Students' Perceptions of the Formative Potential of the National Certificate of Educational Achievement

A thesis presented in partial fulfilment of the requirements of the degree of PhD in Education.

Peter Leslie Charles Rawlins

2007

Abstract

Research evidence suggests that appropriate use of formative assessment promotes effective learning. Improved learning occurs when assessment is viewed as integral to learning, and when it is supported by coherent assessment systems. Although assessment systems designed primarily around the formative purpose can provide both formative and summative information, a tension exists in practice between the summative and formative purposes of assessment.

Using a theoretical framework developed by Sadler (1989), this research project investigated whether New Zealand's new secondary school Standards-Based Assessment qualification—the National Certificate Educational Achievement (NCEA)—has the potential to satisfy both summative and formative purposes of assessment in mathematics. Theorising from a contemporary sociocultural perspective of learning, this project recognised the situated nature and interpersonal dimension of knowledge, and the impact of the social environment in promoting and directing learning. Theorising from this perspective offered opportunities to examine classroom assessment practices from a new perspective.

To date, insufficient attention has been paid to the 'students' voice' concerning educational matters that directly affect them. Given the situated nature of students' engagement with formative practices a case study approach was used to investigate students' perceptions of the formative potential of NCEA mathematics assessment tasks. Three Y12 mathematics classes from an urban secondary school formed the case study singularity for this study. Focus group interviews with nine students were conducted across the year, complemented by classroom observations, a focus group interview with the teachers, and a quantitative questionnaire with all students in each of the three Year 12 mathematics classes.

An examination of the philosophical and structural design of NCEA revealed a strong potential for it to serve a duality of both formative and summative purpose of assessment. However the formative potential of NCEA was yet to be fully realised in the case study classrooms. Students' underdeveloped knowledge of assessment criteria effectively reduced the potential for students' independent use of self-

assessment strategies. This project also identified that teachers and students held differing views on preferred feedback practices. The teachers perceived that students did not read written feedback, and this perception significantly influenced the amount of written feedback that they offered to students. In contrast, students clearly displayed that they read, valued and used scaffolded written feedback to improve their learning. While the teachers preferred to offer oral feedback, students preferred to engage with their peers to use feedback to develop corrective strategies and deepen learning.

The project has made a number of practical and theoretical suggestions to improve students' understandings of the assessment criteria they are working towards, and to more effectively integrate the collaborative use of formative feedback into students' learning experiences. In particular, it has suggested two additional perspectives on the development and use of formative assessment in a sociocultural learning environment. Firstly, that students' knowledge of the role of formative assessment is socially and contextually situated, and develops through the social interactions that occur in the classroom. Secondly, the potential exists for formative assessment practices to stimulate collaborative learning opportunities within communities of practice.

Acknowledgements

I wish to express my appreciation for the guidance, feedback, and support provided by Associate Professor Glenda Anthony and Dr. Jenny Poskitt. Your honesty, support, and confidence in me, have been a great motivation to complete this research. I have greatly appreciated your expertise and expert tutelage. As Sir Isaac Newton said: "If I have seen further, it is by standing on the shoulders of giants".

I also wish to express my gratitude to the nine students and three teachers who took an active part in this research. Without your help this research could not have been undertaken. Your openness, honesty and candour are greatly appreciated. Time is valuable, and I wish to acknowledge the amount of time you willingly gave up to help me with this research. I hope that your expressed views help to improve classroom formative assessment practices in the future. I also wish to thank all of the students in the Year 12 Trad Maths classes for allowing me to be a part of your mathematics class and for helping me with this research.

Lastly I wish to thank Lynda for her tireless support and unswerving love. People have asked me if I will be relieved when my PhD is completed. I always answer that I have enjoyed the process and will be happy, rather than relieved. I am, however, sure that you will be relieved. Thank you for your love and understanding.

Contents

1 Introduction 1

1.1 Background..... 1
1.2 Purposes of assessment 5
1.3 The specific problem..... 8
1.4 The research objective..... 9
1.5 Summary 11
1.6 Structure of this thesis. 12

2 Literature review 13

2.1 Introduction 13
2.2 Assessment systems in New Zealand 13
2.3 Learning and assessment.....18
2.4 Impact of high-stakes assessment.....31
2.5 Assessment as a continuum. 40
2.6 A theoretical way forward.....41
2.7 Conclusion.....61

3 Methodology 63

3.1 Introduction 63
3.2 The nature of reality.....64
3.3 Case study methodology 65
3.4 Quality criteria: validity and reliability.....69
3.5 Selection of the singularity.....74
3.6 The students' voice 76
3.7 Data generating tools 78
3.8 Ethical considerations 84
3.9 Methodology in action 86
3.10 Data Analysis.....95
3.11 Summary. 98

