal with imperfections that, in turn, makes teacher learning possible. A focus on teachers as learners, as they are appraised, signals the relevance of notions of 'professional learning communities', 'community of learners' and schools as 'learning organisations'. These will be discussed further in Chapter 6 in relation to school documentation and the wider appraisal context.

*Indicators as a potential positive influence on school culture*

The analysis of indicators in this study revealed a predominant focus on appraiser inquiry into teacher practice. With thoughtful development, indicators could contribute positively to a school culture (and appraisal culture) that values data-based inquiry into student learning. Indicators could promote the three fundamental aspects of the argument for data-based inquiry into student learning in teacher appraisal. They could promote:

- a culture of inquiry
- a focus on student learning
- a focus on data to inform inquiry about student learning

In relation to the first two of these, a culture of inquiry and student learning focus, there are three main modes of inquiry that indicators might advance – (a) teachers' own inquiry into student learning (Figure 15); (b) an appraiser's inquiry into their appraisee's inquiry into student learning (Figure 16); and (c) appraiser and appraisee joint inquiry into student learning and into each other's inquiry (Figure 17).

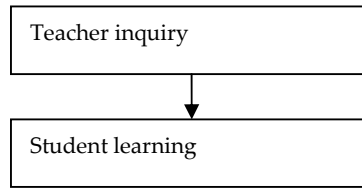


Figure 15 Teachers' Own Inquiry into Student Learning

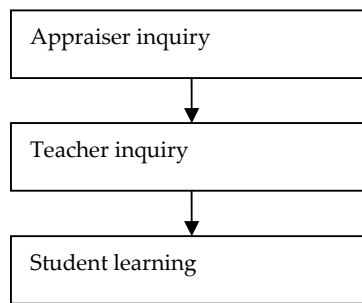


Figure 16 Appraiser's Inquiry into Appraisee's Inquiry into Student Learning

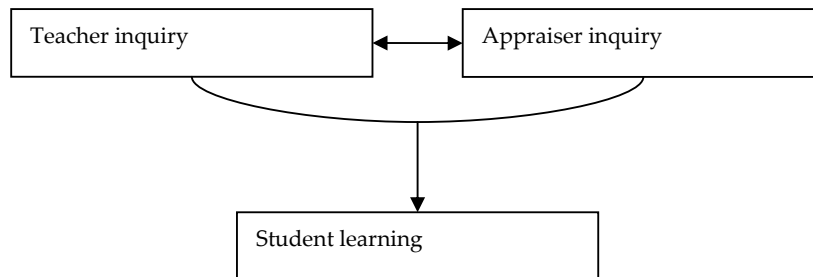


Figure 17 Appraiser and Appraisee Joint Inquiry into Student Learning, and into Each Other's Inquiry

*Limitations of the Study*

There are several limitations of this study that have implications for future research. One is the framework for analysing indicators. The division between indicator categories, in terms of their focus on student learning or other than student learning, could be interpreted as being critical of inquiry into anything other than student

learning. The inclusion of 'other' indicators in appraisal processes is, in fact, considered appropriate, but the relatively minimal inquiry into student learning is not.

Another possible limitation of this study is suggested by the misalignment found between policy statements and documents guiding actual practices. This highlights that limited inferences can be made about practice from policy documents. Policies are unlikely to signal actual practices and would not, in isolation, be an appropriate source of data when considering questions about the reality of practices in schools. The same concern could be applied to the analysis of indicators in this study. The extent to which documents outlining indicators for the professional standards signal actual practices is assumed, but untested. Future investigation, using alternative methods, could delve into a question about the actual significance of these documents, and their indicators, in teacher appraisal. Further studies, reported in the following two chapters, will also inquire into actual appraisal discussions and appraisal goals.

Other questions worthy of further research are also suggested by the results of this first study. The discussion of Study One results proposed that both threat avoidance and school culture are related to the extent to which schools focus on student learning. Subsequent studies will investigate these propositions further and examine actual practices and participant experiences of appraisal

### *Conclusion*

This study has revealed that while improved student learning was espoused as a purpose of appraisal in a number of schools, it was seldom investigated. It has highlighted the limited degree of alignment between the intentions espoused in policies, and the indicators that guide actual practices and discussed these findings in relation to threat avoidance and the organisational culture of schools. I have sought to emphasise the risk of inquiry into student learning being obscured by a

predominance of inquiry into other aspects of teachers' practice, and to draw attention to the significant potential of indicators in determining the nature of appraisal inquiry.

## CHAPTER 4

### Appraisal Discussions – What is inquired into during teacher appraisal?

#### *Introduction*

The purpose of the study reported in this chapter was to investigate the extent to which teacher appraisal discussions, in three primary schools, inquired into student learning. Findings from analyses of schools' appraisal documentation (reported in Chapter 2) revealed that while improved student learning was espoused as a purpose of appraisal, performance indicators, developed in relation to the professional standards, seldom encouraged investigation of student learning. Since the connection between policy and practice cannot be assumed, 11 practitioners were interviewed about their most recent appraisal discussion experience. These interviews provided an insight into whether the concern for student learning evident in policies, but not in performance indicators, may be evident in appraisal discussions themselves.

Appraisal discussions, referred to in many schools as 'appraisal interviews' were chosen as a focus of this study for a number of reasons. Firstly, they are a compulsory element of the appraisal process for every teacher each year. A board of trustees must ensure that each teacher has an opportunity to discuss with their appraiser achievement of their performance expectations and development objective(s) (Ministry of Education, 1997a). Secondly, these discussions are a particularly appropriate context in which data-based inquiry into student learning could occur. A third reason for interviewing about appraisal discussions was to draw on data more closely linked to practice than the previous study (reported in Chapter 3), which analysed appraisal policies and performance indicators.

The purpose of the interviews was to explore the question 'is student learning a focus of inquiry during teacher appraisal discussions?' The following sub-questions were used:

- What topics were included in appraisal discussions?
- Did appraisers and appraisees take the opportunity to inquire into student learning? If so, did appraisers and/or appraisees refer to or use data during the discussion?
- What do practitioners think about the idea of focusing on student learning during an appraisal discussion?

### *Method*

This section will outline the research setting, sources of data, and procedures for gathering and analysing data in relation to the question "Is student learning a focus of inquiry during teacher appraisal interviews?"

#### *The Research Setting*

An ethics application for this study was submitted and approved by the University of Auckland Human Subjects Ethics Committee. I then personally visited six principals in schools that had participated in the previous study and expressed interest in being involved in subsequent research. That interest indicated a likelihood of obtaining rich information from teachers about their experience of appraisal discussions. The approach was similar to that described by Patton as 'purposeful sampling', that is, locating sites that would "yield the most information and have the greatest impact on the development of knowledge" (Patton, 2002, p. 236). The six schools I approached were treated as potential critical cases in a purposeful sampling approach. These cases were critical since if appraisal discussions in these schools did not have that focus (in schools whose teachers had expressed willingness to talk about and share their appraisal experiences), then that focus is unlikely to be occurring more generally.

I discussed with principals from the six schools an information package that included a letter inviting teachers to participate in interviews about their appraisal discussions, a participant information sheet for the school, a participant information sheet for potential interviewees and, for those who chose to participate, a school and interviewee consent form. They were also provided with a summary of key details about the project in question and answer format. Those who gave consent were asked to participate in an interview taking no longer than one hour. The principal was asked to distribute the information sheet to other staff members. Eleven teachers from three schools responded to say they were willing to participate, and hour-long appointments were arranged. Characteristics of the participants' schools are outlined in Table 11.

Table 11 Characteristics of Participating Schools

	School Type	School Authority	School Region	Gender	School roll number (2002)	School Decile
School 1	Contributing	State	Manukau City	Co-Ed	695	10
School 2	Contributing	State	Auckland City	Co-Ed	361	3
School 3	Contributing	State	Auckland City	Co-Ed	358	10

Participants included both appraisers and appraisees, as outlined in Table 12.



Table 12 Interview Participants' Appraisal Roles

School Descriptors	Appraiser	Appraisee	Total
School 1	1	2	3
School 2	3	3	6
School 3	0	2	2
Total	4	7	11

*Data Sources*

Interviews were appropriate for this research question because they provided a means of gathering rich qualitative data relevant to the research question. In interviews participants were able to give detailed descriptions of their appraisal experiences, and were then prompted to recall particular exchanges to add detail to their reports. The semi-structured approach allowed for the interviews to take into account variation in the nature of each participant's appraisal discussion. For example, discussions were held at different stages of the year and of the appraisal process. They were held between different members of staff, and were organised in different ways – for some as a one hour long formal discussion and for others as three short informal discussions. It was important to be able to make adaptations, through the semi-structured interview approach, in order to make the questioning relevant to those different experiences. Semi-structured interviews also provided opportunities for high quality data through the use of prompts, probes and follow-up questions, which encouraged interviewees to clarify or expand their responses (Drever, 1995). Probing during the interview was particularly important in order to elicit the richest descriptive accounts possible and to encourage practitioners to recollect what was discussed and, importantly, what was actually said during their appraisal discussion.

While interviews were an essential source of data, it was important to take into account a number of drawbacks to using interviews to gather evidence (Yin, 1989).

Yin warns that interviews are verbal reports only and are, therefore, subject to a number of problems. These could include bias, poor recall, and poor or inaccurate articulation. To address that limitation interviewees were asked to bring along any written documents pertaining to their appraisal interview. Documents included those self-appraisal documents that were discussed with the appraiser, notes that were taken by the appraisee and/or the appraiser, or formal reports written up as a result of the appraisal interview. All of the participants were willing to share these documents with me during the interview and they were referred to frequently as a means of corroborating the reports of interviewees. The recall limitation was also addressed by requesting participants to refer only to the most recent appraisal discussion in which they had been involved. All of the interviews were within a year of the participant's appraisal discussion, and most were within just a few weeks or months.

I developed semi-structured interview schedules for both appraisers (Appendix A) who were usually school managers, and appraisees (Appendix B) who were all classroom teachers. I attempted to exclude problematic questions by avoiding double-barrelled, two-in-one, restrictive, leading or loaded questions (Anderson, 1990). In all phases of the interviews participants fulfilled a respondent role. In phase one and two they responded to questions relating to their actual experiences. Phase one consisted of open questions to establish what topics were included in their appraisal discussion. Phase two probed questions about the inclusion of a focus on students, their learning and achievement. In the third and final phase of the interview participants were asked to express their views in relation to a suggestion that appraisal interviews could and should focus on inquiring into student learning.

#### *Procedures*

Prior to beginning the interview schedule, I introduced myself, and the goals of the interview. I checked that consent forms were understood and completed, and established if the participants were willing to have the interview tape recorded. I also

took notes as a record of the interview. Participants were reminded that questions would be in relation to their last appraisal discussion, which they were asked to describe in order to set the scene. They were also reminded to refer to their appraisal documentation as necessary throughout the interview. Participants are referred to here according to codes allocated by the researcher to each of them. For instance, [P1] indicates participant number one.

The three phases of interview questions were the basis for each interview. The first phase comprised a set of open questions aimed at eliciting information about the topics and general content of the appraisal discussions. The lead question for this phase was 'what topics were included in appraisal discussions?' Two of the interviews [P6 and P7] were based on 'pre-observation discussions' as opposed to 'post-observation' or 'final' discussions in the other interviews (step 2 as opposed to step 5 of a typical appraisal cycle – see Figure 12, p. 80).

The second phase involved probing questions to find out more detail about what was covered in discussions, why certain topics were discussed and to give participants an opportunity to make clear whether or not student learning was included. The lead questions were 'Did appraisers and appraisees take the opportunity to inquire into student learning?' 'If so, did appraisers and/or appraisees refer to or use data during the discussion?'

In the final 'direct' phase of the interview I shared with the participants an initial analysis of the extent of focus on student learning that I perceived their appraisal discussion to have had. I checked with each participant that my understanding was accurate. This process enabled interview participants themselves to be involved in checking my initial coding of their data, and for them to participate as coding checkers in a reliability process. I then asked 'what do you think about the idea of focusing on student learning during an appraisal discussion?' In most cases participants asked what I meant, or requested I give an example of what this approach to appraisal discussion meant. I gave, where possible, examples that were

connected to their own context and then asked them to respond to the idea. Below is the transcript of a typical example of the kind of scenario given to explain the concept of data-based inquiry into student learning:

How would you feel about an appraisal where data on student learning was the focus? For example, how would you feel if the appraiser took notes on, and talked to you, about how well students achieved those things [pointing to learning intentions on plan]? Which students achieved them? Which didn't? How many achieved them? How many didn't? Why not? How would you feel about those things being focused on in your appraisal discussion?  
[Interviewer]

As well as providing a general type of scenario, such as the one above, I connected the scenario, where possible, to what I knew about the participant's own appraisal/teaching context. For example, one teacher in an earlier phase of the interview raised the issue of differences in appraisal according to her confidence level in the subject the discussion was about. After an initial positive response to the proposed appraisal approach, I was interested to find out if that positive response related only to areas of high self-efficacy. Her concern about appraisal in mathematics, a subject she had lower confidence in than others, was incorporated into a further scenario I gave her about data-based inquiry into student learning:

Now let's say I'm looking at your kids' achievement in maths, and I've taken all this data about who could do it after you taught them that lesson, and who couldn't, what misconceptions there were, and which kids did really well and gained new understanding. How are you feeling now? [Interviewer]

At the conclusion of the interview participants were thanked and their contact details confirmed for communication regarding their interview. All of the participants had agreed to have their interview recorded. The recordings were transcribed, and copies of the transcripts provided to participants in the weeks following to enable them to check and make changes if desired. None of them did so.

The semi-structured approach to interviews in this study, as outlined above, sought to utilise a number of the strengths of interviews outlined by Johnson and Turner. In

particular, interviews were useful, as one of the methods in this research, for measuring attitudes, allowing probing, providing in-depth information, allowing interpretive validity, and as a tool for both exploring and confirming (Johnson & Turner, 2003, p. 308).

#### *Analysis - Coding of Interview Transcripts*

The semi-structured approach to interviews meant that the sequence of questions and responses varied across interviews. The first stage of the analysis involved organising the transcript for each participant on a table according to the phase of the interview the data related to. This ensured that despite the variation in the order that questions were raised, data for all participants were consistently organised ready for analysis. Sections of the transcript that did not relate to questions of the research were removed at this stage.

The second stage of the analysis was to create a column on the table to summarise key concepts (related to the research question) and identify relevant quotes. The research questions dictated from the outset that two important codes be included – ‘student learning’ and ‘data’. All ideas related in some way to these themes were included in the analysis column. An inductive process of forming categories, or clustering, took place (Miles & Huberman, 1994, p. 249). Particulars were subsumed into the general and became the basis for a set of themes that emerged both from the original research question and from the data itself (see Figure 5). The themes outlined in the figure below are examples of themes that relate to both deductive/confirmatory questions (‘student learning’ and ‘data’) and inductive/exploratory questions (the others). Such combination of inductive/exploration and deductive/confirmatory questions is now considered possible in mixed methods research designs (Tashakkori & Teddlie, 2003a).

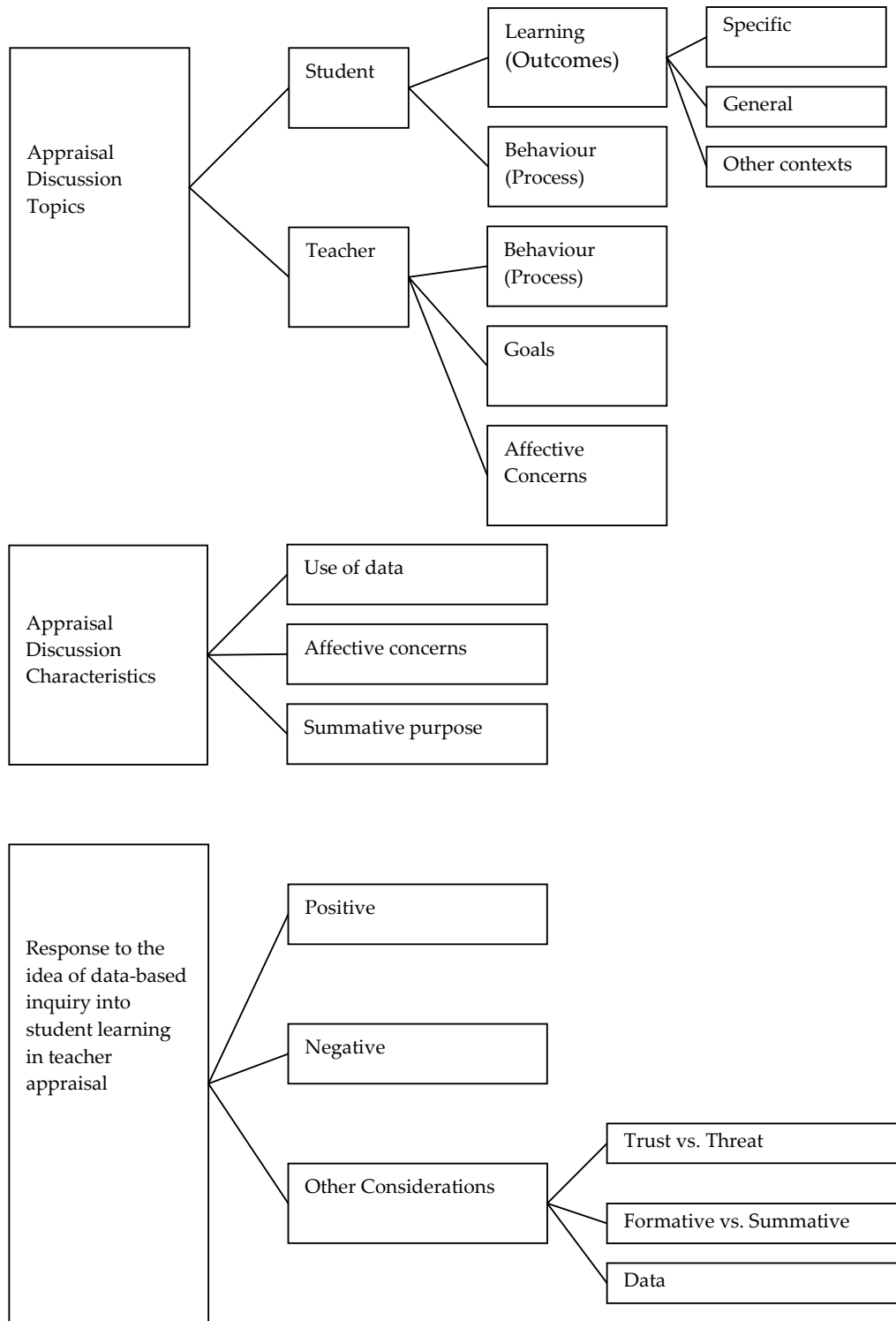


Figure 18 Themes for Teacher Appraisal Discussion Data Analysis

In the third stage of data analysis the transcripts were checked in two ways. I began with a manual approach, in which data from each transcript were scanned, and extracts relevant to the themes were drawn out. I then employed a technical approach using the 'Find' tool on Microsoft Word. Key words likely to relate to each theme were searched in the transcript data for each interview. For example, to make sure that my manual search had elicited all of the extracts relevant to the 'student learning' theme, I searched with the Find tool for 'learning', 'outcome', 'achievement', 'result', 'success', and 'improvement'. Likewise for the 'threat vs. trust' theme I searched for 'threat', 'trust', 'non-threat', 'non-threatening', 'easy', 'easier', 'worry', 'risk', and 'intimidating'. This process identified relevant sections of the transcripts, which were checked against those revealed through the manual process, and was carried out for all of the themes in Figure 18.

In most cases the technical approach verified data already identified manually, but at times new transcript excerpts were highlighted as relating to a particular theme. Each extract was given a numbered reference to the interview it was drawn from, to ensure it could be referred back to in the original transcript, and that the context of the extract could be easily checked during subsequent analysis.

During this process it became apparent that some original themes were too general and that partitioning of themes was required in order for important differences to be emphasised. This was the case in a theme originally labelled 'student learning'. An excerpt about the learning of two particular students which described in detail the nature of their learning in connection with a teacher's practice, was quite different from an excerpt only broadly about student learning, and completely different to an excerpt about the fact that teachers discuss student learning when they meet informally. Therefore, the student-learning category was partitioned to deal with data about student learning that is 1) specific, 2) general and 3) discussed in contexts other than appraisal.

Analysis based on the coding rules outlined in the following tables (Table 13, Table 14 and Table 15) established the number of participants who reported aspects of their appraisal discussion that related to each of the categories.

Rules used for coding transcript excerpts against the themes about appraisal discussion topics are outlined in Table 13.

Table 13 Appraisal Discussion Topics Coding Rules

Category	References to	Example
Student Learning - Specific  (of a particular student, group, class, or year level in relation to a specific learning outcome)	Learning Outcomes Achievement Results Success Improvement Impact	'I was saying [to the appraisee] do you remember when you used the abacus here, and [student A] was sitting across the table from you and he couldn't do it, could he, whereas [student B] was right there with you, didn't need the material and said straight away 24 divided by 6' [P2]
Student Learning - General	Learning Outcomes Achievement Results Success Improvement generally	'I said I can definitely see the children are making progress' [P10]
Teacher Behaviour	The activities a teacher does. Strategies, approaches, techniques used by the teacher. Teacher qualities	'it [the appraisal discussion] was basically surrounding a one-off lesson, the organisation, the preparation, the different ways that we managed it. We just talked about the differences in our styles.... a bit general' [P5]
Student Behaviour	The engagement and attitude of students	'she [the appraiser] said that the children were all totally engaged in the maths activities... they were totally involved in their maths, and really enthusiastic about it' [P3]



Rules used for coding transcript excerpts against the themes about the characteristics of appraisal are outlined in Table 14.

Table 14 Appraisal Discussion Characteristics Coding Rules

Category	References to	Example
Affective Concerns	Descriptions of features of the discussion that were about intentions to send positive messages maintain positive interpersonal relationships	'I think personally the more positivity you can give to your teachers in the appraisal process, the better.... the whole appraisal process is really important for building my self esteem and my confidence, and saying hey, I am doing okay, I'm doing a good job' [P10]
Use of data	Data Results Tracking	(none arose)
Summative purpose		'I do feel like the appraisal interview at the end of the year really is just a sum-up of the year rather than a specific time where he [the appraiser] gives a lot of feedback... I guess I'd never really thought of it as needing to be terribly helpful in terms of changing what you do in your classroom process' [P11]

Rules used for coding transcript excerpts about data-based inquiry into student learning proposition are outlined in Table 15.