4 Results 99

4.1 Introduction 99
4.2 The Y12 mathematics classroom.....100
4.3 Assessment Experience Questionnaire. 105
4.4 Understanding assessment criteria.....112
4.5 Provision of feedback.....120
4.6 Engagement with feedback 133
4.7 Conclusion.....150

5 Discussion and conclusions. 151

5.1 Introduction 151
5.2 Understanding assessment criteria 153
5.3 The Nature of Feedback 160
5.4 Student Engagement with Feedback 166
5.5 The formative potential of NCEA 173
5.6 Assessment in a sociocultural learning environment 179
5.7 Reflections on the research process 184
5.8 Final thoughts 188

References 190

Appendices

Appendix A: Students' information sheet 212
Appendix B: Students' consent form 214
Appendix C: Board of Trustees' information sheet 215
Appendix D: Teachers' information sheet 218
Appendix E: Board of Trustees' information sheet for AEQ 220
Appendix F: Students' information for AEQ 221
Appendix G: Focus group overview 222
Appendix H: Example of follow up questionnaire to students 223
Appendix I: Assessment Experience Questionnaire 225
Appendix J: Students' responses to AEQ 227
Appendix K: Research students' individual responses to AEQ 229
Appendix L: Students' results 230
Appendix M; 'I can do' sheet for Algebra Standard 235

List of figures

Figure 3.1 Annotape working screen 97
Figure 4.1 Display of mean student responses for items of AEQ 105
Figure 4.2 Individual students' responses for factor scores 109
Figure 4.3 95% confidence intervals for the mean factor scores 110
Figure 4.4 Example of part of an 'I can do' sheet. 117

List of tables

Table 4.1 Source of quotes 100
Table 4.2 Statistics for factor scores 109
Table 4.3 Y12 Cohort mean responses to individual items 111

1 INTRODUCTION

> Formative assessment requires that pupils have a central part in it; pupils have to be active in their own learning (teachers cannot learn for them) and unless they come to understand their strengths and weaknesses, and how they might deal with them, they will not make progress. (Harlen & James, 1997)

1.1 Background

According to Crooks (1988), assessment is one of the most potent forces affecting students' learning. Assessment affects their motivation and self-perception of competence. It consolidates learning and guides students' judgements about what is important to learn. Assessment affects the development of enduring learning strategies. These strategies allow students to adapt and respond to new experiences throughout their lifetimes, and may well be more important than the knowledge students accumulate during schooling.

Whilst in the past, assessment was viewed as the endpoint to learning it is now regarded as integral to learning (Black & Wiliam, 1998a; Crooks, 1988, 2006; Sadler, 1989, 1998). Contemporary sociocultural theories of learning increasingly recognise the situated nature and interpersonal dimension of knowledge, and the impact of the social environment in promoting and directing learning (Sfard, 2003). Sociocultural theories reconceptualise the roles and responsibilities of students and teachers in the learning process. Rather than being seen as passive recipients of knowledge, sociocultural theories conceptualise students as active participants in communities of practice (Lave & Wenger, 1991). Theorising from this conception of learning offers opportunities to examine assessment practices from a new perspective. As such, students' active formative use of assessment information within communities of practice is an important component of research into effective student learning.

Educationalists' increasing interest in the interaction between assessment and learning is coupled with the hope that improvements in classroom assessment practices will make a strong contribution to improvements in learning (Black, 2001).

> As reformers dream about changing education for the better they almost always see a need to include assessment and testing in their plans and frequently see them as the main instruments of their reforms. This is because assessment and testing are both ways of expressing aims and means to promote or impose them. (p. 65)

Approaches to improving overall achievement levels, particularly for low achieving students, varies between countries.

> Some countries are using high-stakes pressure on schools to achieve these ends, based on systematic testing throughout schooling (e.g., USA, England); others, like New Zealand, have chosen to focus on improving the capability of schools and teachers, and to offer assessment frameworks that give more choice to students and schools. (Hipkins, Wylie, & Hogden, 2007, p. 2)

1.1.1 The New Zealand context

Within the New Zealand context, the Ministry of Education (MoE) has published a number of policy documents that recognise and emphasise the importance of assessment in informing learning (e.g., Achievement 2001, 1998; Assessment: Policy to Practice, 1994; Mathematics in the New Zealand Curriculum, 1992; The New Zealand Curriculum Framework, 1993; and The New Zealand Curriculum: Draft for Consultation, 2006). According to the current MOE policy:

> The primary purpose of assessment is to improve students' learning and teachers' teaching as both student and teacher respond to the information that it provides. (MoE, 2006, p. 30)

The rhetoric in these policy and curriculum documents is supported by nationwide large-scale initiatives supporting assessment practices in schools (e.g., Assess to Learn (AtoL), Assessment Resource Banks (ARBs), Assessment Tools for Teaching and Learning (asTTle), the New Zealand Curriculum Exemplars).