Table 15 Types of Responses to the Idea of Data-Based Inquiry into Student Learning in Teacher Appraisal

Category	Examples
Positive	'that'd be really good, I'd really like that' [P8]  [it would] give you really good feedback, wouldn't it... it's only good because eventually it'll be more, it'll mean that my teaching will become more effective for the children. And that's the whole idea, isn't it, really? [P5]
Negative	'I would look at it as an imposition' [P3]
Threat	'there was a lot of praising, so it was very, very positive, and she looked at me in a very positive way, a very pleased way, and I felt really relaxed about that... it made me really comfortable... so I didn't feel threatened' [P3]
Data	'I guess you'd use the data to measure the goals, like action learning, where they started at, where they've come to and the process in-between. I mean, that's hard data that you're looking at rather than, well, I think you've catered for this gifted child because he seems to be happy. It's more factual and more fair' [P7]

In the fourth and final stage of analysis, a table was created with a row for each of the themes outlined above (Table 13, Table 14, Table 15). The coding rules in these tables were used to sort excerpts from the transcripts that related to each theme, and a numbered reference to the interview was recorded so that the number of interviewees describing evidence of each category could be counted. A frequency count (Miles & Huberman, 1994), rather than in-depth statistical analysis, was used to establish the relative frequency and significance of the different topics dealt with in appraisal discussions..

Analysis also sought to note relations between variables through, for example, comparing accounts given by an appraisee and appraiser who described the same appraisal discussion.

### *Inter-rater reliability*

The second coder independently analysed two of the 11 interview transcripts to establish whether those transcripts did or did not include data relating to the categories described in Table 13, Table 14 and Table 15. In total there were 22 judgements to be made (11 categories for the two participants) and the second coder's analysis was in agreement with mine for all 22 judgements. Where the coders agreed that transcript data did relate to a category, both coders identified at least one, and typically more than one, common excerpt from the transcript for that category.

## ***Results***

### *Appraisal Discussion Topics*

The most significant finding of this study was that of the eleven teachers interviewed, only one described a discussion that included talk specifically about student learning. In this example the teacher had asked the appraiser to focus on the depth of questioning in a maths lesson, and the engagement of students in the maths thinking and learning. The appraiser reported (see quote below) that discussion was held about two students, and compared the different impact of the teacher's approach on them – one student who achieved the intended learning outcome, and one who did not.

It was very clear that from some of the strategies she used, more learners were accessing the thinking and the learning, or in other instances, not so many were, and so as we talked about that, which instances did it work well, which instances wasn't it working. We very clearly identified some of the kids who were very clearly like, getting it, and the kids who were very clearly not... and why that may have been. [P2]

The appraiser reported that they went on to discuss the teacher's recognition of those students' needs, and use of resources to cater for them. Finally, they decided on strategies that could be used to improve both students' learning. The appraisee involved in this example was also interviewed, but she did not recall the discussion

in the same way. The topics recalled by the appraisee did not include any mention of the inquiry into the learning of the two students that the appraiser described. In contrast, the appraisee reported (see quote below) that the discussion focused on student and teacher behaviours not connected to learning outcomes, and her account emphasised the positive character of the experience, the praise she was given and how supported she felt.

There was a lot of praising so it was very, very positive. She looked at me in a very positive way, a very pleased way, and I felt really relaxed about that... she said 'I saw you reading out of the book and the children were really engaged in it... you used these pegs and you said that this green peg's the frog, and the yellow peg's the kiwi and the kids were really taking it on board, and they were really engaged. She said the children were all totally engaged in the maths activities, they were totally involved in their maths and really enthusiastic about it, and she said that it was really great to see that. Kids feeling really purposeful about their needs and enthusiastic. [P3]

The outcomes focus described by this teacher's appraiser is not evident in this account, from the appraisee, of the appraisal discussion. Rather, the appraisee's recollections were about affective concerns, and the discussion she and her appraiser had about particular practices such as reading the book and using the pegs.

Only three interviewees reported general discussion about their students' learning, though these were particularly fleeting and non-specific. Excerpts were coded to this theme where the reference to student learning did not focus on a particular student, group, class or cohort, and a particular learning outcome. For example, a teacher whose appraisal interview was focused around the observation of her art lesson, recalled talk about the end result being "really pleasing". By this, she explained, she meant, "the art work came out how we wanted it to look" [P5]. Two other teachers recalled mention of "improved numeracy skills" and that "children are making progress" [P10], but there was not mention in the discussion of the particular aspect of learning, or particular students who had improved, and no comment about the nature or extent of the improvement. Again, these were only passing and general references to student learning, and only given by three of the 11 participants.

Discussion about student learning was scarce despite coding which allowed for references to the learning or achievement of the whole school, whole class, group, or individual student learning, including those comments that were not based on formal data. When probed about whether any discussion was had about student learning, more than half of the participants suggested contexts other than appraisal in which student learning was discussed. These included both formal and informal contexts. Examples of formal contexts referred to included syndicate meetings, during classroom observations, in long term planning, and school-wide data analysis sessions. Informal contexts included the likes of 'around the photocopier' and 'at morning tea'.

The lack of emphasis in appraisal discussion about student learning was in contrast to the significant emphasis on teacher behaviour. This category of talk included discussion of processes, approaches, strategies and techniques used by the teacher where the consequences these had for students were not mentioned. All of the interviewees described appraisal discussion about teacher behaviours. There was, then, far greater emphasis on talking in appraisal about what teachers had done than about the connections between what they had done and its impact on student learning. The most commonly discussed teacher behaviours included planning, questioning, using resources, managing behaviour, preparing lessons, modelling, organising the classroom, organising school events and grouping students.

Two of the participants also described interview topics related to the 'teacher behaviour' category, but slightly different in nature. These were qualities, traits or characteristics of the teacher. One described interview topics about the confidence of the teacher and his/her willingness to learn. Another appraisal interview included talk about the effort of the teacher – the extent to which they had "tried to come to grips" with suggestions, and "tried out new ideas, tried out new approaches" [P2]. Again, these discussions made no reference to the way in which the teacher's qualities were perceived to have impacted on student learning.

Figure 19 shows the number of participants who reported topics relating to the four main themes about appraisal discussion topics as described above. It highlights the predominance practitioners give to behavioural and process-based aspects of teachers' practice rather than to inquiry into the learning of specific students and how teachers' practices relate to that learning.

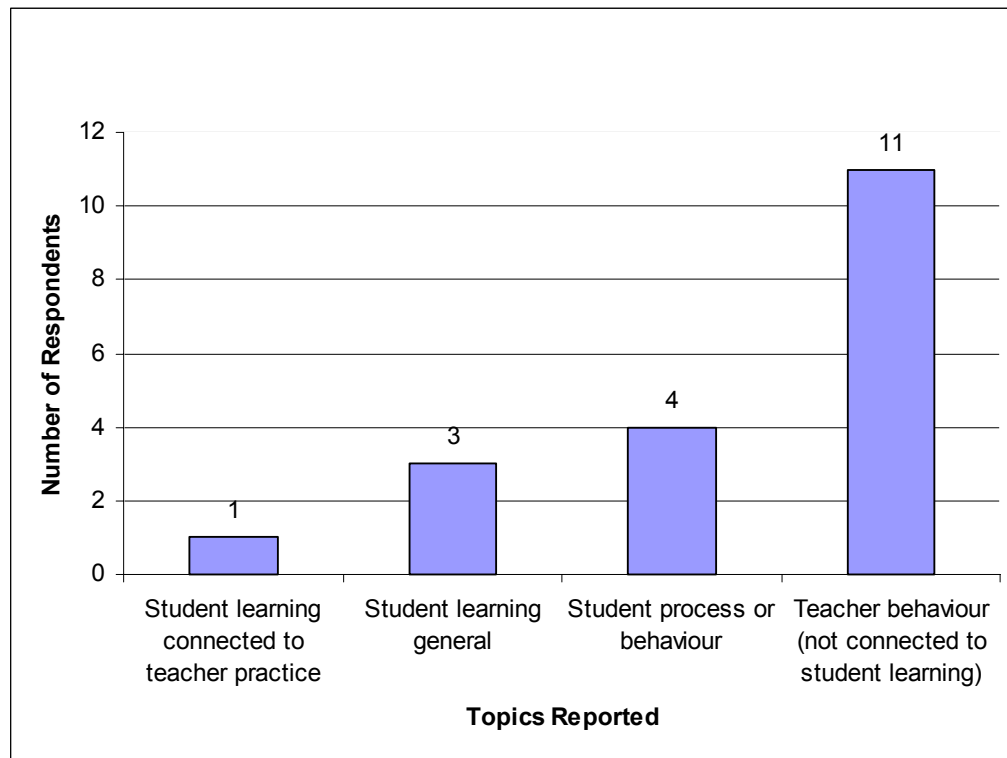


Figure 19 Teacher Appraisal Discussion Topics

When reporting on the topics of their appraisal discussion, nine of the 11 participants described appraisal goals or performance objectives and reported that one or more of these were the basis for the appraisal discussion. Appraisal goals are then a significant factor in determining what is inquired into during an appraisal discussion. Appraisal goals that were the basis of each participant's discussion are recorded in Appendix C as they were reported. Only three of the participants' appraisal goal topics relate clearly and specifically to student learning. These were goals referring to maths learning [P2], to outcomes of an art lesson for students [P6], and to children's reading progress [P11]. The remaining goals focused largely on teacher behaviours

without specific regard for the connection between that behaviour and the impact it has on student learning. There were both general (“classroom teaching” – P7) and specific (“ensuring all three Social Studies processes are used in the programme” – P11) examples of these.

#### *Appraisal Discussion Characteristics*

Questions put to interviewees about the topics of their appraisal discussions frequently led to reports about the characteristics or approach of the discussion (rather than the topic or content). Interviewees were also asked whether data were used or referred to during their appraisal discussion. Results in relation to these themes will be discussed next.

There was a significant reported focus on affective concerns associated with the appraisal discussions participants described. All of the practitioners made at least one comment about affective elements of the process when they were actually asked about the topics covered. The interview schedule did not include items about how they felt about appraisal discussion, or whether they were given positive or negative feedback. Despite this, all interviewees referred to affective consequences of the discussion such as “it gave me confidence” [P3] and “it made me feel really good” [P5]. Overwhelmingly these affective concerns were about positive feelings in the appraisal discussion. Nine of the 11 participants used the term ‘positive’ at least once while describing their experience. These results indicate that the nature of the appraisal discussion, how positive the messages are and how positive it makes participants feel, is of particular significance to practitioners. This has important implications for the appraisal approach argued for in this research, since the surfacing of data about student learning, even if negative, is deemed essential to maximising the learning opportunity for teachers in the appraisal process.

The following quotes are typical of those stressing the positive nature of the discussion and the importance of the affective dimension in appraisal.

We start off with in our appraisal interview talking about the commendations, the things that I would like to commend them for.... [a core purpose is] to acknowledge the, the other good work the teachers are doing. [P1]

I do try to give positive feedback as I go, saying, you know, it's really excellent to see that your long term planning is so well organised and easy to access, or it's excellent to see really current student work that's being displayed. I can see the kids are really proud of their work. You know, so.... I see that as a, a chance to kind of make some really positive statements about what the teacher overall has done this year. [P2]

#### *Use Of Data About Student Learning*

None of the appraisal discussions referred to in the interviews reported that data (other than teacher self report, observations) had been used.

#### *Summative Versus Formative Appraisal Discussions*

Five of the 11 participants referred to the summative nature of the appraisal discussion. Three of those (two teacher appraisees and their principal appraiser) were from the same school and all described appraisal discussions as summative in nature. The appraiser described the discussion component of his school's appraisal cycle as "more of a summative interview" [P1]. Both appraisees at this school also understood the discussion in this way:

I guess I'd never really thought of it as needing to be terribly helpful in terms of changing what you do in your classroom process [P11]

It's almost like a summative thing, it's towards the end, so it wasn't like an opportunity for him to go well have you tried this. It wasn't a bounce off session. It was more of a 'let's evaluate what you've done this year' type thing [P10]

Another teacher suggested that appraisal discussions were not particularly formative due to the timing, at the end of the year, and the time it would take to fulfil a formative purpose well:



Appraisal discussion happens too late to productively use data...in our appraisal cycle in term one it's setting goals, and term two and three you do observation, then term four in the discussion you're kind of wrapping it up and looking ahead [P6]

*Response to the Idea of Focusing On Student Learning during an Appraisal Discussion*

Once interviewees had been given an opportunity to describe their last appraisal discussion experience and had been probed about whether it focused on data about student learning, I asked them to respond to a proposition about how an appraisal discussion might be. The proposal was that appraisal discussions could focus on inquiring into student learning (or achievement or outcomes), and could include examination of the connection between data about students' learning and the teacher's practice. Results for the responses, in terms of being positive or negative, are outlined in Figure 20.

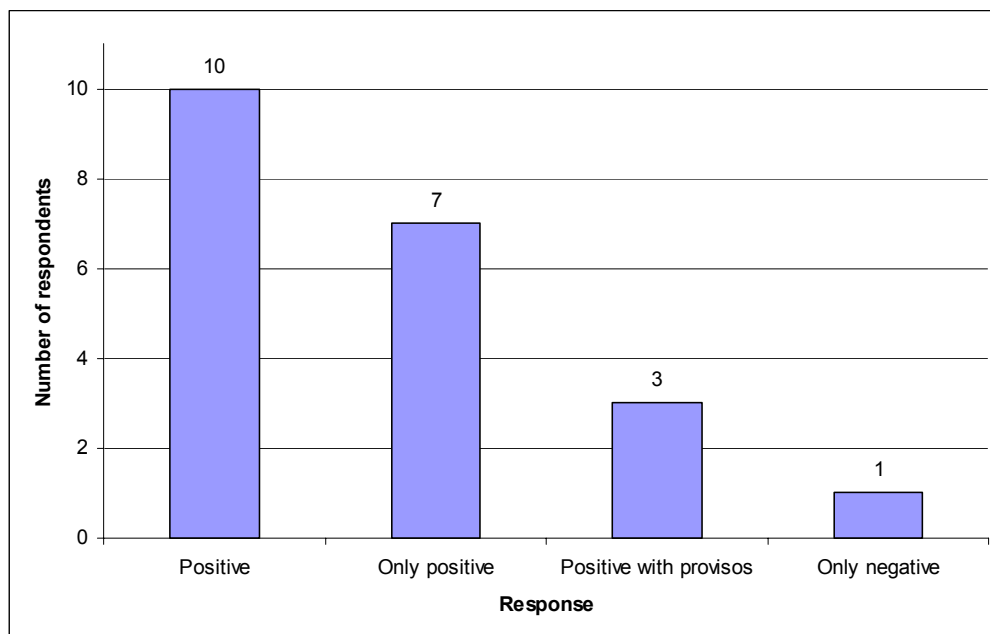


Figure 20 Response to Suggestion of Data-based Inquiry into Student Learning in Teacher Appraisal Discussions

*Teachers' receptiveness - "What a good idea!"*

The suggestion that appraisal discussions should focus on data about student learning received an overwhelmingly positive response from interviewees. Ten participants had a positive response. Seven of these had only positive things to say in

relation to the idea and only one participant showed an exclusively negative response to the suggestion. A typical reaction was “I think that’s great...so you’re using hard data of learning to then get the teacher to focus in on their teaching practice” [P7]. Others said “I actually really enjoy getting that kind of feedback” [P5] and “I think it’s a fair question to ask” [P9]. The positive reactions were typically expressed with a sense of the idea being obvious, reflecting common sense, and with a tone of bewilderment that such a simple, yet important, idea is not in common practice currently. Practitioners generally related the idea to the essential purpose of their work as seen in the statement below:

That’d be really good. I’d really like that. It would be useful ‘cause that’s what we’re here for, ‘cause that’s what really matters. I think that’s actually what it’s all about. That’s so fundamental. That’s actually the guts of what I do. And also it’s what I do really well.  
[P8]

The increasing enthusiasm with which each statement of Participant eight’s response was spoken accentuated her receptiveness to the proposed approach. Several of the other participants who had a positive response, also showed genuine enthusiasm towards the notion of an appraisal discussion in which the learning of their students was a focus. As they pondered the idea, their body language, expression, and questions they posed back to me revealed genuine enthusiasm:

... well, that’d be good. It’d be really good. Like it’d give you really good feedback, wouldn’t it? [P5]

... it’s good because it’ll mean that my teaching will become more effective for the children. And that’s the whole idea, isn’t it, really?  
[P5]

The suggestion also highlighted for one teacher a concern about the invisibility of children in their current approach to appraisal. She laughed at the irony of focusing on student learning being novel as she said, “it’s interesting that you ask about children. I think they rather get lost in our appraisal system.... To ask about students’ learning, I mean it’s so alien, it is so alien”.

The reference in the suggestion to using data was a particular attraction of the model, and a basis for many of the positive responses. Referring to student learning data, one appraiser commented:

That's what you'd use to measure the goals – where they started at, where they've come to and the process in-between. That's hard data that you're looking at rather than, 'well I think you've catered for this gifted child because he seems to be happy'. It's more factual and more fair. [P7]

Another teacher inferred a distinction between pure data and observation:

It would have to come from pure data, observation as well, but data-based, so that you're not talking about emotions. [P6]

This distinction is similar to that made by Piggot-Irvine between informal casual observations and objective information that enhances the validity, fairness, rigorousness and reliability of performance management (Piggot-Irvine, 2001a, 2003b).

Talking about benchmark data in their school, one participant described how potentially valuable it would be to discuss that data as part of the appraisal process:

It's assessment that is actually really useful, it's not just summative, it's not just right at the end, it actually really is informing the learning that's happening, informing your planning, informing everything else that goes on. [P11]

As well as benchmarks, a number of other sources of data were also suggested by teachers in responses to the proposal being useful in the approach. These included student portfolios, work samples, and class, syndicate and school assessment records.

Considering the notion of data-based inquiry into student learning caused one practitioner to reflect on how they miss that approach from their pre-service teacher education experience. She said, "That's something I miss. When we were students,

you know, we were getting that feedback all the time, and [once you're a practising teacher] you don't get it so much. You're out on your own a whole lot more." [P5]

*Provisos*

Three participants were mainly positive in their attitude to the suggestion, but went on to express a consideration that they thought should be taken into account in implementing such an approach.

If the teachers have had professional development in a particular area the year before it's [appraising them in that area] totally relevant. [P6]

Threat was a key theme in the warnings people raised following their positive response to the proposal.

You'd be more tense if you knew people were focusing on what the students were learning.... it would be helpful, once I've got over the tenseness. [P4]

I'd feel fine with that [data-based inquiry into student learning]. Oh but mind you, I'd feel fine with that in art because I kinda feel like that's something I can do quite well. It might be a lot scarier with something I don't know so well. [P5]

I'd be a little bit nervous and a little bit threatened, but I think that's probably quite a natural feeling. [P9]

One appraiser raised some concern about confronting a teacher with data:

To say 'tell me why only 30 per cent of your class is achieving above expectations, where the rest of the school have 70 per cent of the class'. I wouldn't want teachers to be put on that sort of spot because I think there's so many other variables that impact on whole class achievement. [P1]

He referred to the allocation, for instance, of difficult children to capable teachers, and the careful selection of students for newer teachers, suggesting that these contextual factors would need consideration as part of the discussion. He also

mentioned the importance of not losing teachers' support and trust in the appraisal process.

*A solitary negative response*

Only one participant had strong disapproval of the suggestion of appraisal discussions focusing on data about student learning. The interviewee struggled to make sense of the suggestion, as can be seen in the transcript excerpt below:

Interviewer: I'm wondering how you'd feel about the idea of student learning – students' achievement and outcomes - being made the focus of your appraisal

P3: In what ways?

Interviewer: Say, rather than the focus being, for example, your questioning, the focus would be on, for example, student achievement on recognising number patterns.

P3: What do you mean by that?

Interviewer: How would you feel if student learning was the thing that the appraisal observer was to observe and make notes on and to be the focus of the discussion?

P3: How would I feel? Um, I don't really get what you mean though. I don't know what you're getting at.

Interviewer: Yeah. I'm just wondering how you'd feel about the students' achievement being the focus, rather than the focus being on your approaches. Does that make sense, the distinction between the two?

P3: Not very clearly. I'd need a concrete example.

Interviewer: If, say, the observer came in, and made notes on which students achieved a goal about number patterns, and how many students couldn't do certain aspects of the lesson, and what students said that showed that they did or didn't learn

P3: I think also that is, yeah, that a lesson does include that, you know. That's the way I, that's why, what I'm doing it for.

Interviewer: And if those were the things that the, the observer made notes on and were the basis of the appraisal discussion - how would you feel about that?

P3: Oh, oh I'd think that was generated by them, that that's their focus. It wouldn't be my focus.

The same proposition that was quickly and easily understood by all other interview participants (a suggestion that student learning, and data about student learning, be the focus of appraisal) was, to this teacher, ambiguous and difficult to comprehend. Once she grasped the notion, the strength of her opposition was made plain. In

particular, the negative response she went on to make referred to an approach such as this leading her to be suspicious of ulterior motives, and a sense of pressure and imposition. Interestingly, both the large group of participants who responded positively to this notion, *and* the teacher with the negative response (above) highlighted the fundamental nature of student learning as a purpose for what they do.

## *Discussion*

### *Explaining the Lack of Focus on Student Learning in Teacher Appraisal Discussions*

The findings of this study have highlighted that appraisal discussion topics generally focused on teacher practices without exploring connections between those practices and the impact they have on student learning. This was the case even when the underlying intentions and purposes lay in concern for student learning, and yet the appraisal discussions themselves did not explicitly focus on student learning, or data about student learning. It was even the case when an appraiser recalled that there has been a focus on student learning in a discussion when that focus had not been apparent to the appraisee involved. Furthermore, even when teachers had an appraisal goal that related to student learning, the subsequent discussion, while related to that goal, did not inquire into data about student learning. Again, the focus of discussion was limited to teacher practices.

There are a number of likely reasons for the lack of reference to student learning in appraisal discussions. This following section will discuss a number of appraisal characteristics and their role in preventing data-based inquiry into student learning. Characteristics include undemanding appraisal goals, assumed connections between practices and learning outcomes, perceptions of appraisal as ineffectual and summative, limiting document templates, and alignment with appraisal guidelines that are, arguably, themselves flawed.

*Undemanding appraisal goals*

This study established that appraisal goals (or development objectives) are, in nearly all cases, an important basis for appraisal discussion. They were found to be extremely limited in terms of their focus on student learning. Rather, they tended to concentrate on teacher behaviours, practices and actions, without requiring examination of the influence on student learning. The significance of appraisal goals and the implications of goals not being focused on student learning established a need to study this aspect of the appraisal process further. A subsequent study was undertaken to address appraisal goals specifically, and that study is reported in Chapter 5.

*Assumed connections between practices and student learning outcomes*

Several participants, when asked why they had not focused on student learning, gave reasons that suggest teachers assume that fulfilling particular teacher practices will improve the learning of their students. The following teacher's comment highlights the assumptions inherent in the discussion of her involvement in professional development and her planning and assessment practices. It is taken for granted that these have a connection with student outcomes, without those outcomes actually being discussed:

If I am involved in the arts development programme, and if my assessment practices are up to scratch and my planning is good, then does that mean that the student outcomes will be all right? I assume that's the thinking behind it [P8]

A student learning data-based inquiry approach to appraisal would both discourage making such assumptions, and encourage testing of any pre-existing assumptions. For instance, questions would be raised about student outcomes before and after involvement in the professional development programme, about which aspects of the planning helped to improve student learning, in which areas, for which students and why?