In 2002, the New Zealand Qualifications Authority (NZQA) commenced the implementation of a new Standards-Based Assessment (SBA) system—the National Certificate of Educational Achievement (NCEA)—as the principal exit qualification from secondary school. As part of a world wide educational assessment reform movement, SBA has become a central feature of education systems around the world (Zepke, Leach, Brandon, Chapman, Neutze, Rawlins, & Scott, 2006). Research suggests that SBA qualification systems can have a positive effect on teaching and

learning. Proponents of SBA claim that the clear goals inherent in SBA can focus teacher and student attention to the key learning objectives (e.g., Barrington, 2004; Bushnell, 1992; Gipps, 1994; Kannapel, Aagard, Coe, & Reeves, 2001; OECD, 2005; Preece & Skinner, 1999; Wilson & Floden, 2001).

Integrated into the last three years of secondary school, NCEA is comprised of a combination of internally assessed Standards completed during the course of the year, and externally assessed Standards completed at the end of the year[1]. Replacing a series of Norm-Referenced exit qualifications, the implementation of NCEA commenced in 2002 and took three years to complete. The implementation was accompanied by an extensive professional development programme, providing teachers with an opportunity to reflect on the role of assessment in students' learning.

The recent introduction of NCEA has highlighted the inter-relationship between assessment and learning. However it will take time and support before the educational potential of this new assessment system can be realised. Research has an important role to play in examining the theoretical and practical implications for teachers' practices and students' learning.

1.1.2 Mathematics classrooms—situating the problem

Changing perspectives of how students learn, combined with the introduction of a new SBA secondary school exit qualification, have provided an opportunity to examine in-class assessment practices in mathematics from an alternative perspective. Mathematics holds a significant place in society's conceptions of what it means to be 'educated' and 'successful' (Cotton, 2001; Ernest, 2004). Several researchers (e.g., Boaler, 2002; Gates, 2001; Sfard, 2003) have argued that the narrowness with which school mathematics is organised and conceptualised has maintained a system of educational failure, in which only a few ever attain mathematical proficiency.

---

[1] There are two types of assessment Standards; Unit Standards and Achievement Standards. Unit Standards are awarded either a 'Not Achieve' or 'Achieve' grade. Achievement Standards are awarded either a 'Not Achieved', 'Achieved', 'Achieved with Merit' or 'Achieved with Excellence' grade. For further information refer to www.nzqa.govt.nz.

> In many schools, homes, and departments of education, test success is held as the ultimate goal, and students of mathematics often come to believe that their goal is to memorise numerous different, unrelated procedures, so that they can reproduce them when they are given different test questions. (Boaler, 2002, p. 10)

Increasingly, mathematics educators are concerned that students are able to learn mathematics at school for more than 10 years but are not able to transfer their knowledge and skills to unfamiliar out-of-classroom contexts (Boaler, 1998; Sfard, 2003). The need for mathematics reform has been recognised as an important component in the changing nature of school education (Goos, 2004). Notably, greater awareness of the situated nature of mathematics, and the role that social interaction plays in mathematics learning, have informed curriculum and pedagogical development in recent years (Anthony & Walshaw, 2007).

Despite the increasing influence of sociocultural learning theories within mathematics education, in-class assessment practices in secondary school mathematics are slow to align with changing views about how we learn (Pfannkuch, 2001; Watt, 2005). For many teachers, assessment remains an addition to, rather than an integral part of, learning, with few formative assessment opportunities for students (Crooks, 2006; Wiliam, Lee, Harrison, & Black, 2004). Teachers' internal assessment practices often emulate external summative assessments in the belief that this represents good assessment practice (Crooks, 1988; Harlen & Deakin Crick, 2003).

Maintaining traditional assessment practices can reduce the effectiveness of contemporary pedagogical models (James & Gipps, 1998). Given mathematics teachers' propensity to use traditional assessment strategies, further research is needed into how existing assessment practices can be enhanced, or adapted, to help students become aware and in control of their own mathematics learning. Theorising from a sociocultural perspective offers ways to explain practical implications for mathematics teachers' in-class assessment practices. Before outlining the specific problem to be addressed by this study, the purposes of assessment are briefly discussed.