The lack of focus on student learning in appraisal goals was also explained as being due to assumed causation between teacher practices and student outcomes:

... 'cause this one [goal] is all about planning, and planning affects students' learning. 'Cause planning obviously, you know, it must, it must connect with the students' learning 'cause it's part of that planning, learning assessment, teaching [cycle]. [P11]

The study also found that appraisal discussion about student behaviours is undertaken with an assumption that those behaviours indicate positive outcomes for students:

By them [students] actually looking like they're into the task, for us that's an indicator, whether a right indicator or not, that's an indicator that they actually are taking in some learning, more than if they're just sitting there. [P2]

This type of discussion is valid, but it should not stop at considering whether students looked like they 'were into the task', but go on to discuss the learning outcomes that were achieved as a result of the task, and the pedagogical approach of the teacher that promoted or hindered that achievement.

*Perceptions of appraisal as an ineffectual process*

A number of studies have highlighted a perception of appraisal, and appraisal discussions, as ineffectual. This perception may contribute to the tendency to ignore the importance of student learning in appraisal discussions. Down, Chadbourne and Hogan's longitudinal study (2000) examined the stance of Western Australian teachers to performance management, at its introduction in 1997, and again two years later. Findings revealed that teachers' original doubts and concerns about performance management were somewhat reduced at the end of the second year. The reason for the reduction, though, was in the teachers' experience of performance management – "in many instances, it was not happening, happening only superficially, or having no appreciable effect" (p. 217). Teachers reported, "playing the game"... [to achieve] superficial compliance with the mandated aspects of [the] policy, but did not engage with it at a deeper level" (Down et al., 2000, p. 218).



To assist in the validity of the information gathered in this study, participants were asked to bring along and refer to documentation to support their recall of the appraisal discussion. Most did refer to these and it became apparent that the document templates themselves might explain, to at least some extent, the lack of focus on student learning in the discussions. In Table 16 are three typical examples of the headings used on templates which were filled in during an observation and the basis of the appraisal discussion.

Table 16 Appraisal Observation Template Headings

School	Headings
A	Things seen
	Things heard
	Suggestions
B	Teaching Responsibilities
	Management Responsibilities
	School-wide Responsibilities
	Personal Development
	Objectives
C	Focus
	Describe
	Impact
	Coach

These headings provide a framework for the discussion, and as such are an important tool in prioritising and focusing the inquiry that the appraiser and appraisee engage in. While there is the potential to record data and comments about student learning under the heading of these templates, the headings themselves do not require or even encourage such a focus. Even in school C, an appraiser used the section headed 'Impact' to record notes about the impact on students' behaviours rather than learning outcomes:

I said the impact was that some students seemed actively interested in solving the questions she was asking about because they were looking right at the material.... it was looking at the interest, the engagement, kids who were losing focus.... I commented that I felt they were actively interested, or they appeared motivated, or I see students who seem interested in sharing their thinking. Some students seem to lose focus. [P2]

Behavioural engagement, interest, and motivation were used without question as indicators of cognitive engagement and subsequent learning outcomes, and were the basis of the discussion.

*Student learning discussed in contexts other than appraisal*

A number of participants in this study mentioned, when confirming that their appraisal discussion had not focused on student learning, that student learning was talked about in other contexts. This finding echoes similar points from other studies which found that important (student learning focused) discussion occurred in informal exchanges between educators (Down et al., 2000; Hogan et al., 1998; Southwest Educational Development Laboratory, 1999). These informal exchanges, or 'everyday conversations' as Haigh (2000) refers to them, have features that make them particularly conducive to professional learning – they are serendipitous, topics are improvised and influenced moment-by-moment by the participants (rather than pre-determined topics), are focused on personal, local and immediate matters, and often include story-telling. He also suggests that conversations are non-threatening with open agendas that encourage permissiveness and risk-taking (Haigh, 2000, p. 4). Conversations, likely to encompass these features, have been found to function as sites for professional educators to engage in discussion about student learning in relation to their teaching. It may, then, be productive to embrace features of these conversations into appraisal discussion. I would, though, argue that the topics should not be entirely improvised as conversations are. Rather, topics should be broadly pre-determined to ensure they are about the connection between teachers' practice and their students' learning. Valuing conversation as a context for professional learning, and incorporating the features of conversation into appraisal

processes requires those conversations to be rigorous in the examination of impacts on student learning.

#### *Alignment*

Examining the impacts of teachers' practice on student learning in appraisal discussions is likely to be limited where the discussion process is aligned to other aspects of appraisal that do not have a focus on student learning. Alignment is desirable where that which is being aligned to has a focus that is significant. Alignment is not desirable when the focus of that being aligned is flawed in some way. The findings of this study suggest that alignment does exist between the extent to which schools focus on student learning in their professional standards' indicators (see Chapter 3) and in the topics of appraisal discussions. In this case, I would argue, the alignment is not constructive or desirable when it serves to limit the attention given to student learning.

#### *Perceptions of appraisal discussions as a summative phenomenon*

The most current policy resource available to schools, 'Teacher Performance Management: Primary School Teachers, Primary School Deputy/Assistant Principals – A Resource for Boards of Trustees, Principals and Teachers' (Ministry of Education, 1998c), makes clear the role of performance management, and appraisal, in fulfilling both summative and formative functions. The summative function outlined in this text relates particularly to the professional standards, performance expectations, pay progression, career advancement and competency procedures. The formative function is signalled by the many references throughout the document to "development and growth" (p. 2), "improving the quality of teaching" (p. 3), and "[supporting teachers to]...develop their skills" (p. 4) (Ministry of Education, 1998c). Nearly half of the interview participants, though, described an exclusively summative perception of their appraisal discussions.

There were two main reasons for summative perception of appraisal discussions. The first relates to the tasks and topics dealt with in the discussions, and the second reason relates to the timing of the discussion. Appraisees understood the purpose of

the appraisal discussion, as evident in their experience of it, to be, for example, to “fill in the form” [P8], rather than to “be ... helpful in terms of changing what you do in your classroom process” [P11]. This reflects a dominant concern for compliance issues, rather than improvement, development and learning focused issues. It is clear from the data that practitioners perceived appraisal discussions to fulfil, primarily, summative purposes. Participants also reported that appraisal discussions occur most frequently towards the end of an appraisal cycle, which is typically at the end of a school year. Participant 10 suggested that because the discussion occurred “towards the end” it was not treated as an opportunity to consider suggestions, or “bounce off” ideas. Rather, it was a purely evaluative experience. Participant six also suggested that the timing of her appraisal discussion, late in the school year, meant that the discussion focussed on “wrapping up” the current year and was not productive in using data. There is a connection, then, between the timing of an appraisal discussion, and the likelihood of it fulfilling only summative purposes.

Highlighting the mainly summative perception of appraisal discussions is not intended to suggest that there should not be a place for summative evaluation in appraisal discussions, and a focus on data about student learning in those evaluations. I have two criticisms of appraisal discussions that are exclusively summative. The first is that they may decrease the likelihood of appraisers and appraisees engaging in discussion about student learning based on evidence. Rather, ‘summative only’ discussions will increase the emphasis given on generally affirming teachers in appraisal discussions. Raising difficult and challenging issues in an evaluation, is far more likely to be avoided in a process considered summative, than one that is understood to be formative. Davis, Ellett and Annunziata suggest that when assessments of teachers are summative, they are often rated highly and the process is perfunctory. They say that “opportunities are being missed for assessment of the teaching and learning processes to support collegiality and collaboration, to identify professional growth needs, and to acquire the contextually specific data necessary to improve student learning” (Davis et al., 2002, p. 287). I have emphasised the final phrase to highlight that the support of collegiality and

collaboration is not sufficient, but that a focus on student learning, through reference to data about student learning, is required. My second criticism of entirely summative appraisal discussions is that they disregard both the requirement for, and the opportunities presented by, appraisal discussions that function to improve teaching and learning. The lack of attention to formative appraisal may be explained by well-established appraisal systems that have, over time, entrenched a summative focus in the culture of schools' appraisal practices. Davis et al. highlight the key role of leadership in extending teacher evaluation (or appraisal) beyond ritualistic traditions, so that teaching and learning can be improved (Davis et al., 2002). They note that teacher evaluation policies and practices have not improved teaching and learning to the extent they have the potential to. This unrealised potential could be addressed through more thorough attention to the formative function of appraisal. Ensuring that appraisal discussions provide appraisers and appraisees with opportunities to engage with data about student learning for both types of appraisal, formative and summative, is essential.

#### *Appraisal is Threatening*

Much of the literature about teacher appraisal reveals that practitioners report a sense of threat and anxiety about the process (Calabrese, Sherwood, Fast, & Womack, 2004; Cardno, 1994, 1998, 2001; Down et al., 2000; Down et al., 1997; Fitzgerald et al., 2003; Gratton, 2004; Grootenboer, 2000; Gunter, 1996; Hogan et al., 1998; O'Neill, 1997; Townsend, 1995; Youngs & Grootenboer, 2003).

The findings of the study discussed so far, in summary, show that appraisers and appraisees:

- Base appraisals on undemanding goals
- Assume connections between practices and students' learning outcomes without questioning those assumptions
- Underestimate the power and potential of appraisal discussions to examine student learning in ways that impact on, and improve, practice
- Create and rely on document templates that act as barriers to student learning focused appraisal discussions
- Align elements of the appraisal process in ways that limit the focus on student learning
- Discuss student learning in other and informal contexts, but not in appraisal

- Perceive appraisal as a summative phenomenon

These findings show that student learning is rarely examined in teacher appraisal discussions as a means of considering the impact of teachers' practices. The lack of focus on student learning may be explained as a response to the perception of appraisal as a threatening process.

*Student learning as 'undiscussable'*

In the appraisal discussions reported on here, student learning is treated as what Argyris calls an "undiscussable" issue (Argyris, 1986, p. 76). These are, Piggot-Irvine explains, "difficult problems that remain untackled" (2001b). A perception that discussing student learning could potentially be embarrassing, leads to what Argyris describes as 'organizational defensive routines'. These are "any action or policy designed to avoid surprise, embarrassment or threat. But they also prevent learning and thereby prevent organizations from investigating or eliminating the underlying problem" (Argyris, 1986, p. 75).

*Discussing the 'discussable' and avoiding threat*

School leaders face dilemmas in managing tensions between their concern for organisational goals, and their concern for collegial relationships. There are three typical approaches to dealing with these dilemmas: avoidance, soft-sell (defensive) and hard-sell (controlling) (Cardno, 1994, 1998; Piggot-Irvine, 2001b, 2003a). The finding in this study about a lack of reference to student learning data in discussions between appraisers and appraisees represents an avoidance approach. Cardno and Piggot-Irvine also found that appraisers participating in a 'state of play' study generally avoided assembling objective information for appraisal. They highlighted the potential for information that was gathered being sourced from loosely conducted observation. Rather, they advocate that discussions should be "based on factual, objectively collected, 'data-based' information" (Cardno & Piggot-Irvine, 1996, p. 20).

While this study did not analyse transcripts of actual appraisal discussions, the messages that appraisees took away, and remembered, from their appraisal, are in fact quite significant. Their recollections of the appraisal as a positive experience, and the reasons given for this positivity, suggest that appraisers are generally operating in “Model I theory-in-use” (Argyris, 1999), since “Model II theories-in-use” should lead to quite different recollections. While a Model II approach may result in participants having an appraisal experience they recall as positive, the reasons why it is experienced as positive are important. For the participants in this study, positiveness was associated with having ‘got through’ the appraisal without difficulty or challenge, and with having complied with the mandated requirement. The governing values of a Model I theory-in-use are to achieve intended purposes, maximize winning and minimize losing, suppress negative feelings, and behave according to what you consider rational (Argyris, 1999). Under a Model II approach, in contrast, participants’ positive recollections are more likely to be based on satisfaction with appraisal as a learning opportunity. Under a Model II appraisal process the key values one would expect to be recalled are those of high quality analysis and discussion of teaching practice in relation to students’ outcomes. Some evidence for these assertions is found in the study of Calabrese et al. which did examine actual appraisal conferences. Their analysis of both the thinking and dialogue of participants showed that their teachers and principals did operate out of a Model I theory-in-use. The authors concluded that principals and teachers “seek to maximize winning and minimize losing according to a hidden agenda, suppress negative feelings during the summative evaluation conference, and act in a manner that projects consistent rational behavior” (Calabrese et al., 2004, p. 116). None of these is conducive to appraisal discussions that examine the impact of a teacher’s practice on their students’ learning.

Even appraisal discussions that seem on the surface to facilitate a learning approach, such as those outlined by Gunter (1996), can seem less impressive when considered in relation to Model I guiding values/defensive reasoning versus Model II guiding values/productive reasoning (see Table 17).

Table 17 Defensive and Productive Reasoning

Defensive Reasoning	Productive Reasoning
Is concerned with blocking information which we personally feel will create unpleasantness or lessen our control of a situation	Is concerned with generating information in an effort to increase the possibility of critical reflection-in-action
Guiding Values: - win – don't lose - avoid unpleasantness - maintain control	Guiding Values: - seek and give valid information - share control and solutions - monitor solutions jointly
Strategies: - not checking assumptions - giving indirect or mixed messages - not explaining reasoning - using questioning to control	Strategies: - checking assumptions - being forthright - disclosing reasoning - asking questions as genuine inquiry

Source: Cardno, C. (1998). Making a difference by managing dilemmas. *SET: Research Information for Teachers*, 1(13).

In Gunter's study (1996) of appraisal as a learning process in connection with the status of schools' as learning organisations she claimed that the professional appraisal discussions the participants in her study engaged in, facilitated a learning approach. These claims were based on reports that "teachers were reassured, good practice was celebrated, and any appraisee concerns were discussed and solutions sought"(Gunter, 1996, p. 95). These features of the appraisal discussions may promote some learning, but do not, I argue, facilitate the kind of learning that productive reasoning has the potential to generate. The focus on reassurance and good practice being celebrated and no mention of appraiser concerns, suggests a rather more defensive approach, in which unpleasantness and valid information were avoided, and control was predominantly with the appraisee.



*An absence of data about student learning in appraisal discussions*

None of the participants in this study reported any reference to data in their appraisal discussion. This finding, to be expected given that most of the discussions were about teacher practices rather than outcomes, is significant. It reflects a similar finding in Gunter's study (Gunter, 1996) about teacher appraisal: "data collection in addition to classroom observation did not happen....formal review of targets was not taking place amongst a significant majority of the respondents" (p. 96).

In their work on the effectiveness with which data is used in large-scale reform efforts in England and Canada, Earl and Fullan (2003) examined tensions school leaders experience in regard to using data as part of large-scale school reform efforts. Their findings reveal that despite recognition of the positive role data plays in helping to focus thinking and give insights for decision making, leaders have reservations, anxieties and worries about their ability to use data effectively:

School leaders....were very forthright in their anxieties about using data. Even when they were positively disposed to looking at data as part of their decision-making, they expressed insecurity about their skill in gathering, interpreting and making sense of the information about their school. Many of them indicated that they had not had training or experience in research, data collection, data management or data interpretation. (Earl & Fullan, 2003, p. 388)

Earl and Fullan's findings may well give some clues for the reasons appraisers here seem resistant to using data in teacher appraisal discussions. This is an important area requiring further study in the New Zealand context, particularly since the use of data is an essential element of the recently introduced school planning and reporting requirements (Ministry of Education, 2003c, 2003d, 2003e, 2003f, 2003g).

*The 'discussable' – positive messages in appraisal discussions*

While only one participant described an appraisal discussion that included reference to student learning, all 11 participants made at least one comment about affective concerns, such as how the discussion made them feel, or how positive the experience was. These comments were not initiated by questions related to affective concerns,

but since the entire sample raised them, they are clearly and generally considered important. This is perhaps explained by the traditional, and much reported, notion of appraisal (and particularly appraisal discussions) as a threatening endeavour likely to cause anxious responses from those involved.

The appraisal approach argued for in Chapter 2 does not place any particular emphasis on affective concerns such as how appraisal makes participants feel or how positively they view the experience. The approach does not seek to be detrimental to interpersonal relationships, but does privilege the goal of student learning focused appraisal. Mills et al. also recognised the limitations of a focus on affective concerns. Their research examined the ‘curricular conversations’ or ‘professional inquiry’ a group of educators engaged in (2001). As the quote below reveals, a culture of inquiry requires more than concern for interpersonal issues:

More often than not, our conversations revolved around compliments – we took time to validate one another. Looking back, we needed to build trust and respect for one another before we could truly inquire into each other’s beliefs and practices...while we all knew deep inside that the time we took to nurture one another was necessary, we also know it wasn’t sufficient...[we needed a] structure that would push us to take a deeper look. (Mills, 2001, p. 23)

Validations, such as those the teachers above gave each other before moving on to deeper issues, are one aspect of collegiality. But, as Timperley and Robinson warn, “the benefits of collegiality rarely live up to expectations because norms of mutual and uncritical support take precedence over enhancing the validity of information fundamental to quality decision making” (2000, p. 48). The reports given by an appraisee in my study about her appraisal discussion, suggest that the support she was given was uncritical:

So it was very, very positive and she looked at me in a very positive way, a very pleased way. And I felt really relaxed about that. There was a lot of praise and it made me really comfortable. And a lot of positive body language like smiling a lot and saying lots of positive things to me about the lesson.... [It] gave me a lot of confidence. It was within these very supportive, structured guidelines, and I got positive feedback from someone on the management team. [P3]

While she was clearly pleased with the positive focus of her appraiser's feedback, it was unlikely have an impact in terms of her improving her practice, and seeking to enhance the learning of her students. The benefits the discussion had in terms of enhancing collegiality do not compensate sufficiently for the missed opportunities for critical inquiry, and for enhancing teaching and subsequent learning for students.

The attention to maintaining positive interpersonal interactions does not, in itself, reflect a deficiency in appraisal discussion. The quality of an appraisal discussion is, though, inhibited if attention to positive interpersonal interactions is the only focus, or if concern for maintaining those positive interactions prevails over concern for attending to the significant and important matter of student learning. Appraisers who are preoccupied with making sure positive messages are stated and (as in the examples in the section from page 116) stated first, should consider the risk of those positive messages obscuring other more professionally challenging messages raised in an appraisal discussion.

*A Proposed Approach for Teacher Appraisal: Data-Based Inquiry into Student Learning*

In the final phase of questions interviewees were asked to respond to a suggestion that teacher appraisal, as argued for in Chapter 2, should be inquiry based, the inquiry should be focused on student learning and data should be used to provide evidence of connections between a teacher's practice and their students' learning. Responses from both appraisees and appraisers were intensely positive, with only one of the participants giving an exclusively negative response.

*The predominant positive response*

Practitioners were overwhelmingly receptive to the idea of an appraisal that does involve inquiry into data about student learning. It has been suggested here that tackling potentially challenging discussion about student learning (and data) in appraisal might be avoided due to it being perceived as threatening, for either the appraiser and/or the appraisee. The responses of interview participants in this study,

however, suggest that there need not be concern about threat. On the contrary, these educators were particularly responsive and un-threatened by the idea.

The participants here acknowledged the lack of concern for the fundamental aspect of student learning in current appraisal practices. When clarifying (during the interview) that my perception that they had not talked about student learning was correct, many agreed in embarrassment. A similar response occurred in Down et al.'s study, when they were asked to describe an example of a change in their practice as a result of performance management. Replies of 'no' and 'not really' were accompanied by ironic laughter. Recognition of a lack of deep and meaningful learning, as a result of appraisal, is an important catalyst for improvement.

Teachers are, then, generally receptive to the idea of focusing on data about student learning in their appraisal. I would like to revisit here the notion of two main perspectives of appraisal (Fitzgerald, 2001; Fitzgerald et al., 2003; Grootenboer, 2000; Youngs & Grootenboer, 2003), since the perspective reflected in current approaches to appraisal is quite different to the perspective teachers aspire to. The first, bureaucratic, is focused on accountability and the second, professional, is focused on development. A data-based inquiry into student learning approach requires both perspectives to be accounted for. In this model the accountability is not so much for the student learning that is revealed as a result of the inquiry, but for the quality of the appraisee's own inquiry process. Rallis and MacMullen explain:

... to be accountable means to be obligated to understand and explain one's actions. Accountability relies on feedback; it links performance with results. Thus accountability in schools is not only about results but also about every aspect of teachers' actions....practitioners who are accountable evaluate their own practice and then use the information to improve. From this perspective, accountability is the foundation of successful practice because it entails knowing what we do and learning from that knowledge. (2000, p. 24)

What is appealing about Rallis and MacMullen's notion of accountability is the way it merges the bureaucratic and professional perspectives by signalling the importance

of accountability for development, with less emphasis on making a choice between either accountability or development.

*Examining the lone voice of disapproval*

Only one participant expressed an entirely negative response to the suggestion that appraisal discussions could, and should, focus on student learning, and base inquiry about student learning on data. Since this view was a minority one, it would be easy to dismiss as being of little interest. On the contrary, the reluctance this participant had to even make sense of the suggestion, and the strength of her opposition, provide an extremely valuable insight into the issues, constraints and likely misconceptions associated with proposing a data-based inquiry into student learning approach to teacher appraisal.

The notion of focusing on student learning outcomes as a basis for understanding a teacher's practice, and of using data about learning outcomes to inform the discussion, was a completely outlandish idea to participant three. She took some time to grasp the fundamental idea of the suggestion as the transcript of this section of her interview shows (see p. 122). Initially she seemed confused and perplexed, and when the premise of the suggestion became clear, the confusion gave way to frustration and strong disapproval. This finding is important in highlighting that student learning focused teacher appraisal will be, for some, both unfamiliar and unpalatable. Suggesting that teachers adopt such an approach would need careful and thorough explanation.

The main points made by participant three about the appraisal approach argued for here, are paraphrased below. They related to:

- concern about who generated the focus of discussion and whether it was also embraced by the appraisee - the importance of negotiation rather than imposition of appraisal discussion foci and agendas
- concern about the presence of ulterior motives
- a feeling of pressure of outcomes being more important than the process
- perceived loss of support

The concern expressed about who generates the focus of appraisal is an important one. While it is advocated here that the focus be required and made explicit at all levels of the appraisal system, including the national context, teachers should be responsible for generating the specific details of that focus. For example, policies would dictate that the focus on student learning should be evident, but teachers themselves should dictate the particular learning outcomes that are relevant to inquire into, and which learners' outcomes are appropriate to focus on. The second concern, about 'ulterior motives', echoes unease in the literature about the potential for appraisal mechanisms to be connected, even more directly than they are under the current framework, to teachers' pay progression, and to control and compliance measures (Haynes et al., 2003). To address this concern, an absolutely confidential and transparent appraisal process (Piggot-Irvine, 2003b), along with other measures, would need to be in place. Participant three also perceived that outcomes would be more important than process in the proposed appraisal approach. That is a misconception. The purpose of requiring attention on outcomes is so that the best learning and decisions about the qualities of teachers' processes and practices can be made. The proposed approach does not seek to ignore the process aspect of teacher's practice, but does recommend that educators should engage in dialogue that examines connections between processes and outcomes. The final concern participant three raised was about support being withdrawn or being less evident in the proposed appraisal approach. It would be important to emphasise, for those who perceive the approach in this way, the links between appraisal and other aspects of the performance management process, in particular, professional development. While appraisal could be supportive as well as focusing on student learning, professional development would play an important part in providing support and development opportunities.