1.2 Purposes of assessment

Contemporary assessment literature distinguishes between four purposes of assessment: summative, formative, diagnostic, and evaluative. From the earliest use of these terms, it was stressed that they applied not to the assessments themselves, but to the functions they served. The precise nature of these terms, however, is not consistently agreed upon in the literature (Black & Wiliam, 1998a).

The term 'summative' assessment is commonly defined as "assessments given at the end of units, mid-term and at the end of a course, which are designed to judge the extent of students' learning of the material in a course, for the purpose of grading, certification, evaluation of progress or even for researching the effectiveness of a curriculum" (Bloom, Hastings, & Madaus, 1971, p. 117). Summative assessments are designed primarily to inform teachers and employers, and are often used in some form of selection procedure.

Definitions of the term 'formative' assessment, however, are much more problematic. It is a term that has often been loosely defined and frequently misunderstood, especially at the classroom level (Black & Wiliam, 1998a). Historically, formative assessment was often used to distinguish between continuous summative assessment by teachers in the classroom and summative assessment by external examiners (Bell & Cowie, 2001). Early behaviourist views of formative assessment implied that it was primarily used by the teacher to adapt their teaching. From this perspective the learner, seen as a passive recipient, was unlikely to use assessment feedback to inform his or her own learning (Shepard, 2000).

Black (2001) claims that the emergence of a more modern view for formative assessment has been "a slow and tortured development, in part because its vision has been clouded repeatedly by the interference of summative testing, in part because it locates the functioning of assessment more closely within the complexities of pedagogy" (p. 74). In current literature, formative assessment is now generally understood to be assessment that promotes learning by providing feedback to the teacher and to the student about present understanding and skill development (Black

& Wiliam, 1998a). The principal purpose of formative feedback is to guide both the teacher and the student in co-developing the next steps in the learning process. This inclusive definition encapsulates a wide spectrum of teacher and student practices. Indeed, Torrance and Pryor (1998) argue that formative assessment is "a construct, a name that is given to what should more accurately be characterised as a social interaction between teacher and pupil which is intended to have a positive impact on learning" (p. 10). Theorising formative assessment in this way is consistent with the emerging sociocultural perspective of learning adopted in this research.

It has been argued that the ultimate goal of formative assessment should be to teach students to regulate their own learning (Black & Wiliam, 1998a; Gipps, 1994; Sadler, 1989, 1998). Given the strong link between assessment and learning, helping students to become increasingly in control of their learning requires the effective use of assessment. This extends beyond simply measuring students' progress against assessment criteria. We should prepare students for assessment events in such a way that students increasingly take on responsibility and become more autonomous in their learning. "Consequently, instead of being used to gain power over a child, assessment empowers that child" (Anderson, 1993, p. 103). To date, knowledge of pedagogical practices that encourage students to use assessment to become autonomous learners is still developing.

In recent years the terms 'assessment *for* learning' and 'assessment *of* learning' have gained prominence in the international literature. The essential difference between these two terms can be summarised as: "Assessment *for* learning is used to make decisions that affect teaching and learning in the short-term future, whereas assessment *of* learning is used to record and report what has been learned in the past" (Harlen, 2006, p. 104). These definitions are consistent with the definitions of formative and summative assessment cited above.

The final two purposes of assessment, are 'diagnostic' and 'evaluative' assessment. Harlen (2006) argues that diagnostic assessment, which is often taken to be concerned with identification of patterns of learning strengths and difficulties, is subsumed within the inclusive definition of formative assessment as detailed above. Evaluative assessment uses summative assessment information, as well as other information that

does not derive directly from assessment results, (e.g., document analysis, classroom observation) to report on the work of a class, teacher, school, or other group within the educational sector (Harlen, 2006). Evaluative assessment is concerned with the performance of groups of students rather than individual students, and will not be a focus of this research project.

1.2.1 A duality of purpose

Research evidence suggests that effective use of assessment supports and promotes learning (Black & Wiliam, 1998a; Carr, McGee, Jones, McKinley, Bell, Barr, & Simpson, 2000; Crooks, 1988; Natriello, 1987; Sadler, 1989). Simply increasing the amount of assessment, however, does not guarantee improved learning. Research indicates that formative, rather than summative assessment has the greater potential to improve learning (Black & Wiliam, 1998a; Crooks, 1988; Natriello, 1987; Sadler, 1989).