Youngs and Grootenboer's project exploring teachers' perceptions of appraisal in their schools reveals a strategy that could be important in addressing many of the concerns raised by the teacher opposed to appraisal focused on student learning outcomes (2003). It was found, from the questionnaire responses of more than 268

teachers, that teachers involved in the review of their school's appraisal system were far more likely to understand the system, and have a professional rather than bureaucratic view of the system, than those who had not been involved.

### *Conclusion*

A focus on student learning, the fundamental goal of all teachers' practice, was conspicuous in its absence in the appraisal discussions reported in these interviews. There seems, too, a missed opportunity to utilise data on student learning as a basis for discussing teachers' practice, in a way that might lead to improvements in teaching and learning. On a more positive note, teachers were overwhelmingly receptive to a suggestion that appraisal discussions could, and should, focus on data about student learning. This reveals great promise and potential for teacher appraisal discussions.

## CHAPTER 5

### Appraisal Goals – what goals are used in teacher appraisal?

#### *Introduction*

Teachers' performance appraisal is required to be connected to 'development objectives' agreed on by the teacher and their appraiser at the beginning of each appraisal cycle (Ministry of Education, 1997a, 1998c). Development objectives are often referred to in schools as 'appraisal goals', and they provide a basis for appraisal observations and discussions. They are, then, a crucial tool in establishing the parameters for inquiry in teacher appraisal, and ultimately restrict or promote the extent to which those appraisals focus on data-based inquiry into student learning. The purpose of the study reported in this chapter was to investigate the nature of teachers' appraisal goals. Sixty-eight teachers from eight schools completed questionnaires about their appraisal goals. The findings give an insight into the extent to which appraisal goals focus on student learning, and the degree to which they are likely to improve teachers' practice.

In the study about appraisal discussions (see Chapter 4), nine of the 11 participants reported that appraisal goals were addressed in those discussions. This highlights the important role they have in setting the course for subsequent phases of the appraisal process. Appraisal goals that are disconnected from student learning will act as gatekeepers to inquiry discussions that focus on student learning. Appraisal goals clearly connected to student learning will, on the other hand, steer discussion between an appraiser and appraisee towards those important concerns.

The appraisal discussions reported in Chapter 4, in the vast majority of cases, did not focus on inquiry into student learning. An explanation for this finding given by a number of participants in that study was that the teachers' appraisal goals simply did



not require that focus. Since their appraisal goals were not student learning focused, their appraisal discussion was not student learning focused. This study, sought to establish whether the tendency for goals to not focus on student learning was unique to that particular sample, or whether that omission is a more widespread problem. By investigating what teachers' appraisal goals are about, this study will examine the extent to which the findings from the appraisal discussion study (that the nature of goals explained the limited focus on student learning in appraisal discussions) are generalisable. Goals are significant since they are required in the guidelines to be a focus of appraisal discussions. If goals are not about student learning, then appraisal discussions are unlikely to be about student learning.

This rationale given above for undertaking this study reflects the sequential exploratory design of the series of studies reported in this thesis. Sequential exploratory design of mixed methods research typically serves to "use quantitative data and results to assist in the interpretation of qualitative findings" (Creswell, Plano Clark, Gutmann, & Hanson, 2003). While the data gathered in this study was largely qualitative, the data were 'quantitized' (qualitative data converted into numerical codes for statistical analysis (Miles & Huberman, 1994)), in order to strengthen claims made based on the qualitative findings reported in Chapter 4.

The intention of this study is two-fold. Firstly, I aim to reveal, through analysis of the content of teachers' goals, the extent to which teachers' appraisal goals promote or prevent appraisal inquiry that focuses on student learning, and data about that learning. Secondly, I will investigate, through analysis of the characteristics of teachers' goals, the quality of goals and the subsequent likelihood of them improving performance. The main basis for considering the quality of goals will be Locke and Latham's goal setting theory.

Locke and Latham (1990) reviewed the results of five meta-analyses of goal setting effects. Empirical findings in the studies they analysed demonstrated a linear relationship between degree of goal difficulty and specificity and people's

performance in achieving the goal. According to them, goals should be specific rather than vague to avoid ambiguity and to avoid the potential for there to be doubt when analysing performance. Goals should be challenging, they say, because more difficult goals lead to greater effects on performance. While the meta-analysis was drawn from studies in a range of contexts, there is, Locke and Latham claim, overwhelming evidence that their goal setting effects findings are generalisable across a range of tasks, settings and subjects. It follows, then, that teachers' goals, when specific and challenging, are more likely to impact positively on both their own practice, and on the student learning that results from that practice.

Other writers, who have focused on the teacher appraisal context, also report the importance of goals. Piggot-Irvine's research between 1996 and 1999, relating to appraisal implementation and the impact of appraisal training, highlights, among other findings, that establishment of 'deep' rather than 'surface' appraisal objectives and improvement plans are a key to effective appraisal (2003b). Deep objectives involve an action research approach, in which the appraisee examines the current situation, plans changes and improvements, carries those out and evaluates their effectiveness through data-based reflection. Surface objectives, on the other hand, involve what Piggot-Irvine refers to as a "quick-fix approach to objective setting: one that is concerned with getting the objective out of the way as quickly as possible rather than focusing substantially on something that results in considerable improvement to learning and teaching" (Piggot-Irvine, 2003b, p. 175).

While the importance of teachers' goals and their potential for enhancing practice is widely acknowledged, there is also evidence that appraisal goals, in practice, are not utilised as a means for improving performance. In a study undertaken by Gratton (2004), none of the seven teachers interviewed about their perceptions of appraisal revealed a deep objective as the basis for his/her appraisal and several could not remember their objective. Most of the participants reported that development objectives were only referred to once or twice each year. Since it was generally the appraisal discussion that prompted teachers to refer to the objectives, they "were not

actively considered during the year as a whole" (Gratton, 2004, p. 295). Limitations of this kind were also reported by Down, Chadbourne and Hogan (2000) in relation to their research on the implementation of performance management in Western Australia. They concluded that "schools and teachers practiced a superficial compliance with the mandated aspects of policy, but did not engage with it at a deeper level" (p. 218). This was highlighted in a quote from one of the participants in their study:

Teachers are not putting forward things they really need to work on, they're not admitting to genuine faults, they find something that's popular, that's easily demonstrated, so it's strategic selection. Nearly everyone chose Information Technology because you could go from "I can't even turn on the computer" to "look I can use the Internet" in a very short time, and you've got all sorts of very clear evidence. I guess they're acquiring skills along the way, but it becomes an exercise in showcasing (Teacher interview, 10/98) (Down et al., 2000, p. 218)

Student learning focused teacher appraisal, based on evidence about teachers' impacts on students (the approach argued for in Chapter 2), would be limited by surface goals and superficial compliance, such as that reported by both Gratton and Down et al. Rather, student learning focused appraisal requires deep, significant objectives, which teachers commit to and engage with in a meaningful and productive way throughout the appraisal process. This study investigated the content and quality of teachers' appraisal goals and, particularly, their emphasis on student learning.

### *Method*

This section will outline the procedures for gathering data and the sources of data, as well as the approach taken to data-analysis. The focus question for this study was "to what extent do teachers' appraisal goals focus on student learning?"

### *Procedures*

Two of the schools involved in previous studies had expressed an interest in being involved in any subsequent research. When it became apparent that an investigation into teachers' appraisal goals was necessary, I made an application to the University of Auckland Human Subjects Ethics Committee (UAHSEC), which was approved, and approached the principals from those two schools. One agreed to participate. That principal also suggested that I attend a principals' cluster group meeting to invite other schools to be involved. All 10 of the principals at that meeting agreed to participate. Sampling for this study was 'opportunistic or emergent' (Patton, 2002, p. 240) since new opportunities were taken advantage of during data collection. The unforeseen opportunity of the principal's cluster group enabled the sample to emerge during the study.

Principals were provided with participant information sheets, a summary question and answer sheet about the study, a copy of the questionnaire and a consent form to be signed by the principal. They were also given enough copies of these documents for each of the primary, fully registered or experienced teachers. Since the study met the University of Auckland Human Subjects Ethics Committee's requirements for an anonymous questionnaire, individual participant consent was not required. The principals were asked to distribute the information to teachers, and for those willing to participate, the questionnaires. I suggested that a staff meeting would provide the most ideal opportunity to do this, so that questionnaires could be completed in a few minutes, and returned in self-addressed envelopes I had provided.

Principals were given an opportunity to ask questions about all aspects of the research. They were asked to look closely at the questionnaire, so that they could clarify any queries about how to complete it. They were 'walked through' each question, and I highlighted that the main question was 'what are your appraisal goals for this year?' I suggested that participation in this research could provide a valuable opportunity for teachers to focus on their appraisal goals by referring to the goals in their appraisal documentation.

### Data Sources

Responses were received from more than half of the 12 schools that were invited to participate. Responses were from seven of the 10 schools represented at the principals' cluster meetings, and from one of the two schools I approached directly at the outset. This study is based on data from 68 teachers across eight schools. On average, there were just over eight questionnaires returned from each of the schools that had participating teachers. Questionnaires were numbered to indicate those that had been returned from the same school, and also to indicate a questionnaire number from each school. This provided the researcher with a reference to the original data source. For instance, [2-7] indicates school number two, questionnaire number seven. The number of years' teaching experience and the class year level for respondents is outlined in Table 18.

Table 18 Characteristics of Questionnaire Respondents

Characteristics	Number of respondents	
Year teaching experience	<4 years	14
	4-6 years	8
	7-9 years	8
	10+ years	35
	Not answered	3
Students' year level	Year 1-2 (junior primary)	19
	Year 3-4 (middle primary)	18
	Year 5-6 (senior primary)	19
	Year 1-6 (all primary levels)	4
	Year 7-8 (intermediate levels)	7
	Not answered	1

The majority of teachers participating in the study (63%) had seven or more years' experience, and more than half had been teaching for in excess of 10 years.

The questionnaire (see Appendix D) included five key questions:

1. which year level band do most of your student fall into? (Y1-2, Y3-4, Y5-6, Y7-8)

2. how many years' teaching experience do you have? (<4, 4-6, 7-9, 10+)
3. what are your appraisal goals for this year? (the main question)
4. what focus do you expect the goal to be given in your appraisal interview? (only, main, minor, don't know)
5. who has been involved in deciding on these goals?

While the questionnaire included several closed questions, the most significant question ('what are your appraisal goals for this year?') was open-ended with a qualitative focus. Since the final two questions yielded no data of particular relevance, they will not be discussed here.

Questionnaires were considered an important means of gathering data about teachers' appraisal goals since they are characterised by a number of strengths as outlined by Johnson and Turner (2003, p. 306). They are inexpensive, quick to turn around, able to be administered to groups, and, most importantly for this study, are perceived by potential participants as being genuinely anonymous. This increased the likelihood of teachers recording actual appraisal goals without modifying them according to any perceived criteria for their quality. The risk of encountering vague answers in the open-ended items, a weakness according to Johnson and Turner (2003), was avoided by requesting that teachers refer to the written record of their appraisal goals in their personal appraisal file.

#### *Analysis of Appraisal Goals*

There were two main phases of analysis in this study. The first was underpinned by the appraisal approach argued for in Chapter 2. Categories were developed to address key aspects of that argument – inquiry; data; student learning outcomes; student process; and teacher process.

Table 19 Coding Rules - Appraisal Goals in Relation to Data-based Inquiry into Student Learning

Category	References to
Student (a particular individual or group) learning outcome	improvement; increased achievement; learning gains; student learning; student success; student achievement; student outcomes; learning gains; improvement; achievement (in a specific aspect of learning)
Student - behavioural	Student behavioural processes
Teacher process	Teachers' actions, activities, tasks,
Inquiry	see if; find out; check ; ask; query; question; look at; examine; explore; analyse
Data	data; evidence; assessment; data; results; scores; reading ages; percentages; assessment grades; achievement ratings; test results

There was potential, in the analysis, for a goal to be coded to more than one category. For example, a goal such as “implement shared reading approach on a daily basis to improve ‘Tiger’ group’s ability to use picture visual cues in predicting text most times they encounter an unknown word’ would be coded to ‘data’, ‘student learning outcome’ and ‘teacher process’.

Goals coded to the ‘student learning outcome’ category were analysed further. The analysis, according to the categories in Table 19, had established the content of the goals. The next step was to establish, for those relating to student learning, the quality of the goal. Locke and Latham’s extensive research about goal setting (1990), and their findings about the significance of specificity and challenge, provide the rationale for analysing the quality of the goals. I have developed categories that apply the notions of specificity and challenge in goals to the teaching and learning context. This approach is similar to that used by Brereton (2001) who analysed goals according to whether they were specific, clear, measurable, challenging and flexible. In trial analysis of the goals in the present study, it became apparent that classifying them simply according to whether they met those criteria or not, was problematic.

For example, the goal to “develop independent learners”, while focused on learners, is not specific or challenging enough to be likely to impact on a teacher’s practice or their students’ learning. It was necessary, then, to establish a coding system that addressed (for the goals which had already been categorised as including reference to student learning – see Table 19) codes for specificity and challenge, as outlined in Table 20.

Table 20 Coding Rules - Student Outcome Goal Specificity and Challenge

	Degrees of specificity and challenge	The goal refers to:
Specificity Criteria	Learner – specific	A particular student or group of students is stated
	Learner - general	A particular student or group of students is not stated
	Learning – specific	A particular aspect of a learning area
	Learning – general	Learning or learning area generally
	Extent of improvement – specific	The quantity, quality or frequency of improvement
	Extent of improvement - general	General improvement - the extent of desired improvement unclear
Specificity	Specific	All aspects (learner/learning/extent of improvement) stated specifically
	General	2 aspects (learner/learning/extent of improvement) stated specifically
	Vague	0-1 aspects (learner/learning/extent of improvement) stated specifically
Challenge	Challenge explicit	Outcomes that are demanding in their nature and extent, with challenging learning improvements explicitly stated
	Potentially challenging	Outcomes that are potentially challenging but (since wording of challenge is not explicit) where negligible improvement would also represent achievement of the goal
	Not challenging	Outcomes that are straightforward and undemanding, unlikely to be a challenge to achieve



The main purpose of this study was to investigate if appraisal goals are focused on student learning (in a way that will promote inquiry into student learning in subsequent stages of the appraisal process). It was expected that references to student learning, where apparent, would relate to learning in the essential learning areas of the New Zealand Curriculum Framework. The learning areas were used as a basis for the next phase of goal analysis.

Findings from Down et al.'s study revealed that teachers engaged in a superficial compliance manner with appraisal policy. In particular, Down et al. identified a tendency for genuine needs to be avoided by teachers selecting their appraisal goals in favour of goals that are easy to achieve and demonstrate. The reference they made to an information communication technology (ICT) goal being used in this way (Down et al., 2000, p. 218) prompted me to include ICT categories along with the essential learning areas of the New Zealand Curriculum Framework (Ministry of Education, 1993) in this analysis (see Table 21). The purpose of this analysis was to reveal the emphasis given to learning in or teaching of particular curriculum areas in appraisal goals, and (through inclusion of the 'Teachers ICT skills' category) the potential for a superficial compliance approach such as that identified by Down et al. I was particularly interested in establishing the relative emphasis on goals focused on a curriculum learning area and those about teachers' ICT skills.

Table 21 Coding Rules - Appraisal Goals and Their Focus on ICT and Learning Areas

Category	References to
English	English; literacy; reading; writing; language
Te Reo	Te Reo, Maori language
Mathematics	Maths; numeracy; numeracy project
Science	Science
Technology	Technology, but not ICT (see below)
Social Studies	Social Studies
The Arts	The Arts: dance; drama; music; visual arts
Health and Physical Well Being	Health; PE; fitness;
Teachers' ICT skills	Teachers' own skills, confidence, abilities, knowledge in relation to ICT
ICT	ICT implementation in class programmes; ICT with students; ICT for learning
ICT generally	ICT

Further analysis of the total set of goals was undertaken to ascertain the orientation of teachers' appraisal goals. The orientation of teachers' goals towards some kind of action was established by analysing the goals according to the presence or absence of a part of speech that encapsulates action, such as a verb. Three categories were developed for goals that encapsulated action – improvement oriented, task oriented and inquiry oriented. Those that did not encapsulate action were categorised as topic oriented. The four orientation categories are described in Table 22.

Table 22 Coding Rules - Orientation of Appraisal Goals

Category	References to
Improvement Oriented	improve; consolidate; upskill; develop; teach; refine; come to grips with
Task Oriented	oversee; create; implement; be part of; establish; cater for; promote; use; address; plan; coordinate; display; support; ensure; integrate; try; respond to; finish; complete; model; manage; present; maintain; motivate; provide; share; devolve; organise; get; place; work; assist; continue; attend; update; pursue; re-vamp
Inquiry Oriented	Observe; familiarise; look at; investigate; review
Topic Oriented	No verb – topic rather than action or outcome oriented

*Inter-rater reliability*

Twenty-four of the 244 goals were coded by a second rater to establish the reliability of my data analysis. The 24 goals were selected by highlighting every tenth goal on a spreadsheet. The second rater then coded the goals (by ticking the categories on the spreadsheet) based on the rules outlined in Table 19 and Table 21. The coding decision was to identify the goals that did and did not refer to:

- Student'/s' learning outcome/s
- Student'/s' behavioural processes
- Teacher process
- Inquiry
- Data
- English
- Te Reo
- Mathematics
- Science
- Technology
- Social Studies
- The Arts
- Health and Physical Well Being

- Teachers' information communication technology (ICT) skills
- ICT in the classroom
- ICT generally

In my analysis of the 24 goals that were also used for inter-rater reliability I coded them into 48 categories. Forty-three of those coding decisions were agreed by the second rater, giving 90 per cent inter-rater reliability on these aspects of appraisal goal analysis. Only one of the disagreements (2%) related to the student-learning category (the most significant category) in which I did not code a particular goal to student learning that the other rater did. The second rater also analysed the entire set of 11 goals that I had coded to the student outcomes category against the coding rules for specificity and challenge (see Table 20):

- Learner - specific
- Learner – general
- Learning – specific
- Learning – general
- Extent of improvement – specific
- Extent of improvement – general
- Challenge explicit
- Potentially challenging
- Not challenging

We had agreement on 32 of the 33 coding decisions for goal specificity (97%), and on all 11 of the coding decisions for goal challenge (100%).

## ***Results***

### *Appraisal Goals' Connections to Data-Based Inquiry into Student Learning*

The results of analysis of teachers' appraisal goals according to the key elements of the appraisal approach argued for in Chapter 2 are outlined in Table 23.

Table 23 Percentage of Appraisal Goals Focused On Student Learning and Based On Data

Goals' refer to:	Number of goals	
	N=244	Percentage
Student outcomes	11	4.5%
Student behaviours	2	1.0%
Teacher process	220	90.2%
Inquiry	9	3.7%
Data	0	0.0%

*Student outcome focused teacher appraisal goals*

Only nine of the 68 teachers who completed the questionnaire recorded an appraisal goal that related to student learning. These goals, (less than 4% of the total) are listed below: (throughout this chapter references to the school number and questionnaire/teacher number is given in square brackets)

- develop independent learners [1-6]
- developing literacy in Y1 [2-2]
- to upskill all the children in my class in the use of ICT [2-5]
- improve numeracy skills and teaching and numeracy thru NUMP [2-8]
- to upskill literacy at yr1/yr2 level (reading/writing) [4-2]
- improve written language throughout whole school (school wide goal) [4-4]
- develop cognitive and cooperative skills in Philosophy for Children creating a circle of inquiry [6-1]
- extending gifted and talented children in my classroom [6-5]
- develop and implement teaching practices which develop information literacy [8-1]
- students answering own questions [5-6]
- getting lower level students talking math [5-6]

Further analysis of this set of goals, since they represent those most closely aligned with the argument for student learning focused appraisal, was undertaken. These results will be outlined later in this section.

#### *Student process and teacher process focused goals*

Two teachers recorded goals that related to students' behaviour rather than learning outcomes. These goals were to "ensure ready student response to instructions" [3-9] and "encourage students to always do their best" [4-5]. The remaining 93% of goals did not make any explicit connection to either students' learning or their behaviour. The vast majority of goals (90%) were about elements of teachers' practice - either a teacher's behaviour, an aspect of their style, or a particular process or approach. In this category were goals such as "implement new arts curriculum" [1-8], "support teacher aides with autistic student" [3-8], "maintain a student focused physical classroom environment" [5-2], "make and use ENP resource throughout the year" [6-9].

#### *Inquiry*

Nine of the 244 goals were about teacher's own inquiry. Six of the nine categorized here were from one school, and they all related to reviewing the school science programme. The other three were to "look at management and content of the maths programme" [3-9], "investigate different learning styles" [3-10] and to "monitor professional reading" [5-7]. In no instance was an inquiry-oriented goal specifically related to inquiry into teachers' actual practice, or the impacts practices had on students.

#### *Data*

None of the 244 goals referred to data, or to specific measures by which the achievement of the goal could be established. This finding relates to the findings regarding specificity and challenge reported below.

### *Goal specificity and challenge*

Having established that 11 of the 244 goals included content about student learning (Table 23), it was important to consider those goals in terms of their degree of specificity and challenging. Each of the 11 goals was included on a checklist to identify how specific or general they were in their reference to learners, learning and extent of learning improvement. This checklist is included as Appendix E and the overall results are given in Table 24.

Table 24 Specificity of References to Learners, Learning and Extent of Improvement in Student Outcome Focused Teacher Appraisal Goals

Goal aspect	Specificity	Number of the student outcome coded goals (n=11)
Learners	Specific	6
	General	5
Learning	Specific	1
	General	10
Extent of improvement	Specific	0
	General	11

There was a fairly even split between the number of goals that were specific and those that were general in regard to the learners. References to specific learners included “Y1” [2-2], “gifted and talented children in my classroom” [6-5], and “lower level students” [5-6]. The way in which goals referred to the aspect of learning, and extent of improvement desired, though, was overwhelmingly general. The one specific reference to an aspect of learning was about the cognitive and cooperative skills children would use to create a circle of inquiry. Of the other 10 goals, two referred to learning generally, four referred to literacy, two to maths/numeracy, one to ICT and one to questioning. While nine of the 11 goals analysed here were about aspirations to “develop”, “upskill” and “improve” and “extend” children and their learning, not one made clear the desired extent of improvement. No specific reference was made to how much students might improve or how regularly they might demonstrate, for example, a skill. It was merely stated that they would “improve”, suggesting that any improvement, however small, was the goal. The

extent of improvement teachers were aiming for was vague and open to subjective interpretation.

Analysis of how specific or how general goals were for the three aspects (learners, learning, extent of improvement) enabled me to assign an overall specificity rating to each goal. They were classified as either 'specific', 'general' or 'vague'. Each one was also classified as either 'explicitly challenging', 'potentially challenging' or 'not challenging' (see Table 25).

Table 25 Specificity and Challenge in Student Outcome Focused Teacher Appraisal Goals

		Number of the student outcome coded goals (n=11)	Percentage of student outcome goals
Specificity	Specific	0	0%
	General	0	0%
	Vague	11	100%
Challenge	Challenge explicit	1	9%
	Potentially challenging	10	91%
	Not challenging	0	0%

*School Level Variable*

The 11 goals that referred to student learning were developed in just six of the 11 schools that participated in this study. Table 26 shows that three of those schools had more than one teacher focus on such a goal (the two student learning focused goals at school 5 were from just one teacher). School level variables are, then, likely to be important. It is important to note, that the 91% figure for potentially challenging



goals relates only to the 11 goals (out of the total 244) that were analysed for challenge.

Table 26 Number of Student Outcome Focused Appraisal Goals by School

School number	Number of student outcome appraisal goals
1	1
2	3
4	2
5	2
6	2
8	1

*Orientations of Appraisal Goals*

Where goals were stated with a verb, that verb indicated the orientation of the goal to a task, to inquiry or to improvement. A topic orientation was apparent in goals stated without a verb. All 244 goals were analysed to establish their orientation, and the results are summarised in Table 27.

Table 27 Orientation of Teachers' Appraisal Goals

Goals	Number of goals N=244	Percentage
Improvement oriented	48	19.7%
Teacher improvement	39	16.0%
Student improvement	9	3.7%
Task oriented	127	52.0%
Inquiry oriented	9	3.7%
Topic oriented	60	24.6%

While nearly 20% of teachers' goals were improvement oriented, these mostly related to teacher improvement unrelated to student improvement. Examples of these include, "Improve/consolidate/upskill my ICT knowledge" [1-8], "develop own level of Te Reo" [1-5], and "improve facilitation skills e.g. taking staff meetings" [2-1]. Only 10 of the 68 teachers, 15% of the sample, had goals focused on student learning, and in particular improvements in learning. These include those goals listed on page 154.

More than half of the goals were task oriented, relating to what the teacher would 'do'. They comprised tasks of a short-term 'quick-fix' nature, as well as those likely to occur over an extended period of time. These included "to have a re-vamp of the co-ordinator file" [2-10], "to manage the junior syndicate effectively" [4-1], "to learn to use the digital camera and to print out images on my own" [6-4] and "provide a safe, stimulating classroom environment" [4-5], "develop maths profile for data collection" [4-4] and "attend writing and reading courses" [7-9]. Many of the task oriented goals were particularly unchallenging. Some reflected tasks that might be generally accepted as fundamentals of good practice, and not aspects of practice that should need addressing through an appraisal goal. For example, "encourage students to always do their best" [4-5], "to place all new students to the school" [5-7] and "consider our outdoors facilities, e.g. gardens, care of plants, trees" [7-1].

The category with the fewest goals, only nine of the 244, reflected aspirations to inquire in some way. Interestingly, none of these was based on inquiry into student learning, and were not even directly about teachers' practice. In nearly all cases they were about inquiry into the broader aspects of a programme. Goals of this type included "review science programme" [7-3], "look at management and content of maths programme" [3-9] and "refresh/review the classroom writing programme" [7-1]. Two of the nine inquiry oriented goals were about investigating a learning styles approach. The potential for impacting positively on teachers' and students' learning through a goal such as "review science programme" is severely limited since "reviewing the science programme" is not specific or, necessarily, challenging in

terms of improving student learning. There is a vast range of approaches to reviewing a science programme, and the degree of rigour, scope, and impact of reviews could vary enormously. With a general goal statement like “review science programme”, though, even the lowest quality programme review would be sufficient, for appraisal purposes, to deem the goal met.

One quarter of the teachers’ goals were topic oriented. These did not include a verb about what the teacher would improve, do, or inquire into. Rather, these goals were stated as a general context or topic. Examples include “time management” [6-8], “ICT” [3-4], “behaviour management” [5-4], “stress management” [7-8], “De Bono’s Thinking Hats’ [3-6], and “special needs” [5-4].

*Learning Areas and Information Communication Technology in Teacher Appraisal Goals*

Table 28 summarises results of the analysis according to the curriculum areas that appraisal goals referred to, and also their focus on Information Communication Technology (ICT). Only 66 of the 244 goals (27%) included reference to one of the curriculum areas of the New Zealand Curriculum Framework or to Te Reo.

Table 28 Appraisal Goals Focus on Curriculum Areas and ICT

Learning Area	Number of goals	Percentage
English	13	5.3%
Te Reo	7	2.9%
Mathematics	27	11.1%
Science	6	2.5%
Technology	0	0.0%
Social Studies	0	0.0%
The Arts	10	4.1%
Health and Physical Well Being	3	1.2%
Teachers’ ICT skills	13	19.1%
ICT		
ICT in the classroom	12	17.6
ICT generally	6	8.8%

The figure below shows the relative frequency of mention of learning areas and aspects of ICT referred to in appraisal goals.

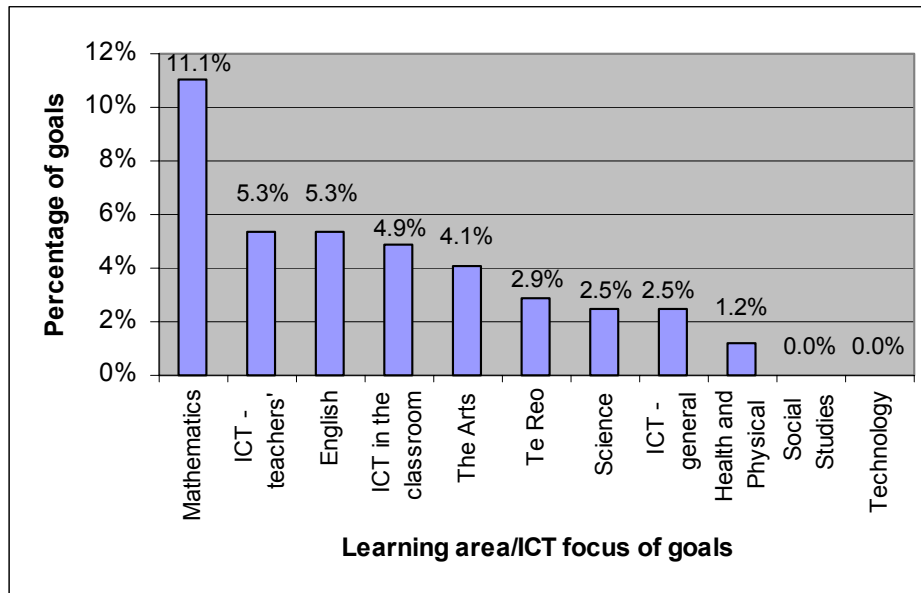


Figure 21 Percentages of Goals Focused on Learning Areas and ICT

Only one learning area, mathematics, was referred to in more goals than teachers' ICT skills. Nearly half of the teachers (31/68) had an ICT goal. Most related to teachers' own 'upskilling' (13), nearly as many (12) related to the integration of ICT into classrooms, and six were stated generally as a topic (ICT). The technology figure of zero indicates that there were no references to technology beyond ICT.

### *Discussion*

The implications of the findings about teachers' appraisal goals outlined in the previous section will be discussed here. In particular, I will address the lack of focus on student learning in appraisal goals, the quality of goals in terms of their specificity and challenge, and the potential for goals to facilitate appraisal that is focused on student learning and informed by evidence.

### *Lack of Focus on Student Learning*

Student outcomes were referred to in less than five per cent of the appraisal goals analysed in this study. In contrast, more than 90% of goals related to a teacher process, or something a teacher would do. This signals a lack of concern for investigating the links between teacher practices and impact on student learning. The finding that three schools (of the six) had two or three teachers who had goals that referred to student outcomes, indicates some potential for school characteristics, such as school culture, as influences on the nature of appraisal goals.

The lack of focus on student learning evident in teachers' appraisal goals suggests that appraisal matters are thought of quite broadly and are considered, in practice, an opportunity to deal with a wide range of issues, practices, ideas, tasks and topics of interest. Appraisal goals rarely reflect the pedagogical imperative, the professional expectation or the ethical obligation (see Chapter 2) to make student learning a fundamental concern of all aspects of a teachers' work. Furthermore, the lack of focus on student learning in appraisal goals conflicts with recent moves in the wider educational context to emphasise student learning. The national education priorities require all providers to "be much more deliberate, explicit and strategic about what improvements in students' learning they will achieve and the basis from which they are making these judgements" (Ministry of Education, 2001c).

Setting targets is a key aspect of planning and reporting requirements for schools in New Zealand. From 2003 schools have been required to submit charters to the Ministry of Education, which include specific targets for improving student achievement in the year following (Ministry of Education, 2002b). In a strategically aligned environment, at least some of teachers' goals should reflect their schools targets. Schools are also encouraged to "investigate and respond to any differences between their planned improvement targets and the actual outcomes achieved by students" (Ministry of Education, 2003g, p. 4). This "analysis of variance" process takes into account the target, the outcome (or what happened), analysis and evaluation. Appraisal goals could provide the basis for a parallel or at least similar

process in relation to the practice of individual teachers (Ministry of Education, 2003g, p. 4).

The potential for a process like this is highlighted in a study reported by Joyce and Showers (2002). Goals were an important means of maintaining focus on student achievement in an initiative that succeeded in significantly improving the amount of independent reading students engaged in. Data were used to examine the extent to which goals were achieved through the initiative, a cultural climate conducive to the goal achievement was established, analysis of goal achievement was undertaken, and achievement of goals was celebrated.

The findings also highlighted the prevalence of goals relating to teacher learning as opposed to student learning. While teacher learning, particularly in relation to student learning, is not to be discouraged, the nearly 20% of teachers with a goal relating to their own ICT skills (without reference to the classroom) was surprising. A focus on student learning in appraisal goals was also hindered by the lack of reference to learning areas. Only just over a quarter of the goals related to a particular learning area (English, Te Reo, Mathematics, Science, The Arts and Health and Physical Well Being). More explicit reference to learning areas, and their curriculum statements, when developing appraisal goals may be an important strategy for promoting a focus on student learning.

This study confirmed, then, through an enumerative type of induction (Miller, 2003, p. 428), at least partial generalisability from the discussion study findings (see page 112). Through the accumulation of further examples of goals that are unrelated to student learning, a hypothesis can be made about a wider tendency for professional norms that do not emphasise student learning-focused appraisal goals.

#### *Lack of Specificity and Challenge*

As previously mentioned there is a strong body of evidence that suggests that “goals that are specific and difficult lead to a higher level of performance than vague, non-

quantitative goals such as ‘do your best’” (Locke & Latham, 1990, p. 29). Teacher performance management guidelines, though, rather than stressing the potential of challenging and difficult goals, stress a need to avoid unrealistic goals – “It [performance management] is not about setting people up to fail by setting unrealistic goals” (Ministry of Education, 1998c, p. 4). While it is possible for a goal to be both realistic and challenging, the emphasis given to just the first of those aspects does not encourage constructive appraisal goals. It is not surprising, given guidelines such as these, that the appraisal goals in this study were not found to be particularly challenging. More recent documentation, pertaining to achievement targets, reflects less emphasis on being realistic (potentially interpreted as unchallenging), and more emphasis on targets that are challenging – “Setting challenging targets for students’ achievement gives focus and meaning to a school’s planning and teaching” (Ministry of Education, 2001c, p. 3).

Even goals that did include an element of student learning lacked specificity, so the potential for improving student learning was minimised. While half of the student learning goals were specific in terms of the learners the goal related to, with respect to the aspect of learning and the desired extent of improvement, in nearly all cases they lacked specificity in relation to learning or to the extent of improvement targeted. This led to an overall specificity classification of all 11 student-outcome goals as vague rather than general or specific. Coe’s explanation of what it means for a goal to be specific highlights the limitations of general goals:

You will know whether you have achieved it – there is no room for doubt or interpretation. This specificity must therefore include information about when something must be achieved, as well as precisely what. The achievement of the target must be verifiable by some objective or accepted criterion. (Coe)

Coe’s ideas about developing specific goals prompts many questions about the appraisal goals reported here. For instance, in relation to the goal “consolidate Maths programme” [1-8] – at what point would you know that the Maths programme is consolidated? Are there objective and accepted criteria for determining how ‘consolidated’ the Maths programme has become? How might different people interpret ‘Maths programme consolidation’?, and what aspect of the Maths programme is to be consolidated? The appraiser missed the opportunity,

however, to prompt this teacher [1-8] to consider those questions as they developed the goal. This might have led to the goal itself being framed in a specific and challenging way, so that it could be a constructive goal for an appraiser and appraisee to engage with, throughout the appraisal cycle.

Related to the finding about a lack of specificity in teachers' appraisal goals, was the complete absence of references to data. None of the 244 goals mentioned data, evidence, assessment, results, scores, reading ages, percentages, assessment grades, assessment comments, achievement ratings or test results. While 48 of the total set of goals, and nine of the 11 student outcome related goals, referred to improvement they did not suggest that any of the data sources listed about might provide evidence of the improvement.

Ten of the eleven appraisal goals about student outcomes were also limited because the outcomes they referred to were only *potentially* challenging, but not *necessarily* challenging. Lack of explicit reference to the challenge would allow even negligible improvements to deem a goal achieved. For example, the goal to "develop literacy in Year 5" is undoubtedly a goal with potential to be challenging, if for instance it refers to a goal to increase the reading ages of all boys in the class (currently 80% of whom are reading at an age 1-2 years below their chronological age) to at least their chronological age, by the end of term three. Achieving this would certainly indicate that literacy for Year 5 students had been developed. However, the vagueness of the challenge in the "develop literacy in Year 5" goal means that even the slightest improvement for just one of the students, perhaps learning to scan the text without pointing one to one with a finger, could also technically deem that the goal has been met. While the second example of an improvement given here would indeed be important for that student, it is significantly less challenging than the first. Unless appraisal goals clearly state the challenge, evaluation of the extent to which goals are achieved will not be possible.

The findings of this study suggest that there is the potential for significant improvement in the quality of teachers' appraisal goals. To use a hurdling analogy, jumping the first hurdle (having a focus on student learning) is not enough. The



second hurdle (making the goal specific) needs to be jumped too, and the hurdles must also be challenging. Otherwise, the teacher and their learners minimise their chance of finishing the race well. Just as a professional athlete would not expect to jump hurdles at ankle height, professional educators need to expect to jump hurdles that involve appropriate challenge in terms of their practice, their learning and their students' learning. There are also other hurdles (connected to other aspects of the performance management process) that involve working towards achieving the goal – engaging in professional development to support improvement in practice, collecting, analysing and inquiring into the extent to which the goal has been achieved, and so on.

This study has revealed a predominance of task and topic oriented goals. These do not reflect the characteristics of goals advocated by Cwikla (2003), who claims that “developing explicit learning goals for teachers that focus on a depth of understanding rather than discrete teacher behaviours will help guide.... teachers who are focused on improving their classroom practice” (p. 50). Most of the goals in this study, unfortunately, did encourage discrete behaviours rather than deep understandings about improving practice. Goals such as “to place all new students to the school” [5-7] and “completing unit plans before end of term” [7-9] are examples of discrete behaviour focused goals. These goals may reflect a challenge for the particular teachers involved and be appropriate to their needs. It is concerning, though, given that teachers had an average of 3.5 goals each, that these were the goals considered most pressing. It is questionable whether goals such as these deserve the focus of a yearlong developmental appraisal cycle, or if they should be addressed through other means.

#### *Possible Explanations*

There are a number of possible explanations for there being only scarce reference to student learning in appraisal goals, and for the lack of specificity and challenge in goals. Two that will be discussed here are the notion of superficial compliance and norms of professional autonomy and supportive collegiality

### *Superficial Compliance*

The appraisal goals analysed here suggest an appraisal system that is engaged in with only superficial compliance. Superficial compliance, a notion mentioned at the beginning of this chapter (Gratton, 2004), reflects a perception of appraisal, and appraisal goals in this instance, as a tedious process in which teachers go through the motions in order to comply with mandated requirements. A superficial compliance approach to appraisal goals would explain the predominance of teacher process rather than student outcome goals, of vague rather than specific and challenging goals, and task rather than improvement oriented goals. The goals reported in this study allow teachers and schools to show compliance with the mandated policy, but prevent them from engaging with the goals in ways that will impact positively on teaching and learning. Appraisal goals that are specific, challenging, and oriented to improving student learning would also require compliance, but they would require deep, rather than superficial, engagement with data, and constructive dialogue to establish the extent to which goals had been met.

### *Norms of professional autonomy and supportive collegiality*

Teachers' autonomy and a penchant for collegiality explain, at least to some extent, the findings about teachers' appraisal goals reported here. Timperley and Robinson (2000) highlight two norms that can limit the effectiveness with which educators engage with each other that are relevant here. The first is a norm of professional autonomy that can reduce educators' capacity to hold each other to account for their practice. Cwikla (2003) agrees and suggests that the nature of teaching as a province of individual teachers (professional isolation), rather than a public activity open to examination, is a possible explanation for teachers' learning goals remaining vague and unspecified. The second norm highlighted by Timperley and Robinson is one of supportive collegiality that can prevent difficult issues being engaged with and ensure that conflict is avoided. Norms of both professional autonomy and supportive collegiality could bring about uncritical acceptance of teachers' appraisal goals, despite their lack of reference to student learning, their lack of specificity and lack of challenge.

The warning about unrealistic goals in the Ministry of Education performance management guidelines themselves could be seen to support norms of supportive collegiality. It is suggested that performance management should occur “in supportive working environments where there is a high level of communication and trust. It [performance management] is not about setting people up to fail by setting unrealistic goals” (Ministry of Education, 1998c, p. 4). I agree that goals should not be unrealistic but would argue that emphasis should be given to goals that are going to make the most difference for teachers and learners, that is, challenging and specific goals. Focusing just on realistic goals may lead to goals that stray further from challenge than they need to, in order to be realistic. Rather than focusing on an approach that will avoid teachers experiencing a sense of failure, focus should be given to the kind of goals that are most likely to avoid failure for students.

Figure 22 shows how the aspects of superficial compliance, norms of professional autonomy and supportive collegiality and bureaucratic perceptions of appraisal work together to minimise the extent that appraisal focuses on student learning. In Figure 23, alternatives are suggested that could maximise the potential of goals to contribute to improvement in teaching and, most importantly, student learning.

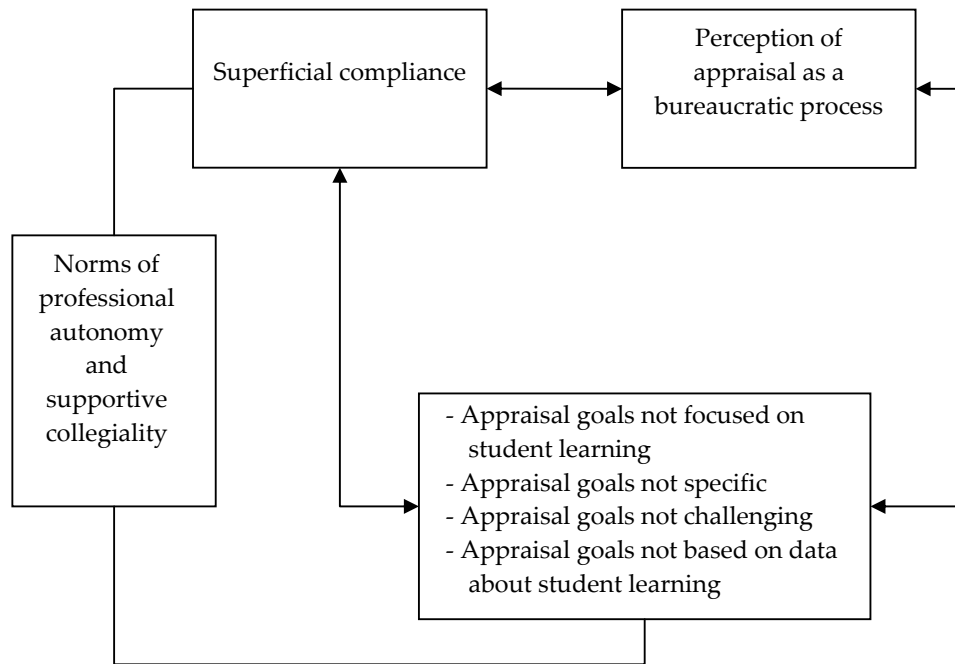


Figure 22 Appraisal Goals' Minimising Positive Impact on Student and Teacher Learning

An alternative, one that I argue will promote improvements to teaching and learning through particular appraisal system characteristics is presented in Figure 23.

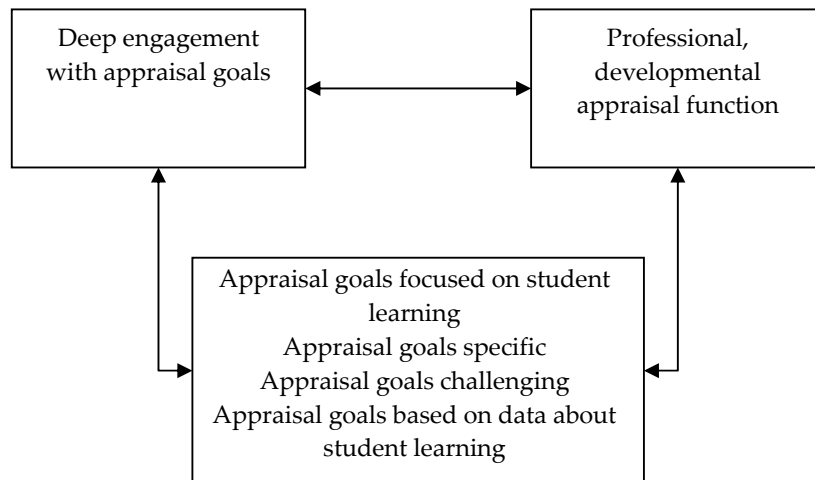


Figure 23 Appraisal Goals' Maximising Potential to Impact on Teacher and Student Learning

*The Risks*

Having given much attention to the potential for and benefits of student learning focused appraisal goals that are specific and challenging, I will give attention here

also to some potentially less positive aspects of this approach. Hopkins (2001) alerts educators to the prospect of well-intentioned targets for improving achievement resulting in undesirable tactical responses:

The emphasis on target setting in many educational systems puts pressure on schools and teachers to raise levels of achievement in the short term. This widespread tendency often results in a tactical response such as homework clubs or teaching to the test, and leads to a reductionist and impoverished interpretation of what constitutes learning. Powerful learning is more than just results and scores, it subsumes a range of cognitive and affective processes and outcomes (Hopkins, 2001, p. 71)

While the above quote relates to target setting at the school level, there are potentially similar issues for goal setting at the teacher/classroom level. Hopkins' point is similar to, though less extreme than, the notion put forward by Schweitzer, Ordonez and Douma (2004) who suggest there is a "dark side" to goal setting, that is, the role goals may play in motivating unethical decision making. While they acknowledge the substantial literature on the benefits of setting goals, they warn that they should be used cautiously. They found that people close to achieving a specific goal are more likely to engage in unethical behaviour in regard to demonstrating achievement of the goal. While the scenario and laboratory studies they undertook were not based in an educational context, these findings do have implications for the use of teachers' goals. In particular, the principle of "flexibility" (Brereton, 2001) in appraisal goals would be important. That is, there needs to be acceptance of the possibility that goals may need amending, or replacing, during the year according to circumstances.