Several researchers (e.g., Broadfoot, Pollard, Osborn McNess, & Triggs, 1998; Harlen & James, 1997; Pollard, Triggs, Broadfoot, McNess, & Osborn, 2000; Reay & Wiliam, 1999) have argued that the summative and formative roles of assessment are not always compatible and that institutional demands mean that the summative role often takes precedence, resulting in limited use of formative assessment practices. Other researchers (e.g., Biggs, 1998; Black, 2000; Black & Wiliam, 1998a, 2006; Brookhart, 2001; Tanner & Jones, 2003) contend that the formative and summative purposes of assessment are not mutually exclusive. Successful students do not necessarily make neat distinctions between formative and summative assessment, and have learned to use feedback from assessment events, designed primarily to provide summative information, for their own formative purposes (Brookhart, 2001; Tanner & Jones, 2003).

To encapsulate the complex nature of the interrelationship between formative and summative assessment, Wiliam and Black (1996) proposed that, rather than a dichotomy, formative and summative assessment can be thought of as the ends of a continuum along which all assessment lies. Theorising assessment in this way explains how individual assessment events can serve a duality of purpose.

Accordingly, it is not the instrument that is formative or summative, but how, and when, the information is used, emphasising the distinction between the gathering of student achievement evidence and the interpretation of that evidence. This perspective reinforces the argument that, provided assessment is designed primarily from a formative approach, assessment events can serve both formative and summative purposes (DES/WO, 1988; Harlen, 2006; Wiliam & Black, 1996). Informed by this argument, the current project examines the formative potential of the summative exit qualification NCEA from a sociocultural perspective.

1.3 The specific problem

Providing students with valid and reliable judgements about the quality of their work, does not necessarily lead to improvement in learning and achievement (Sadler, 1989). Tanner and Jones (2003) claim that, although most mathematics students realise that they should learn from their assessments, the majority "only look at their mark" (p. 280). This well established norm is, according to Tanner and Jones, at the heart of the problem. In such circumstances students desire to "work out how to do better next time is likely to be restricted to unfocussed targets like 'try harder' or 'be more careful" (p. 280). Moreover, other researchers (e.g., Sadler, 1998; Tunstall & Gipps, 1996) also note that students often fail to recognise formative feedback as a helpful signal and guide, and remind us that it cannot be assumed that when students are given feedback they will know what to do with it.

The provision of feedback designed to be formative is a necessary, but not sufficient, first step in effective formative assessment. Within the sociocultual perspective of learning adopted in this research, feedback is considered formative only if students use it to improve their learning (Sadler, 1989). As Sadler notes:

> If the information is simply recorded, passed to a third party who lacks either the knowledge or the power to change the outcome, or is too deeply coded (for example, as a summary grade given by the teacher) to lead to appropriate action, the control loop cannot be closed, and 'dangling data' substituted for effective feedback. (p. 121)

Engagement with feedback is a complex and multi-facetted issue. Successful students' natural formative use of summative assessment information invites the

question: to what extent can the effective formative use of summative assessment be expanded to include all students? This research project considers this question and investigates systematic features and pedagogical strategies that encourage the formative use of high-stakes assessment information as a natural part of the learning process within a sociocultural framework. It seeks to answer the question: can a single assessment system effectively serve a duality of purpose and positively impact on students' learning for all students?

Given mathematics teachers' propensity to use traditional assessment methods in the senior secondary school (Pfannkuch, 2001: Watt, 2005), an investigation of this question will potentially expand the range of strategies available for teachers to improve mathematics learning for all students.

Current research on effective formative feedback has centred on specific characteristics of the feedback offered to students and not *how* the students engage with the feedback. Specifically, there is a paucity of research examining classroom activities and pedagogical practices that encourage both individual and collective students' engagement with feedback as a valid learning activity.

A key component of sociocultural theory is that knowledge is socially and contextually situated. Within the context of this project it is conjectured that students develop knowledge of *how* to engage effectively with formative feedback through the activities and practices that they are exposed to in their senior mathematics classroom. These activities and practices are influenced by the conceptions of assessment held by the teacher and the students, and collectively form the social structure that constitutes the classroom assessment environment. Theorising from a sociocultural perspective provides an opportunity to examine social structures and pedagogical practices that strengthen students' effective formative use of NCEA assessment events within New Zealand's unique educational context.

1.4 The research objective

This current research investigates the extent and nature of students' formative engagement with assessment information derived from NCEA assessment events. To

frame the research questions a theoretical model of formative assessment proposed by Sadler (1989) was adopted. Sadler's theoretical model consisted of three key learner-based components, namely:

> the learner has to (a) possess a concept of the *standard* (or goal, or reference level) being aimed for, (b) compare the *actual* (or current) *level of performance* with the standard, and (c) engage in appropriate *action* which leads to some closure of the gap. (p. 121, emphasis in original)

The three components of Sadler's theory are interconnected and are discussed in more depth in the literature review (chapter 2). Specifically, this research is guided by the following questions.