#### *Conclusion*

In this study, goals were extremely unlikely to include references to student learning outcomes and, even when they did, they tended to be vague rather than specific, and potential challenges were not explicitly stated. The lack of focus on student learning is unlikely to have come about because teachers are not concerned about student learning or because they do not have aspirations to improve the impact they have on their learners. On the contrary, as discussed in Chapter 2, there is evidence that

educators do intend for appraisal to improve teaching and learning. There are constraints, though, in the practice of appraisers and appraisees that limit those intentions being met. It is hoped that this study might increase the extent to which educators focus on the fundamental goals for student learning, and state these in specific and challenging terms, as they develop and engage with appraisal goals.

## CHAPTER 6

### Appraisal for Learning

The approach to appraisal argued for in the preceding chapters – appraisal focused on inquiry about student learning and based on data – is fundamentally about ‘appraisal for learning’ (as it is referred to from here). ‘Appraisal for learning’ positions appraisal as an important context for teacher learning and recognises the significant potential teachers have to effect improvement in students’ learning, as a consequence of their own learning. In order for that learning to take place educators must focus their attention, throughout the appraisal process, on student learning. Empirical studies reported in chapters 3, 4 and 5, though, all reveal that only limited attention was given to student learning in several critical elements of the appraisal process (indicators, goals, discussions). Furthermore, references to the examination of data about student learning in these elements of appraisal were almost non-existent. On a more positive note, responses to the suggestion that appraisal should focus on data about student learning (see Chapter 5) received overwhelming support from both appraisers and appraisees. This signals that the argument for evidence-based inquiry into student learning in teacher appraisal has not only a theoretical basis, but also has validity in terms of teacher’s responsiveness and connection to the idea.

This chapter will outline the contribution of this research, including the implications that arise from the studies for further research and for practice. It goes on to discuss policy issues that arise from the research, and suggestions for how the gap between current practices and ‘appraisal for learning’ might be closed. I propose that notions of alignment and capacity building are central to understanding how to improve appraisal policies and practices in the New Zealand school system.

### *The contribution of this research*

This research has emphasised the importance of practitioners focusing on student learning as they engage in data-based inquiry during teacher appraisal – ‘appraisal for learning’. It has highlighted that, in the main, opportunities for learning in teacher appraisal have been missed. In doing so, it is hoped that some light has been shed on how those opportunities might instead, in the future, be captured.

Focusing on student learning and engaging in data-based inquiry are clearly not, in themselves, new concepts in the educational literature. The contribution of this research has been to examine how these qualities of educational practice are, and how they could be, utilised in the teacher appraisal context in New Zealand. ‘Appraisal for learning’ has the potential to enable teacher appraisal to achieve what is claimed by many to be its key purpose – to improve student learning (Bailey, 1997; Battersby, 1991; Irons, 1993; Leithwood & Louis, 1998; Ministry of Education, 1998c; Stewart, 1997; Townsend, 1995). The triangulated studies reported here, though, collectively provide evidence that appraisal practices in many schools currently, are unlikely to achieve that purpose, since relatively little attention is given to student learning or to the examination of achievement data in appraisal. Furthermore, the little attention that is given to student learning is typically obscured by attention to other matters, particularly those relating to teachers’ practices (without attention to the impact on student learning).

It is anticipated that this research will alert educators and policymakers to the, as yet unrealised, potential for appraisal to contribute to the fundamental goal of education. A key purpose of this thesis is to prompt educators and policymakers to engage in dialogue around the notion of ‘appraisal for learning’, an approach that is centred on improving student learning and necessitates data-based inquiry into student learning. Many of those involved in this research responded positively to the suggestion of this approach – one that is more focused and rigorous than typical



appraisal processes and likely to generate greater learning for teacher learners and student learners alike.

While the message is clear that attention should be given to the policy and practice of teacher appraisal, limitations to the studies should also be noted. The studies were of the permanent records of appraisal and reports from individuals of what happened in their discussions. While attempts were made to enhance participants' recall through reference to documentation from their actual discussion, the data presented here are inevitably indirect assessments of the content of appraisal discussions. This does limit, to some extent, the claims that can be made since the participants' retrospective recall may not entirely accurately represent what occurred in their appraisal discussion. That recall does, though, provide an insight into the enduring aspects of the appraisal experience. A further limitation was that information on the school-wide appraisal culture was not accessed. Hence, it is hard to comment upon the cultural context of the schools that explain the findings.

#### *Implications for further research*

This research has provided findings and explanations that have, in turn, generated further questions that would be relevant to pursue. The most important area for further research, and a potentially methodologically complex one, would be to examine the impact on student learning of an 'appraisal for learning' approach. Examination of the issues around implementing 'appraisal for learning' would also be of interest, in particular a comparative study to explore implementation in both schools functioning, and not functioning, as professional learning communities.

Findings from each of the studies relating to particular elements of the appraisal process could also be developed into further research topics about, for instance:

- the ways in which specific and challenging goals about student learning impact on other elements of the appraisal process and on teaching and learning.
- the ways in which educative appraisal based on productive reasoning contributes to appraisal discussions focused on data about student learning
- the ways in which schools might align complementary initiatives that are focused on student learning and achievement data and the impact of this

- the ways in which schools might address professional standards within an 'appraisal for learning' framework
- the capacity educational leaders have to effect the personal, interpersonal and organisational capacities central to implementing effective 'appraisal for learning'

*Implications for practice*

While further research would make a valuable contribution to the knowledge base around teacher appraisal, there are also practical implications, which could potentially be implemented immediately. It is hoped that the normative 'appraisal for learning' model, presented theoretically here, might provide a useful basis for school leaders and teachers to review their policies and practices. Schools could use the elements of 'appraisal for learning' to examine their appraisal policies and indicators and other documentation against, and practitioners could use these to reflect on their appraisal goals, observations, discussions and reports. This examination could bring about a number of shifts in practice in which educators:

- refer to student learning data as a basis for setting (or revising) appraisal goals.
- critique each other's appraisal goals against criteria relating to their focus on student learning and their specificity and challenge.
- incorporate shorter term appraisal cycles within the annual accountability framework as a means of regularly discussing the impacts of teachers' practices on their students' learning.
- establish a process for collecting and collating evidence in relation to appraisal goals (such as teacher portfolios).
- engage in appraisal discussions for which there is agreement that the giving of positive feedback is not the overriding purpose of the discussion.
- gather and manage assessment information in ways that enable educators to engage with information that is relevant to each teacher's learners and teaching/learning context.
- include on staff meeting agendas opportunities for professional dialogue about 'teacher learning about student learning' that has resulted from an appraisal experience. Ultimately, these discussions might also, given the appropriate professional learning climate, refer to teachers' learning about inadequate teaching practices and the implications that learning has for shifts.
- make shifts in appraisal approaches as a means of developing or enhancing a professional learning community.
- reflect a focus on student learning and an evidence-based approach in appraisal documentation

Shifts such as these should also, clearly, be reflected in the policies and procedural guidelines that are associated with appraisal in schools. It is also hoped that this work might impact at a wider policy level. It might, if it achieves its purpose, impact on the wider policy context for appraisal in New Zealand, and inform dialogue around future revisions to performance management policy.

## *Policy Issues*

A notable reason for the lack of focus on student learning and evidence based inquiry in teacher appraisal is that the education policy context in New Zealand does not emphasise such a focus. Official appraisal documentation, published mainly in the late 1990s, does not reflect more recent and widespread evidence that teachers are a greater source of variance in students' achievement outcomes than the school (Hattie, 1999b, 2002; Ministry of Education, 2004a). Hattie advocates the importance of considering magnitude of influence, and particularly the relative magnitude of influences. This suggests that the focus in current education policy and school level initiatives should also be applied to other initiatives, directly relating to and involving teachers, such as appraisal. National appraisal policy, since it relates to teachers (whose impact on learners is greater than that of the school itself), should also require the focus on student learning based on evidence that is promoted in school planning and reporting policy (Ministry of Education, 2002b, 2003c, 2003d, 2003e, 2003f, 2003g). Currently, teachers are positioned in the policy as somewhat disconnected to students and their learning, and tend, as a result, to be appraised against their fulfilment of certain responsibilities relating to their own behaviour and performance, without regard for whether that behaviour or performance promoted student learning. It is not surprising, then, that the studies reported here did not reveal appraisal practices that encourage educators to explore connections between teaching practice and student learning. Appraisal, in practice, is focused on a compliance function, to provide minimum quality assurance within an accountability framework, rather than on improving the learning of all teachers (and their students). Current policy does not meet the needs of teachers for whom minimum quality assurance is not an issue.

### *Closing the Gap between Current Practice and the Potential for 'Appraisal for Learning'*

Appraisal policies and practices, at national and local school levels, are in serious need of modification. They should at least include, and ideally be centred on, three important aspects. Firstly, the improvement of student learning should be explicitly

expressed as a key purpose for teacher appraisal. Secondly, appraisal inquiry should provide appraisees and appraisers with opportunities to engage in critical dialogue about teachers' practice in relation to their students' learning. Thirdly, that dialogue should be based on and informed by evidence about student learning. This notion of data-based inquiry into student learning in teacher appraisal is a notion that made sense to, and was well-received by, practitioners who were interviewed about the prospect. It is timely, nearly 10 years since the requirement for teacher appraisal in schools was mandated, for renewed focus and rigour in teacher appraisal. Just as the need to reframe, refocus and revitalise the New Zealand curriculum has been addressed, through The Curriculum/Marautanga Project (Ministry of Education, 2005), performance management systems could benefit from being reframed, refocused and revitalised.

Such change is likely to evoke dilemmas and challenges that are typical of appraisal (Cardno, 2001). The potential, though, for student learning focused and evidence-based appraisal to impact on teachers' practice in ways that enhance student learning, will make those dilemmas and challenges worth enduring.

#### *Closing the Gap Through Alignment Within Appraisal*

The findings reported here emphasise the significance of alignment, and misalignment, to the way in which appraisal is enacted in schools. Both alignment and misalignment were problematic. Well-intentioned policy statements (those that referred to improving student learning) were not aligned to at the level of performance indicators, discussions and goals. In this case, misalignment was problematic. Alignment was also problematic – goals that were not related to student learning were the basis of most appraisal discussions. To address this issue I propose that all elements of the appraisal process should be aligned, and should be focused on data about student learning (see below).

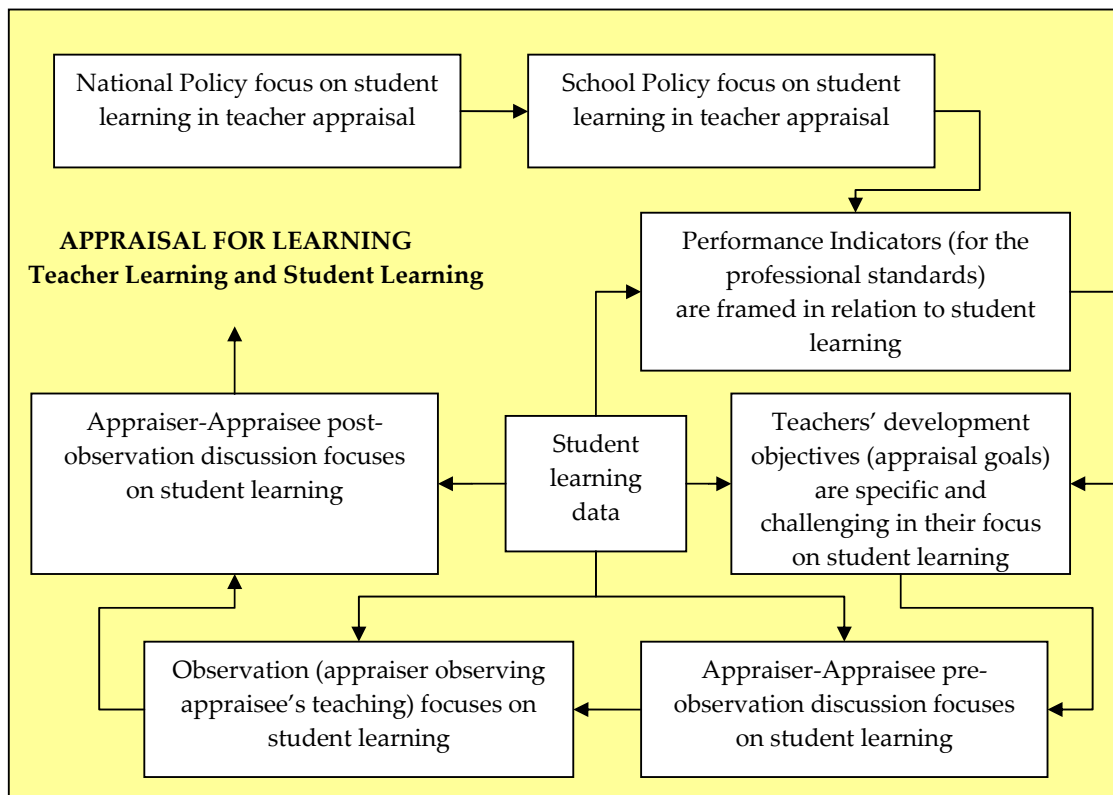


Figure 24 Appraisal for Learning Alignment Model

The strength of the model above is the coherence across the elements of the appraisal process. The repetition of references to student learning serves to maintain the focus on student learning beyond the level of policy rhetoric. The consistent focus on student learning and the use of data limit the potential for policy implementation slippage, making explicit, throughout the process, the purpose of appraisal. This focused and uncomplicated model seeks to avoid the risk, inherent in more complex models of appraisal (those that encompass multiple domains of practice, indicators, and criteria, for example), of practitioners selecting only certain aspects and rejecting others.

*Closing the Gap Through Alignment Beyond Appraisal*

In addition to aligning elements of an appraisal system to each other, ‘appraisal for learning’ should also align to educational initiatives beyond appraisal that are complementary in their focus. Stronge and Tucker describe this as “systemic

compatibility". They suggest, though, that "teacher evaluation is often approached as a discrete, isolated activity that is disconnected from the central purposes and personnel activities of a school system, and yet performance evaluation achieves relevance only by becoming an integral component of other system-level initiatives" (1999, p. 347). In New Zealand, there are a number of other initiatives (that focus on student learning and promote an inquiry stance that is based on data) which appraisal could be connected to. These include recent trends and research on professional development (Timperley, Phillips, & Wiseman, 2002b; Timperley & Wiseman, 2002), school planning and reporting policy (Ministry of Education, 2002b, 2003d, 2003e, 2003f, 2003g) and the national assessment strategy (Ministry of Education, 2001a). Links that already exist between these initiatives are represented by the black arrows in

Figure 25. Potential linkages between appraisal and the other initiatives (that are already in place) are represented by the grey lines.

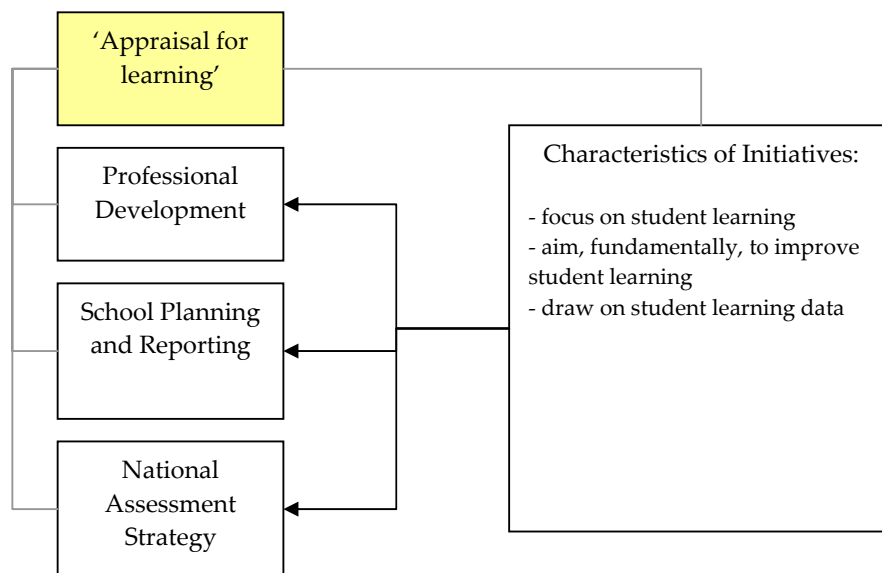


Figure 25 Aligning 'Appraisal for Learning' to Other Initiatives

Education practitioners have already been challenged to focus explicitly on student learning and student learning data as a means for improving teaching and learning. This focus is required in their approach to assessment (National Assessment

Strategy); it is required in their school achievement targets, annual reports and analyses of variance (School Planning and Reporting). It is also an increasingly common feature in professional development programmes. There has, then, already been a significant degree of capacity building, which would support the application of characteristics outlined in the figure above.

Shifts in appraisal policy and guidelines could be made more manageable for those implementing the process by explicitly aligning to these other initiatives. That alignment could be evident in the terminology used and also by making connections to these processes, such as those outlined in the following examples.

*Aligning to school planning and reporting policy and practices*

New Zealand's 'Planning for better student outcomes' initiative encompasses the suggestions made by Hopkins and MacGilchrist in relation to development planning. They suggest development planning should be used as "a strategy to enhance directly the progress and achievement of students. The crucial difference between this and previous approaches to development planning is that it is rooted in classrooms. The focus is on students' learning, their progress and achievement, what is needed to improve it and how this is best supported. The plan begins with learning goals for students" (Hopkins & MacGilchrist, 1998, p. 410). The strategies for focusing on outcomes outlined by the Ministry of Education (Ministry of Education, 2003d) already being utilised in schools (see the left-hand column of Table 29), could be paralleled in teacher appraisal in ways that promote 'appraisal for learning' (see the right hand column of Table 29).

Table 29 Connecting Teacher Appraisal to the Outcomes Focus in School Planning and Reporting Policy

Strategies for Focusing on Outcomes →	Teacher Appraisal Parallels: Links to 'Appraisal for Learning'
Making learning outcomes central to all debates about education and focusing on the impact of what we do on learning	Appraisal could 'focus on the impact of what we do on learning'
Requiring all providers to be much more deliberate, explicit and strategic about what improvements in student learning they will achieve and the basis from which they are making these judgements	Teachers' appraisal goals could deliberately, explicitly and strategically relate to improvements in student learning
Increasing the systematic use of information to continuously improve institutional performance	Information about student learning could be used to improve not only institutional, but also individual teacher's performance
Developing leadership capacity for outcomes-focused management across the education system, especially in the schooling sector	Appraisers' capacity for outcomes-focused appraisal could be developed since it is a key management task.
Implementing a planning and reporting regime in the schooling sector	Aspects of the appraisal process could support or complement the planning and reporting regime

*Aligning to the national assessment strategy*

Similarly, the purposes, processes and uses of assessment information, as outlined in the national assessment strategy framework, could be broadened to encompass appraisal. The summary of the national assessment strategy (p. 44) highlighted the missed potential to integrate teacher appraisal into the framework for assessment. The table below outlines how appraisal could be integrated into the assessment strategy in ways that make clear for educators how assessment data can contribute to 'appraisal for learning'.



Table 30 Integrating the Appraisal Context into the Assessment Strategy Framework

	The purpose of gathering assessment information	Assessment information is needed by	Tools and Processes	Uses of assessment information
Appraisal Context	<p>Provide learning opportunities for teachers;</p> <p>Provide feedback to teachers and identify their next learning steps;</p> <p>Inform teaching appraisal;</p> <p>Assess the extent to which goals for student learning have been met</p>	<p>Appraisers</p> <p>Appraisees</p>	<p>Student work</p> <p>School developed assessments</p> <p>Benchmarked assessments</p> <p>Aggregated data</p> <p>AsTTle data</p> <p>Exemplars</p> <p>etc</p>	<p>To inform: appraisal goals, appraisal discussions, appraisal observations, appraisal reports</p> <p>Focused: Feedback and guidance to teacher learners</p> <p>Changes to teaching practices</p> <p>Reliable, clear and valid information for appraisers and appraisees</p>

The explicit connection of ‘appraisal for learning’ to compatible current notions of formative assessment (Black & Wiliam, 1998; Clarke, 2001; E. Hargreaves, 2005; Timperley & Parr, 2004) is important since both assessment and appraisal ought to be *for* learning. Compatible assessment approaches are those in which teacher learning about student learning is a focus. For example, E. Hargreaves, (2005) suggests that conceptions of assessment should shift from ‘assessment-as-measurement’ to ‘assessment-as-inquiry’. Assessment-as-inquiry involves a search or investigation in order to develop teacher learning and understanding about student learning so that “the emphasis is not only on what or who is being assessed, but also on the assessor, the inquirer” (E. Hargreaves, 2005, p. 219). It emphasises, also, the importance of teachers learning about student learning to inform the next steps in teaching and learning.

*Aligning to effective professional development approaches*

Likewise, current notions of effective and sustainable professional development could be applied to teacher appraisal, since ‘appraisal for learning’ emphasises the professional development aspect of the process. In a study of school-based factors associated with high student achievement, Timperley and Wiseman reported that schools whose professional development resulted in the highest student achievement were those characterised by the features of learning communities. In those schools “teachers had shared norms that their children could and should be at national levels of achievement, the focus was constantly on student learning and teachers talked about and reflected on their professional practice. Practice was deprivatised in the sense that the achievement data was available for all to see, with the dialogue among teachers based on that data” (Timperley & Wiseman, 2002, p. 99). Professional learning communities are, then, influential in facilitating improvement in student learning, through teacher learning. Five elements that typify professional communities (Kruse, Louis, & Bryk, 1995; Louis & Marks, 1998) are:

- shared norms and values
- focus on student learning
- collaboration
- deprivatised practice
- reflective dialogue

Professional learning communities that encompass these characteristics will also be essential for ensuring that appraisal practices impact positively on student learning.

*Closing the Gap through Building Capacity for ‘Appraisal for Learning’*

Mitchell and Sackney (2001) signal the need to deliberately and explicitly build the capacity of educators to function as a learning community. A learning community will not occur simply because it has been required. The same can be said for ‘appraisal for learning’. Mitchell and Sackney (2001) describe three critical capacities in relation to building a learning community that are also critical for the implementation of ‘appraisal for learning’. These capacities – personal capacity, interpersonal capacity and organisational capacity – are explored next.

Personal capacity involves deep, critical deconstruction and reconstruction of professional knowledge (Mitchell & Sackney, 2001). In the appraisal context, personal capacity will relate to both the professional knowledge appraisers and appraisees have *about* appraisal, and the knowledge they bring *to* appraisal. Building educators' personal capacity to use data in appropriate and meaningful ways is central to 'appraisal for learning'. Practitioners will need a well-developed understanding of the type of student outcomes to use, and of how to measure student learning. 'Appraisal for learning' necessitates not only the use of data to examine the consequences that teaching has for student learning, but also the examination of the quality of the data and decision-making about appropriate shifts in teaching practice. Both pre-service and in-service teacher education (and professional learning opportunities) are important sites for building personal capacity for 'appraisal for learning'. Most importantly, potential teachers need to learn "how to learn in and from practice" (Ball & Cohen, 1999, p. 10).

Interpersonal capacity, according to Mitchell and Sackney (2001), refers to collegial relations and collective practices in which norms of professional learning are of ultimate importance to the group. In appraisal, this means giving explicit attention to ensuring appraisal interactions and practices (for example, discussions, observations and written reports) serve to promote professional learning. Piggot-Irvine's work (2001a) explains the significance of open, trust-based relationships to these interactions. It provides evidence of the effectiveness of establishing educative relationships between appraisers and appraisees. These relationships are characterised by "bilateralism (shared control, shared thinking, shared evidence, shared planning and monitoring) leading to appraisers having more confidence to help appraisees to confront and resolve problems" (Piggot-Irvine, 2001a, p. 297). Building the capacity for appraisers and appraisees to interact in ways that reflect productive, rather than defensive, reasoning ((Cardno, 1994, 1998) explored in Chapter 2) is important since productive reasoning promotes educative processes (Piggot-Irvine, 2003a).

Interpersonal capacity also relates to the norms that schools establish around the nature of interactions practitioners engage in. Analysis and critical examination of data during appraisal and the investigation of connections between teacher learning and student learning should be considered norms of appraisal interaction. Those norms require building interpersonal capacity for jointly examining data.

The third critical capacity is organisational capacity – the “structures and systems that support and value personal learning and that facilitate and encourage collective learning” (Mitchell & Sackney, 2001, p. 9). Central to the capacity for an organisation to implement ‘appraisal for learning’ is the nature of the context or community in which appraisal takes place. As outlined earlier (p. 183), a professional learning community is particularly conducive to this approach, and the development of such a community will strengthen the organisational capacity to implement ‘appraisal for learning’.

At a more tangible level, organisational capacity for ‘appraisal for learning’ could be promoted through, for example, the policy framework, documentation and templates that guide appraisal practices. These artefacts, which Halverson (2003) refers to as entities that shape and enable organisational practices, are an important part of the appraisal system of practice. A local system of practice is a network of artefacts that, together, shape the context of instruction and also indicate opportunities for school leaders to shift instructional practices (Halverson, 2003). Appraisal system artefacts (policies, indicators, processes for observations, discussions and reports) need to align with the elements of ‘appraisal for learning’ if they are to shape and enable appraisal that is focused on student learning and based on data. For instance, the headings on observation templates and written reports should explicitly reflect those elements by incorporating sections that relate to student learning, and to associated data. Guidelines on appraisal discussions should outline the type of data that will be examined during the discussion. Appraisal goal templates should prompt teachers to develop goals that are specific and challenging in terms of learning for students that will result as a consequence of a teacher achieving an agreed goal.

Developing systems and structures that support 'appraisal for learning' may require schools to think differently about the nature of the appraisal process. The findings of the study about appraisal discussions (reported in Chapter 4) drew attention to the impact that scheduling has on the function of those discussions. Discussions held later in the year were perceived to be summative in nature rather than an opportunity for teacher learning to take place. Alternative schedules may need to be considered that allow the formative function of appraisal to flourish. Consideration could also be given to implementing mini-appraisal cycles within the annual appraisal cycle, so that teachers' opportunities to develop goals, self-assess, be observed, and discuss progress occur many times throughout the year.

*Implications for leadership*

School leadership, clearly, has a fundamental role to play in promoting the capacities (particularly organisational capacity) outlined above, as Davis et al. explain:

...school-based administrative and professional leadership play essential roles in determining the meaning and value of teacher evaluation in schools, and how teacher evaluation can extend beyond its ritualistic traditions to improve teaching and learning. In the extreme case, leadership *makes the difference* between perfunctory summative teacher evaluation and meaningful assessment of the teaching and learning process that has the potential to enhance the quality of teaching and student learning. (Davis et al., 2002, p. 288, emphasis added)

The way in which educational leadership 'makes a difference' is largely through indirect, rather than direct, effects on student achievement. Educational leadership is extremely significant since it "is related to school organization and culture as well as to teacher behaviour and classroom practices and these factors are related in turn to student achievement" (Witziers, Bosker, & Krüger, 2003, p. 418). Teddlie (2005) describes this as a "causal pathway between educational leadership, teacher behaviours and values, and student achievement" (p. 219).

Leadership is a key mechanism in determining whether appraisal serves learning or compliance purposes and, as such, is an indirect mechanism for affecting student achievement. There needs, then, to be mechanisms in place within the wider education system to ensure that school leaders, too, have the capacity to develop those personal, interpersonal and organisational capacities. Principals' capacity to build others' capacity is essential if effectual, rather than perfunctory, appraisal (such as 'appraisal for learning') is to transpire. This is evident in the Government's "Education Priorities" document which signals the need to focus on outcomes by "developing leadership capacity for outcomes-focused management across the education system" and "increasing the systematic use of information to continuously improve institutional performance" (Ministry of Education, 2003b, p. 5).

As educators develop the personal, interpersonal and organisational capacities that in turn enable 'appraisal for learning', teachers' ability to learn "in and from practice", as Ball and Cohen put it, will be strengthened (1999). Attention to "particulars" is especially important for learning in and from practice – "teaching occurs in particulars – particular students interacting with particular teachers over particular ideas in particular circumstances" (Ball & Cohen, 1999, p. 10). 'Appraisal for learning' will provide learning opportunities for teachers that allow them to grapple with the particulars and complexities of their own circumstances – particular teachers and particular students whose particular learning outcomes are reflected in particular data which has particular implications for improving teachers' practice and student learning.

This research has sought to contribute to an understanding of the nature of, and potential for, appraisal in schools. 'Appraisal for learning' is theoretically and pragmatically justified, aligned to the direction of other education initiatives, and eminently possible. I hope that teachers, school leaders and policy makers alike will at least consider, hopefully have dialogue about, and ideally implement and examine, an 'appraisal for learning' approach to teacher appraisal.

## APPENDICES

### Appendix A Appraiser Interview Schedule

Focus Question: What is the focus of inquiry during teacher appraisal?

Reminders: recording, focus on last typical discussion, look at/share question schedules, look at the file to refresh memory, context for all questions is the particular appraisal discussion

Scene Setting: I'd like you to think back to the last appraisal discussion you had/did? When was it? (month/time of day). How long did it last? Did you have any particular documents with you during that interview? And where was the meeting held? Did you take any notes?

#### Phase One: Open Questions

What topics did you discuss during the appraisal discussion you did? (create a list)

I'd like you to rank those topics in order of most to least attention they were given during the discussion.

Which of those topics was given the most time? Why? Also importance/priority/most productive? Which of those topics was given the least time? Why? Also low importance/priority/most productive?

Let's focus on the topic [x] (any that have potential for inquiry into student learning)

What did you say or ask about [x]? What did you want to find out about [x]?

(Return to Q4 for any other topics with potential for inquiry into student learning)

Are there any other topics/questions you would have liked to raise or considered raising but did not? Why?

Were these topics and questions typical of those you have recently used in other appraisals? Why/why not?

Did you refer to any data or documents during the appraisal discussion? What were they?

Did you refer to your school's professional standards' indicators during the appraisal discussion?

What conclusions did you make about the effectiveness of this teacher?

What performance objectives were established for this appraisal?

What messages were conveyed to the teacher about their practice?

(School specific questions)

Phase 2 – probing – student learning, good teaching, programmes, how did you know, progress, achievement, success, effectiveness

Focus on student learning absent:

Did you find out anything about the teacher's effectiveness?

Did you find out anything about students' learning? If yes – what? How? If no - How do you think the teacher's students are doing?

Focus on student learning present: (probing would occur as student learning was raised, during phase 1)

Why did you focus on....?

Why did you ask about...(e.g. grouping)?

What did you find out the learning of that teacher's students?

Phase 3 – Direct (according to whether any focus on student learning was present or absent)

A) Focus on student learning absent:

I notice from what you've told me that student learning was not a particular focus of the appraisal discussion.

Would you agree with that?

Why do you think that is?

Are there other times when student learning is focused on?

What do you think about the idea of focusing on student learning/achievement during an appraisal discussion?

What do you think about the idea of that inquiry being based on data about your student's learning?

B) Focus on student learning present:

I notice from what you've told me that you focused on student learning when you talked/asked about \_\_\_\_\_.

Was there a particular reason that you did that?

What do you think about focusing on student learning/achievement during an appraisal discussion?

What do you think about that inquiry being based on data about your student's learning?



## Appendix B Appraisee Interview Schedule

Focus Question: What is the focus of inquiry during teacher appraisal?

Reminders: recording, focus on last typical discussion, look at/share question schedules, look at the file to refresh memory, context for all questions is the particular appraisal discussion

Scene Setting: I'd like you to think back to the last appraisal discussion you had/did? When was it? (month/time of day). How long did it last? Did you have any particular documents with you during that discussion? And where was the meeting held? Did you take any notes?

### Phase One: Open Questions

What topics were discussed at your last appraisal discussion? (create a list)

I'd like you to rank those topics in order of most to least attention they were given during the discussion.

Which of those topics was given the most time? Why? Also importance/priority/most productive? Which of those topics was given the least time? Why? Also low importance/priority/most productive?

Let's focus on the topic [x] (any that have potential for inquiry into student learning)

What were you asked about [x]? What was discussed about [x]

(Return to Q4 for any other topics with potential for inquiry into student learning)

Are there any other topics/questions you would like to have been raised that were not? Why?

Were these topics and questions typical of those you have experienced in other appraisals? Why/why not?

Were any data or documents referred to during the appraisal discussion? What were they?

Were your schools' professional standards' indicators referred to during the appraisal discussion?

What conclusions did you think were made about your effectiveness as a teacher?

What performance objectives were established for this appraisal?

What messages were conveyed to you about your practice?

(School specific questions)

Phase 2 – probing – student learning, good teaching, programmes, how did you know, progress, achievement, success, effectiveness

Focus on student learning absent:

Did you discuss your students' learning at all?

What kinds of things were asked about your students?

Focus on student learning present: (probing would occur as student learning was raised, during phase 1)

What exactly was asked about/discussed?

Phase 3 – Direct (according to whether any focus on student learning was present or absent)

A) Focus on student learning absent:

You did not particularly discuss your students' learning.

What do you think about the idea of focusing on student learning/achievement during an appraisal interview?

What do you think about the idea of that inquiry being based on data about your student's learning?

Do you think this might be useful?

B) Focus on student learning present:

You discussed your students' learning.

What do you think about focusing on student learning/achievement during an appraisal interview?

What do you think about that inquiry being based on data about your student's learning?

Appendix C Appraisal goals in the context of appraisal discussions

Participant Reference	Appraisal Role	Appraisal Goals focused on as Appraisal Discussion Topics
P1	Appraiser	Classroom teaching Numeracy project KPA Peer mediation – linkages with the junior school Leading of curriculum project team, responsibility for peer mediation (general -) development objectives set by teachers themselves
P2	Appraiser	Depth of questioning in maths Engaging students in maths, quality maths thinking and maths learning
P3	Appraisee	Questioning Including and engaging all students
P4	Appraisee	Numeracy project: grouping according to testing including ANP ideas in planning extending children, recognising and catering for needs (according to test data)
P5	Appraisee	Supporting other staff members in implementing visual arts Supporting the staff with innovation, resources and time (arts)
P6	Appraiser	Dance curriculum – the lesson, lesson content, involvement of students, outcome of the lesson for students
P7	Appraiser	Classroom teaching Management as a senior teacher Other school-wide
P10	Appraisee	Maths – working on ENP (also a school wide focus) ICT – familiarisation with junior level programmes (also a school wide focus) Special needs – the challenges of particular children and accessing appropriate support
P11	Appraisee	Social Studies – ensuring all three SS processes were used in the programme ICT – to try and find more resources of how to use ICT as a thinking tool rather than just a publishing sort of medium Reading – to re-evaluate how we ran our reading programmes “as to whether what we were doing ... how we were doing things was the most effective way of doing it and whether the children were progressing Leadership – “to create a lead team which utilised strengths to become self-managing...basically highlighting different people’s strengths and allowing them to use their strengths rather than being a boss, a boss leader”

Appendix D Teachers' appraisal goals anonymous questionnaire

1. Which year level band do most of your students fall into? (please circle)

Year 1-2      Year 3-4      Year 5-6      Year 7-8

2. How many years' teaching experience do you have? (please circle the number of full time equivalent years)

Less than 4 years      4 – 6 years      7 – 9 years      10+ years

3. What are your appraisal goals for this year? (please list these below – you may have fewer goals than the spaces provide. If you have more goals than the spaces provide, please record them on the back of this form)

4. What focus do you expect the goal to be given in your appraisal interview? **The goal will be...** (please circle)

	the <b>only</b> focus	the <b>main</b> focus	a <b>minor</b> focus	Don't know
	the <b>only</b> focus	the <b>main</b> focus	a <b>minor</b> focus	Don't know
	the <b>only</b> focus	the <b>main</b> focus	a <b>minor</b> focus	Don't know
	the <b>only</b> focus	the <b>main</b> focus	a <b>minor</b> focus	Don't know
	the <b>only</b> focus	the <b>main</b> focus	a <b>minor</b> focus	Don't know
	the <b>only</b> focus	the <b>main</b> focus	a <b>minor</b> focus	Don't know

5. Who has been involved in deciding on these goals? (please circle)

Myself (teacher)      Senior Teacher      Assistant/Deputy Principal      Principal      Other (please indicate) \_\_\_\_\_

Completion and return of this anonymous questionnaire indicates your consent to participate in this research. Please return the questionnaire in the envelope provided to Claire Sinnema, PO Box 26519, Epsom by October 3<sup>rd</sup>, 2003. Thank you for your time!

Appendix E Analysis of Specificity and Challenge of Student Outcome Focused Teacher Appraisal Goals

	Specificity Criteria						Specificity			Challenge		
	Learners		learning		Extent of improvement		Vague	General	Specific	Challenge explicit	Potentially challenging	Not challenging
	Specific	General	Specific	General	Specific	General						
develop independent learners [1-6]		✓		✓		✓	✓				✓	
developing literacy in Y1 [2-2]	✓			✓		✓	✓				✓	
to upskill all the children in my class in the use of ICT [2-5]	✓			✓		✓	✓				✓	
improve numeracy skills and teaching and numeracy thru NUMP [2-8]		✓		✓		✓	✓				✓	
to upskill literacy at yr1/yr2 level (reading/writing) [4-2]	✓			✓		✓	✓				✓	
improve written language throughout whole school (school wide goal) [4-4]	✓			✓		✓	✓				✓	
Students answering own questions [5-6]		✓		✓		✓	✓				✓	
Getting lower level students talking math [5-6]	✓			✓		✓	✓				✓	
develop cognitive and cooperative skills in Philosophy for Children creating a circle of inquiry [6-1]		✓	✓			✓	✓		✓			
extending gifted and talented children in my classroom [6-5]	✓			✓		✓	✓				✓	
develop and implement teaching practices which develop information literacy [8-1]		✓		✓		✓	✓				✓	

## REFERENCES

- Alton-Lee, A. (2004, April 19-20). *Improving educational policy and practice through an iterative best evidence synthesis programme*. Paper presented at the Evidence-based Policy Research. OECD-US Seminar, Washington D.C.
- Anderson, G. J. (1990). *Fundamentals of educational research*. London ; New York: Falmer.
- Argyris, C. (1986). Skilled incompetence. *Harvard Business Review*.
- Ministry of Education Argyris, C. (1999). *On organizational learning* (2nd ed.). Mass: Malden.
- Australian Council for Educational Research. (2004). Evaluating Australian teachers. In *Research Highlights*.
- Bailey, H. (1997). Doing appraisal. In O'Neill, J. (Ed.), *Teacher appraisal in New Zealand* (pp. 17-46). Massey: ERDC Press.
- Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In Sykes, G. & Darling-Hammond, L. (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco: Jossey Bass.
- Battersby, E. J. (1991). Staff appraisal: Preparation and climate setting. *New Zealand Journal of Educational Administration*, 6, 15-19.
- Bennett, H. (1999). One drop of blood: Teacher appraisal Mark 2. *Teacher Development*, 3(3), 411-428.
- Berliner, D. C. (1987). Simple views of effective teaching and a simple theory of classroom instruction. In Gage, N. L., Berliner, D. C. & Rosenshine, B. (Eds.), *Talks to teachers: a festschrift for N.L. Gage* (1st ed., pp. xiv, 411). New York: Random House.
- Black, P., & Wiliam, D. (1998). *Inside the black box*. London: Kings' College.
- Bower, M. (1996). *Will to manage*. New York: McGraw-Hill.

- Boyd, V. (1992). *School context: Bridge or barrier to change [Electronic version]* (Vol. 2002). Texas: Southwest Educational Development Laboratory.
- Brereton, P. (2001). Target setting: Science or fiction. In West-Burnham, J., O'Neill, J. & Bradbury, I. (Eds.), *Performance management in schools: how to lead and manage staff for school improvement* (pp. xiv, 220). Harlow: Pearson Education.
- Brock, P. (1998, 1999). *Ethics and professional teaching standards*. Paper presented at the Ethics of the Teaching Profession Standards Conference, Lewisham, New South Wales.
- Calabrese, R. L., Sherwood, K., Fast, J., & Womack, C. (2004). Teachers' and principals' perceptions of the summative evaluation conference: An examination of Model I theories-in-use. *International Journal of Educational Management*, 18(2).
- Cardno, C. (1994). *Dealing with dilemmas: a critical and collaborative approach to staff appraisal in two schools*. Unpublished PhD thesis, Auckland University, Auckland.
- Cardno, C. (1998). Making a difference by managing dilemmas. *SET: Research Information for Teachers*, 1(13).
- Cardno, C. (1999). Appraisal policy and implementation issues for New Zealand schools. *The International Journal of Educational Management*, 13(2), 87-97.
- Cardno, C. (2001). Managing dilemmas in appraising performance: An approach for school leaders. In Middlewood, D. & Cardno, C. E. M. (Eds.), *Managing teacher appraisal and performance : a comparative approach* (pp. xvi, 200). London ; New York: Routledge Falmer.
- Cardno, C., & Piggot-Irvine, E. (1996). Incorporating action research in school senior management training. *International Journal of Educational Management*, 10(5), 19-24.
- Cardno, C., & Piggot-Irvine, E. (1997). *Effective performance appraisal: Integrating accountability and development in staff appraisal*. Auckland: Longman.

- Carnegie Forum on Education and the Economy. (1986). *A nation prepared: Teachers for the 21st century*. New York.
- Chamberlain, M. (2000, July/October). *The national assessment strategy and the relationships between teaching, learning, and student achievement*. Paper presented at the National Assessment Regional Seminars, Christchurch, Palmerston North & Auckland.
- Clark, J. (2001). Ethical issues in managing teacher performance. In West-Burnham, J., O'Neill, J. & Bradbury, I. (Eds.), *Performance management in schools: how to lead and manage staff for school improvement* (pp. xiv, 220). Harlow: Pearson Education.
- Clarke, S. (2001). *Unlocking formative assessment*. London: Hodder & Stoughton.
- Cochran-Smith, M., & Lytle, S. L. (1999). The teacher research movement: A decade later. *Educational Researcher*, 28(7), 15-25.
- Coe, R. (n.d.). *Target setting: The evidence from research*. Retrieved September, 2003, from <http://www.cemcentre.org/ebeuk/experiments/targetsetting.pdf>
- Creemers, B. P. M. (1994). *The Effective Classroom*. London: Cassell.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, California: Sage Publications.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W., E. (Eds.). (2003). *Advanced mixed methods research designs*. California: Sage Publications.
- Cwikla, J. (2003). The importance of setting learning goals to investigate the effectiveness of teacher professional development. *Educational Research Quarterly*, 27(2), 43-59.
- Darling-Hammond, L. (1996). The right to learn and the advancement of teaching: Research, policy, and practice for democratic education. *Educational Researcher*, 25(6), 5-17.



- Darling-Hammond, L. (1998). Standards for assessing teaching effectiveness are key: a response to Schalock, Schalock and Myton [Electronic version]. *Phi Delta Kappan*, 79(6), 471-472.
- Davis, D. R., Ellett, C. D., & Annunziata, J. (2002). Teacher evaluation, leadership and learning organizations. *Journal of Personnel Evaluation in Education*, 16(4), 287-301.
- Deal, T. E., & Peterson, K. D. (1999). *Shaping school culture: The heart of leadership*. California: Jossey-Bass.
- Department for Education and Employment. (1998). *Excellence in schools, Command Paper 3681*. London: DfEE.
- Dewey, J. (1933). *How we think*. New York: D.C. Heath.
- Dimmock, C. A. J. (2000). *Designing the learning-centred school: A cross-cultural perspective*. London: Falmer.
- Down, B., Chadbourne, R., & Hogan, C. (2000). How are teachers managing performance management? *Asia-Pacific Journal of Teacher Education*, 28(3), 213-223.
- Down, B., Hogan, C., & Chadbourne, R. (1997, December). *Making sense of performance management in schools: Official rhetoric and teachers' reality*. Paper presented at the Australian Association for Research in Education Conference, Brisbane.
- Drever, E. (1995). *Using semi-structured interviews in small-scale research: a teacher's guide*. Glasgow: Scottish Council for Research in Education.
- Earl, L., & Fullan, M. (2003). Using data in leadership for learning. *Cambridge Journal of Education*, 33(3).
- Earl, L., & Katz, S. (2002). Leading schools in a data rich world. In Leithwood, K., Hallinger, P., Louis, K. S., Furman-Brown, G., Gronn, P., Mulford, B. & Riley, K. (Eds.), *Second International Handbook of Educational Leadership and Administration*. Dordrecht: Kluwer.