1. How do students derive understandings of the assessment criteria they are working towards and how does this impact on their learning?

2. What feedback do students receive from in-class assessment events and in what ways do students interpret this feedback?

3. In what ways do current assessment practices impact on students' engagement with feedback and subsequent learning behaviours?

The research is framed around students' perspectives since ultimately the action to improve learning must be taken by the student (Sadler, 1989). To date, insufficient attention has been paid to the 'student voice' concerning educational matters that directly affect them (Fielding, 2006) and this study partially addresses this concern. Given the situated nature of students' engagement with formative practices, and the situated nature of mathematics learning, a case study approach is used to investigate students' perceptions of the formative potential of NCEA. Three Y12 mathematics classes from a medium to large (1300 students) decile 7[2] urban secondary school formed the case study singularity for this study. Focus group interviews with nine students were conducted across the year, complemented by classroom observations, a

[2] The Ministry of Education uses a decile rating system for school funding purposes. Each decile contains approximately 10% of schools. Schools in decile 1 have the highest proportion of students from low socio-economic backgrounds. Schools in decile 10 have the lowest proportions of students from low socio-economic backgrounds.

focus group interview with the teachers, and a quantitative questionnaire with the three mathematics classes. This approach is discussed in more detail in methodology chapter (chapter 3).

This project does not seek to examine all of the interactions involving teachers and students that might be classified as formative feedback. This project concentrates on the interactions resulting from structured formal assessment events that form part of the students' progress towards NCEA. This includes unit assignments, end of unit assessments, and the written and oral feedback offered to students subsequent to these assessment tasks.

1.5 Summary

This chapter has identified that assessment is an integral part of learning. A major direction in current learning theory is the recognition of the situated nature of knowledge and the social dimension of learning. As such, improving our understanding of the students' perspective of formative assessment in a sociocultural learning environment contributes to the development of a more complete theory of how students learn.

Increasing calls for reforms in how school-mathematics is learned and assessed has added to the impetus for this research. Given mathematics teachers' preferential assessment practices, greater understanding of how existing practices can be more effectively used as valid learning activities is an important addition to our current knowledge base.

New Zealand has recently introduced a new SBA high-stakes summative exit qualification, the NCEA. Rather than seeing formative and summative assessment as mutually exclusive events, this chapter has argued that these two purposes of assessment can co-exist in a single assessment system. Accordingly, this project adopts a theoretical framework proposed by Sadler (1989) to examine students' perspectives of the formative potential of NCEA from a sociocultural perspective of learning.

1.6 Structure of this thesis.

The following literature review (chapter 2) situates the current study in the contemporary assessment research. Chapter 3 outlines and justifies the methodology used in this project. Chapter 4 details both the qualitative and quantitative results of the current project. Chapter 5 discusses the results with reference to key themes identified in the literature review; conclusions are drawn, and limitations of the current study and suggestions for future research are explored.

2 LITERATURE REVIEW

2.1 Introduction

This chapter examines research literature on relevant theories of learning and assessment. It initially looks at an historical overview of the changes in secondary schools assessment systems in New Zealand. This situates the present study in a historical context. Section 2.3 draws on the research associated with three broad families of learning theories and examines their inherent conceptions of mathematics and assessment. This is followed by an examination of the research literature surrounding the impact of high-stakes assessment on learning (section 2.4). Section 2.5 discusses the contention that, rather than being thought of as a dichotomy, formative and summative assessment can be more usefully thought of as end points of a continuum along which all assessment lies. The final section (2.6) examines the research literature supporting Sadler's (1989) model for effective formative assessment. This model has been adopted in this research project as a theoretical framework supporting the examination of the key research questions.

2.2 Assessment systems in New Zealand

During the 1950s and 1960s high-stakes assessment in New Zealand secondary schools developed into a system governed predominantly by three years of externally assessed national exams, starting at Year 11 (Y11). All of these examinations were Norm-Referenced in nature, with a specified quota of passes, combined with inter-subject statistical scaling (Crooks, 2002). During this period the school leaving age was set at 15 years. At a time when New Zealand was considered to be an affluent society where graduating students could readily get a job, the senior years of secondary school could be characterised as being highly selective. There was no social promotion through the senior year groups; students had to 'pass' their Y11 School Certificate course by securing an average of 50% or better in four subjects. Students were then able to progress to the Y12 University Entrance course, which they had to pass before moving onto an optional Y13 in preparation for university

study. At the end of Y13 students sat national external examinations with University Bursaries and University Scholarships being awarded based on mark aggregates across five subjects. Students who did not pass a particular course of study had to repeat the entire year in all subjects. Less than 10% of the Y9 intake went on to complete their Y13 (Crooks, 2002).