- The Education Act, New Zealand Government Sections 60A and 61 (1989).
- Education Standards Act, New Zealand Government sections 61-63 (2001).
- Educational Assessment Research Unit. (2003). *National education monitoring project forum comment*. Dunedin: University of Otago.
- Ellett, C. D. (1997). Classroom-based assessments of teaching and learning. In Stronge, J. (Ed.), *Evaluating teaching: A guide to current thinking and best practice* (pp. 107-128). Newbury Park, CA: Corwin Press.
- Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, 17(1), 101-128.
- Ellis, E. S., & Worthington, L. A. (1994). *Research synthesis on effective teaching principles and the design of quality tools for educators* (No. 5). Oregon: National Center to Improve the Tools of Educators.
- Evans, R. (1996). *The human side of change*. San Francisco: Jossey-Bass.
- Fancy, H. (2004). *Education reform: Reflections on New Zealand experience*. Retrieved March, 2005, from <http://www.minedu.govt.nz/index.cfm?layout=document&documentid=9750&indexid=5451&indexparentid=5437>
- Farrell, C., & Morris, J. (2004). Resigned compliance. *Educational Management & Administration*, 32(1), 81-104.
- Fitzgerald, T. (2001). Potential paradoxes in performance appraisal. In Middlewood, D. & Cardno, C. (Eds.), *Managing teacher appraisal and performance: a comparative approach*. London ; New York: Routledge Falmer.
- Fitzgerald, T., Youngs, H., & Grootenboer, P. (2003). Bureaucratic control or professional autonomy?: Performance management in New Zealand schools. *School Leadership & Management*, 23(1), 91 - 105.
- Fullan, M. (1988). Performance appraisal and curriculum implementation research. In Hickcox, E. S., Lawton, S., Leithwood, K. & Musella, D. (Eds.), *Making a*

*difference through performance appraisal* (pp. viii, 202). Toronto: Ontario Institute for Studies in Education Press.

Fullan, M. (1992). *Successful school improvement: the implementation perspective and beyond*. Buckingham [England] ; Philadelphia: Open University Press.

Fullan, M. (1999). *Change forces: the sequel*. London: Falmer Press.

Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.

Fullan, M. (2002). *Change forces with a vengeance*. London: Routledge Falmer.

Fullan, M., & Mascal, B. (2000). *Human resource issues in education: A literature review. Prepared for the New Zealand Ministry of Education: Ontario Institute for Studies in Education of the University of Toronto*.

Furtwengler, C. B. (1995). State actions for personnel evaluation: Analysis of reform policies, 1983-1992. *Education Policy Analysis Archives*, 3(4).

Ghaye, T., & Ghaye, K. (1998). *Teaching and learning through critical reflective practice*. London: D. Fulton Publishers.

Glass, G. V. (1999). Solving the policy implementation problem: The case of Arizona charter schools. *Education Policy Analysis Archives [Electronic version]*, 7(26).

Goodlad, J., Soder, R., & Sirotnik, K. (1990). *The moral dimensions of teaching*. San Francisco: Jossey-Bass.

Gordon, L. (2003). School choice and the social market in New Zealand: Education reform in an era of increasing inequality. *International Studies in Sociology of Education*, 13(1), 17-34.

Gratton, R. (2004). Teacher appraisal: a lesson on confusion over purpose. *International Journal of Educational Management*, 18(5), 292-296.

- Grootenboer. (2000). Appraisal for quality learning. *Waikato Journal of Education*, 6, 121-131.
- Gunter, H. (1996). Appraisal and the school as a learning organisation. *School Organisation*, 16(1), 89-100.
- Gunter, H. (2002). Teacher appraisal 1988-1998: A case study. *School Leadership & Management*, 22(1), 61-72.
- Haigh, N. (2000). Teaching teachers about reflection and ways of reflecting. *Waikato Journal of Education*, 6.
- Hall, A. (2003). *Notes on a code of ethics for New Zealand registered teachers*. Paper presented at the New Zealand Teachers Council Education Summit, Victoria University, Wellington.
- Halverson, R. (2003). Systems of practice: How leaders use artifacts to create professional community in schools. *Education Policy Analysis Archives*, 11(37).
- Hannay, L. M., & Telford, C. (2003). Making the conceptual shift: Teacher performance appraisal as professional growth. *Educational Action Research*, 11(1), 121-137.
- Hargreaves, A., & Goodson, I. (1996). Teachers' professional lives: Aspirations and actualities. In Goodson, I. & Hargreaves, A. (Eds.), *Teachers' professional lives* (pp. 1-27). London: Falmer Press.
- Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education*, 35(2), 213-224.
- Hattie, J. (1999a). *Aotearoa Board for Professional Teaching Standards: A Proposal*. Retrieved June 21, 2003, from <http://www.arts.auckland.ac.nz/edu/staff/jhattie/NZBPTS.html>
- Hattie, J. (1999b). *Influences on Student Learning: Inaugural Professorial Address*. Paper presented at the Inaugural Lecture: Professor of Education, University of Auckland.

- Hattie, J. (2002). *What are the attributes of excellent teachers?* Paper presented at the Teachers make a difference: What is the research evidence? New Zealand Council for Educational Research Conference 2002, Wellington.
- Haynes, G., Wragg, T., Wragg, C., & Chamberlin, R. (2003). Performance management for teachers: Headteachers' perspectives. *School Leadership & Management*, 23(1), 75-89.
- Higgins, J. (2001). *Developing numeracy: Understanding place value*. Wellington: Ministry of Education.
- Hodkinson, H., & Hodkinson, P. (2005). Improving schoolteachers' workplace learning. *Research Papers in Education*, 20(2), 109-131.
- Hogan, C., Down, B., & Chadbourne, R. (1998, December). *How are teachers managing performance management?* Paper presented at the Australian Association for Research in Education, Adelaide.
- Hopkins, D. (2001). *School improvement for real*. London: Routledge Falmer.
- Hopkins, D., & Levin, B. (2000). Government policy and school development. *School Leadership & Management*, 20(1), 15-30.
- Hopkins, D., & MacGilchrist, B. (1998). Development planning for pupil achievement. *School Leadership & Management*, 18(3), 409-424.
- Hoyle, E. (1982). The professionalization of teachers: A paradox. *British Journal of Educational Studies*, 2, 161-171.
- Hoyle, E., & John, P. D. (1995). *Professional knowledge and professional practice*. London: Cassell.
- Irons, B. (1993). *Appraisal: worthwhile or ineffective? A case study of the implementation and development of performance appraisal in four New Zealand primary schools* (No. 5). Palmerston North: Massey University.
- James, M. (2005). Insights on teacher learning from the teaching and learning research programme. *Research Papers in Education*, 20(2), 105-108.

- Jennings, K., & Lomas, L. (2003). Implementing performance management for headteachers in English secondary schools: A case study. *Educational Management & Administration*, 31(4), 369-383.
- Johnson, B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In Tashakkori, A. & Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioral research*. California: Sage Publications.
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development*. Alexandria: Association for Supervision and Curriculum Development.
- Kleinhenz, E., & Ingvarson, L. (2004). Teacher accountability in Australia: Current policies and practices and their relation to the improvement of teaching and learning. *Research Papers in Education*, 19(1), 31-49.
- Kruse, S. D., Louis, K. S., & Bryk, A. S. (1995). An emerging framework for analyzing school-based professional community. In Louis, K. S. & Kruse, S. D. (Eds.), *Professionalism and community: Perspectives on reforming urban schools*. Newbury Park, California: Corwin.
- Lange, D. (1988). *Tomorrow's Schools: the reform of education administration in New Zealand*. Wellington: Government Printer.
- Leithwood, K. A., & Louis, K. S. (1998). *Organizational learning in schools*. Lisse [Netherlands] ; Exton, PA: Swets & Zeitlinger Publishers.
- Levin, B. (1998). An epidemic of education: (What) Can we learn from each other? *Comparative Education*, 34(2), 131-141.
- Levin, B., & Wiens, J. (2003). There is another way: A different approach to education reform. *Phi Delta Kappan [Electronic version]*, 84(9).
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. New Jersey: Prentice Hall.

- Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, 106.
- Lovat, T. (1997, July). *Ethics and ethics education: Professional and curricular best practice*. Paper presented at the Paper Presented at the Biennial Conference of the Australisan Curriculum Studies Association, University of Sydney.
- Middlewood, D., & Cardno, C. (2001). *Managing teacher appraisal and performance : A comparative approach*. London ; New York: Routledge Falmer.
- Miles, M. B., & Huberman, M. A. (1994). *Qualitative data analysis: An expanded sourcebook*. California: Sage.
- Miller, S. (2003). Impact of mixed methods and design on inference quality. In Tashakkori, A. & Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioral research*. California: Sage Publications.
- Mills, H. (2001). When teachers have time to talk: The value of curricular conversations. *Language Arts*, 79(1), 20-28.
- Ministry of Education. (1992). *Mathematics in the New Zealand curriculum*. Wellington: Learning Media.
- Ministry of Education. (1993). *The New Zealand curriculum framework*. Wellington: Learning Media.
- Ministry of Education. (1994). *English in the New Zealand curriculum*. Wellington: Learning Media.
- Ministry of Education. (1995). *Technology in the New Zealand curriculum*. Wellington: Learning Media.
- Ministry of Education. (1997a, 10 February supplement). Performance management systems: PMS1: Performance appraisal. *Education Gazette*.
- Ministry of Education. (1997b). *Science in the New Zealand curriculum*. Wellington: Learning Media.

Ministry of Education. (1997c). *Social studies in the New Zealand curriculum*. Wellington: Learning Media.

Ministry of Education. (1998a). *Interim professional standards: Primary school deputy/assistant principals, primary school teachers*. Wellington: Ministry of Education.

Ministry of Education. (1998b, October). Introduction of interim professional standards: Primary school deputy/assistant principals, primary school teachers. *New Zealand Education Gazette*, 5.

Ministry of Education. (1998c). *Teacher performance management: Primary school teachers, primary school deputy/assistant principals. A resource for boards of trustees, principals and teachers*. Wellington: Government Printer.

Ministry of Education. (1999). *Health and physical education in the New Zealand curriculum*. Wellington: Learning Media.

Ministry of Education. (2000). *The arts in the New Zealand curriculum*. Wellington: Learning Media.

Ministry of Education. (2001a). Curriculum Update He Korero Marautanga: Assessment - Te Aro Matawai. *Issue 47*.

Ministry of Education. (2001b). *PISA 2000: Overview of selected New Zealand findings*. Wellington: Ministry of Education.

Ministry of Education. (2001c). *Primary Teachers' (including Deputy and Assistant Principals and other unit holders) Collective Agreement 2001-2003*. Retrieved 15 April 2002, 2001, from [http://www.minedu.govt.nz/web/document/document\\_page.cfm?id=6156&p=1003.1010.1013.1014#P1577\\_21365](http://www.minedu.govt.nz/web/document/document_page.cfm?id=6156&p=1003.1010.1013.1014#P1577_21365)

Ministry of Education. (2002a, May 2004). *Assessment tools for teaching and learning*. Retrieved June, 2004, from [http://www.tki.org.nz/r/asttle/index\\_e.php](http://www.tki.org.nz/r/asttle/index_e.php)

Ministry of Education. (2002b). *Planning for better student outcomes*. Wellington.



- Ministry of Education. (2003a). *Annual Report 2003*. Wellington.
- Ministry of Education. (2003b). *Education priorities for New Zealand*. Retrieved May, 2003, from <http://www.minedu.govt.nz/index.cfm?layout=document&documentid=8354&indexid=1108&indexparentid=2107>
- Ministry of Education. (2003c). *The national administration guidelines*. Retrieved January, 2005, from <http://www.minedu.govt.nz/index.cfm?layout=document&documentid=8187&data=l>
- Ministry of Education. (2003d). *Planning for better student outcomes*. Wellington.
- Ministry of Education. (2003e). *Planning for better student outcomes: Analysing student achievement data*. Wellington.
- Ministry of Education. (2003f). *Planning for better student outcomes: Improvement, plans, and charters*. Wellington.
- Ministry of Education. (2003g). *Planning for better student outcomes: Reporting on student achievement - refocusing annual reports*. Wellington.
- Ministry of Education. (2003h). *A Study of the Use of School Entry Assessment (SEA) in Schools*. Wellington: Research Division, Ministry of Education.
- Ministry of Education. (2004a). *Impact of teaching and schools on variance in outcomes*. Wellington: Ministry of Education.
- Ministry of Education. (2004b). *Ministry of education annual report 2004*. Wellington.
- Ministry of Education. (2005). *Curriculum/Marautanga project online*, from <http://www.cmp.ac.nz/>
- Mitchell, C., & Sackney, L. (2001). Building capacity for a learning community. *Canadian Journal of Educational Administration and Policy*, 19.

- Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving schools in socioeconomically disadvantaged areas - A review of research evidence. *School effectiveness and school improvement*, 15(2), 149-175.
- Muijs, D., & Reynolds, D. (2001). *Effective teaching: evidence and practice*. London: Paul Chapman.
- National Commission for Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: US Government Printing Office.
- New South Wales Government. (1989a). *Report of the Committee of Review of NSW Schools (The Carrick Report)*. Sydney.
- New South Wales Government. (1989b). *School-centred education: Building a more responsive State School System (The Scott Review)*. Sydney.
- New South Wales Ministry of Education and Youth Affairs. (1989). *Excellence and equity: New South Wales Curriculum Reform. A White paper on curriculum reform in New South Wales schools*. Sydney.
- New South Wales Parliament. (1990). *Education Reform Act*. Retrieved October, 2004, from <http://www.legislation.nsw.gov.au/maintop/search/sessional>
- New Zealand Council for Educational Research. (2005). *Assessment resource banks*. Retrieved February, 2005, from <http://arb.nzcer.org.nz/nzcer3/nzcer.htm>
- Novlan, J. F. (1998). New Zealand's past and tomorrow's schools: Reasons, reforms and results. *School Leadership & Management*, 18(1), 7-18.
- O'Brien, P., & Down, B. (2002). What are teachers saying about new managerialism? *Journal of Educational Enquiry*, 3(1), 111-133.
- O'Neill, J. (1997). Teach, learn, appraise: An impossible triangle. In O'Neill, J. (Ed.), *Teacher Appraisal in New Zealand* (pp. 109-122). Palmerston North: ERDC Press.

- Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In Tashakkori, A. & Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioral research*. California: Sage Publications.
- Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). California: Sage Publications.
- Peterson, K. D., & Deal, T. E. (2002). *The shaping school culture fieldbook*. San Francisco: Jossey-Bass.
- Picot, B. (1988). *Administering for excellence: Effective administration in education*. Wellington: Government Printer.
- Piggot-Irvine, E. (2000). Appraisal: the impact of increased control on the "state of play" in New Zealand schools. *Journal of Educational Administration*, 38(4), 331-349.
- Piggot-Irvine, E. (2001a). *Appraisal - Reducing Control: Enhancing Effectiveness*. Unpublished PhD thesis, Massey University, Auckland.
- Piggot-Irvine, E. (2001b). *Tackling the hard stuff!* Retrieved January, 2004, from [http://www.nzpf.ac.nz/resources/magazine/2001/june/tackling\\_the\\_hard\\_stuff!.htm](http://www.nzpf.ac.nz/resources/magazine/2001/june/tackling_the_hard_stuff!.htm)
- Piggot-Irvine, E. (2003a). Appraisal training focused on what really matters. *The International Journal of Educational Management*, 17(6/7), 254-261.
- Piggot-Irvine, E. (2003b). Key features of appraisal effectiveness. *The International Journal of Educational Management*, 17(4), 170-178.
- Poskitt, J., Brown, M., Maw, N., & Taylor, K. (2003, December). *New Zealand national curriculum exemplars : Mist or must for teachers?* Paper presented at the New Zealand Association for Research in Education/Australian Association for Research in Education Conference, Auckland.
- Rallis, S. F., & MacMullen, M. M. (2000). Inquiry-minded schools - Opening doors for accountability [Electronic version]. *Phi Delta Kappan*, 81(10), 766.

- Reagan, T. G., Case, C. W., & Brubacher, J. W. (2000). *Becoming a reflective educator: How to build a culture of inquiry in the schools* (2nd ed.). Thousand Oaks California: Sage Publications.
- Reynolds, D., Muijs, D., & Treharne, D. (2003). Teacher evaluation and teacher effectiveness in the United Kingdom. *Journal of Personnel Evaluation in Education*, 17(1), 83-100.
- Riordan, G., & Weller, S. (2000, December 4-7). *The reformation of education in New South Wales: The 1990 Education Reform Act*. Paper presented at the Australian Association for Research in Education, Sydney.
- Robinson, V. (1993). *Problem-based methodology: Research for the improvement of practice*. Oxford: Pergamon Press Ltd.
- Robinson, V., Phillips, G., & Timperley, H. S. (2002). Using achievement data for school based curriculum review: A bridge too far? *Leadership and Policy in Schools*, 1(1), 3-29.
- Rowan, B., Correnti, R., & Miller, R. J. (2002). *What large scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools*. Pennsylvania: Consortium for Policy Research in Education, University of Pennsylvania.
- Sachs, J. (1997). Reclaiming the agenda of teacher professionalism: An Australian experience [Electronic version]. *Journal of Education for Teaching*, 23(3), 263-275.
- Schmoker, M. (1996). *Results: The key to continuous school improvement*. Alexandria, VA: ASCD.
- Schmoker, M. (2004). Tipping point: From feckless reform to substantive instructional improvement. *Phi Delta Kappan [Electronic version]*, 85(6).
- Schön, D. (1983). *The reflective practitioner. How professionals think in action*. London: Temple Smith.
- Schweitzer, M., Ordonez, L., & Douma, B. (2004). Goal setting as a motivator of unethical behavior. *Academy of Management Journal*, 47(3), 422-432.

- Semler, S. W. (1997, March 6-9). *Intervening to align: Approaches to improve organizational alignment*. Paper presented at the Academy of Human Resource Development, Atlanta, Georgia.
- Senge, P. M. (2000). *Schools that learn: A fifth discipline fieldbook for educators, parents, and everyone who cares about education* (1st ed.). New York: Doubleday.
- Shapiro, J. P., & Stefkovich, J. A. (2001). *Ethical leadership and decision making in education: Applying theoretical perspectives to complex dilemmas*. Mahwah, N.J.: L. Erlbaum Associates.
- Shulman, L. S. (1998). Theory, practice, and the education of professionals [Electronic version]. *Elementary School Journal*, 98(5), 511-526.
- Shulman, L. S. (2002). *Messing with the world of the ethics of teaching and scholarship*. Retrieved March, 2003, from [http://www.carnegiefoundation.org/elibrary/docs/ethics\\_of\\_inquiry\\_fore.pdf](http://www.carnegiefoundation.org/elibrary/docs/ethics_of_inquiry_fore.pdf)
- Sockett, H. (1993). *The moral base for teacher professionalism*. New York: Teachers College Press.
- Soder, R. (1990). The rhetoric of teacher professionalization. In Goodlad, J. I., Soder, R. & Sirotnik, K. (Eds.), *The moral dimensions of teaching* (pp. 35-86). San Francisco: Jossey-Bass.
- Soltis, J. A. (1986). Teaching professional ethics. *Journal of Teacher Education*, 37(3), 2-4.
- Southwest Educational Development Laboratory. (1999). Principals and teachers: Continuous learners. *Issues*, 7(2).
- Stake, R. E. (1989). The evaluation of teaching. In Simons, H. & Elliott, J. (Eds.), *Rethinking Appraisal and Assessment* (pp. 13-23). Philadelphia: Open University Press.
- The State Sector Act, New Zealand Government Part VII  
Section 77C (1988).
- Stewart, D. (1997). Appraisal and quality learning circles. In O'Neill, J. (Ed.), *Teacher Appraisal in New Zealand* (pp. 47-63). Palmerston North: ERDC Press.

- Stoll, L. (2000). School culture. *SET: Research Information for Teachers*, 3, 9-14.
- Stoll, L., Fink, D., & Earl, L. (2003). *It's about learning (and it's about time)*. London : New York: Routledge Falmer.
- Stronge, J. H., & Tucker, P. D. (1999). The politics of teacher evaluation: A case study of new system design and implementation. *Journal of Personnel Evaluation in Education*, 13(4), 339-359.
- Sykes, G. (1999). Teacher and student learning: Strengthening their connection. In Darling-Hammond, L. & Sykes, G. (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 151-180). San Francisco: Jossey-Bass.
- Tashakkori, A., & Teddlie, C. (2003a). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In Tashakkori, A. & Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioral research*. California: Sage Publications.
- Tashakkori, A., & Teddlie, C. (2003b). The past and future of mixed methods research: From data triangulation to mixed model designs. In Tashakkori, A. & Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioral research*. California: Sage Publications.
- Teddlie, C. (2005). Methodological issues related to causal studies of leadership: A mixed methods perspective from the USA. *Educational Management Administration and Leadership*, 33(2), 211-227.
- Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In Tashakkori, A. & Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioral research*. California: Sage Publications.
- Timperley, H. S. (2004). *Analysis and use of student achievement data: Final evaluation report prepared for the Ministry of Education*. Wellington: Ministry of Education.
- Timperley, H. S., & Parr, J. (2004). *Using evidence in teaching practice*. Auckland: Hodder Moa Beckett.

- Timperley, H. S., Phillips, G., & Wiseman, J. (2002a). *The sustainability of professional development in literacy. Part one: Changing and sustaining teachers' expectations through professional development in literacy.*
- Timperley, H. S., Phillips, G., & Wiseman, J. (2002b). *The sustainability of professional development in literacy. Part one: Changing and sustaining teachers' expectations through professional development in literacy:* Ministry of Education.
- Timperley, H. S., Phillips, G., & Wiseman, J. (2002c). *The sustainability of professional development in literacy. Part two: School-based factors associated with high student achievement.*
- Timperley, H. S., & Robinson, V. (2000). Workload and the professional culture of teachers. *Educational Management & Administration, 28*(1), 47-62.
- Timperley, H. S., & Wiseman, J. (2002). *The sustainability of professional development in literacy. Part two: School-based factors associated with high student achievement:* Ministry of Education.
- Townsend, F. (1995). Appraisal: A process of betrayal? *New Zealand Journal of Educational Administration, 10*, 12-19.
- Wang, M. C., Haertael, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*(3), 249-294.
- Warren-Little, J. (1999). Organizing schools for teacher learning. In Darling-Hammond, L. & Sykes, G. (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 233-262). San Francisco: Jossey-Bass.
- West-Burnham, J., & O'Neill, J. (2001). Perspectives on performance management. In West-Burnham, J., O'Neill, J. & Bradbury, I. (Eds.), *Performance management in schools: How to lead and manage staff for school improvement* (pp. xiv, 220). Harlow: Pearson Education.
- Witziers, B., Bosker, R. J., & Krüger, M. L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly, 39*(3), 398-425.

Yin, R. K. (1989). *Case study research: Design and methods* (Rev. ed.). Newbury Park, CA: Sage Publications.

Youngs, H., & Grootenboer, P. (2003). Primary and secondary teachers' perceptions of their appraisal: Implications for school leaders. *New Zealand Journal of Educational Administration and Leadership*.

Zeichner, K. M., & Liston, D. P. (1996). *Reflective teaching: An introduction*. Mahwah, N.J.: L. Erlbaum Associates.