By the 1970s there was growing concern that traditional assessment strategies did little to encourage the development of skills required for a modern society. Critics argued that existing qualifications did not fit either the curriculum, the needs of users, or the full range of students in New Zealand schools (Crooks, 2002). Central to these debates were; changes to conceptualisations of learning, a changed direction and purpose for education, recognition of the impact of assessment on learning, dissatisfaction with present assessment practices, and ongoing attempts to improve the quality of assessment (Strachan, 2001). The School Certificate Examinations board consulted widely throughout the 1970s about the future shape of the senior secondary school qualification system (Lennox, 2001). The removal, or replacement, of School Certificate was called for by a number of national educational bodies, (e.g., the Post-Primary Teachers' Association (PPTA), the Secondary Principals' Association, and the Department of Education). The limitations of School Certificate, and a preference for assessment against 'Standards', were also noted in a survey of employers' groups in the early 1980s and the New Zealand Employees' Federation submission to the Committee of Inquiry into Curriculum, Assessment and Qualifications, in 1986 (Lennox, 2001).

By the late 1980s a number of societal and structural changes had occurred in the senior secondary school. For example, the relaxing of official pass-marks in favour of subject promotion meant that schools had a greater number of multilevel learners. Additionally, it was more difficult for students to gain employment without suitable qualifications and, as a result, students were staying longer at school. Approximately 40% of Year 9 students were now staying on to Y13, although a significant proportion of the Y13 students were not intending to progress to university study (Crooks, 2002).

While School Certificate and University Bursary still existed, a number of courses now included some internal assessment, statistically moderated against the

performance of a school's student cohort on the final examination. An internally assessed qualification, the Sixth Form Certificate, was introduced to replace the University Entrance Examination, with grades also statistically moderated by the cohort of students' School Certificate results. A constant in all of this change was that all three years of high-stakes national assessments were still Norm-Referenced and predominantly summative in nature. In addition to the nationally accredited course, schools were developing alternative programmes to cater for the wide range of abilities, and career aspirations, amongst their student body. This was particularly prevalent in mathematics and whilst many of these alternative mathematics programmes were regionally accredited (e.g., Manawatu Mathematics Certificate), none were recognised as national qualifications.

In the early 1990s the New Zealand Ministry of Education (MoE) identified that the primary purpose of school-based assessment was "to improve students' learning and the quality of learning programmes" (MoE, 1993, p. 24). The Ministry of Education provided a number of curriculum policy and support documents, recognising and emphasising the importance of assessment in informing learning (e.g., Assessment for Better Learning: A public discussion document, 1989; Assessment: Policy to Practice, 1994; The New Zealand Curriculum Framework, 1993). Mathematics in the New Zealand Curriculum (MoE, 1992) reflected this focus by including suggested assessment activities for the specified learning outcomes within the main body of the curriculum document.

In 1989, the Government and educational sector embarked on an ambitious structural reform to create a seamless qualification system: The National Qualification Framework (NQF). It was intended that this framework would incorporate all upper-secondary, vocational, and tertiary qualifications. In the following year the Government created a new organisation—the New Zealand Qualification Authority (NZQA)—to oversee the development, and management, of the NQF. The framework was designed to provide as many students as possible with sensible and varied pathways towards appropriate qualifications (Philips, 2003). Hopes were also expressed during the reforms that the NQF would prove "more engaging for students because it could acknowledge a much wider range of levels and types of learning success" (Hipkins et al., 2007, p. 2). The task of implementing the NQF was

complicated by the number of educational providers involved, and their well-established, but often incompatible, systems and practices. To date, there are over 800 National Certificates, mostly vocational, registered on the Framework.

NZQA adopted a unitary Standards-Based approach for the assessment of all qualifications on the NQF. Assessment was via a tool called 'Unit Standards' which clustered together learning outcomes and their associated assessment criteria. Students would either 'pass' or 'fail' the Unit Standard with no recognition of individual excellence. Students could have multiple attempts to pass a given Standard, allowing them the opportunity to further their learning between attempts. Provision was also made for recognition of prior learning, with students able to be assessed without being required to complete the learning programme associated with that Unit Standard. Each Unit Standard was worth a specified number of 'credits', designed to reflect the relative amount of content assessed by the Standard. These credits accrued towards National Certificates.

Despite the fact that the Educational Development Conference in 1974 had called for the abandonment of external examinations within 5 years (Strachan, 2001), and the employers' survey of 1981 had supported teacher assessment using a range of methods (Lennox, 2001), the newly developed NQF did not find universal favour. The implementation of the NQF was conducted in an environment of debate over the relative merits of Standards-Based and Norm-Referenced methods of assessment (Dobric, 2006). Although welcomed by some areas of trade and vocational education, where its mastery approach seemed a suitable choice, it was not welcomed by all in the educational community (Codd, McAlpine & Poskitt, 1995; Crooks, 2002). Resistance in the tertiary sector restricted the widespread adoption of this NQF by this sector (Vlaardingerbroek, 2006).

In spite of expressed concerns, assessment by Unit Standards was extended to traditional school subjects in 1995. Initially, schools could assess students against the NQF as well as against the three existing conventional qualifications (Vlaadingerbroek, 2006). Not unexpectedly, this change to senior secondary school qualifications was greeted with a wave of protest (Lee & Lee, 2000; Philips, 2003). Critics argued that the lack of provision for graduation of excellence gave students

insufficient encouragement to excel, and therefore Unit Standards were not suited to academic courses. Concerns were also raised that, without national examinations, the reliability of assessment information and qualifications would be compromised (Lee & Lee, 2000; Philips, 2003).

NCEA

In 1998, partially in response to these criticisms, the Achievement 2001 Policy announced the development of the NCEA. This policy introduced a new Standards Based Assessment (SBA) tool—Achievement Standards—to complement the existing Unit Standards. Achievement Standards provided for the recognition of a range of achievement—Not Achieved, Achieved, Achieved with Merit, and Achieved with Excellence— and consisted of both internally and externally assessed Standards.

NCEA has three levels: Level 1 has replaced School Certificate (Y11), Level 2, Sixth Form Certificate (Y12), and Level 3, University Bursary (Y13). Students sit a variety of internal Achievement Standards or Unit Standards during the year and external Achievement Standards at the end of the year, usually in the form of end of year examinations. Additionally, students sit 'practice' assessments for the external Achievement Standards during the course of the year. For internally assessed Standards, students can be offered as many reassessment opportunities as the provider can practically accommodate. Students' results are reported for each individual Standard, with each Standard worth a prescribed number of credits that accrue towards a level certificate of NCEA. NCEA also saw the removal of the compulsory course content present in the previous system. Accordingly, schools were now freely able to devise programmes of study assessed by any combination of the two types of assessment Standards.

In policy documents, NCEA is strongly underpinned by a philosophical position of enhancing students' learning. During the political launch of NCEA a number of key statements that indicated support for embedding assessment in learning were made. "All of this is about learning first. NCEA results are a consequence of enhanced learning opportunities" (Meek, 2001, p. 4). Consistent with the MoE's National Education Goals, NCEA is purported to provide schools with an "enhanced

flexibility…to offer broader and deeper learning for all students" (Mallard, 2001, p. 5). In supporting the Government's stated goals of encouraging and supporting lifelong learning, both for the development of the individual as well as the wider community, NCEA was heralded as "the crucial first step on the ladder of lifelong learning, and we simply can't be a prosperous, confident and assertive knowledge society and knowledge economy without lifelong learning" (West, 2001, p. 3). Such comments highlight an increasing recognition of the importance of assessment for learning.

The introduction of NCEA has arguably been the most contested part of the NQF reforms "perhaps because high-stakes assessment at the senior secondary school-level is more visible in the wider community than are tertiary-level assessment practices" (Hipkins et al., 2007, p. 4). Certainly the negative commentary associated with NCEA has far outweighed the positive commentary in the media (Brooking, 2006). However, Hipkins et al. (2007) contest that many of the concerns expressed in the media are not new to NCEA; they had existed for a number of years, hidden within the less visible functioning of the older Norm-Referenced assessment system.

2.3 Learning and assessment

Underlying all discussions of assessment are "assumptions about the psychology of learning" (Black & Wiliam, 1998a, p. 16). In recent years traditional views of assessment have been challenged through increased recognition of the interactions between assessment and classroom learning. Assessment is not an isolated activity operating independently of teaching and learning practices (Crooks, 1988). How we construe learning influences how we conceptualise teaching as an activity, and how we conceptualise assessment practices (Harlen, 2006; James, 2006). The last 20 years have seen many debates and changes in how learning is viewed (Lerman, 2000; Sfard, 1998). However, these changes in the conceptions of learning have not been fully realised in current assessment practices in mathematics (Cavanagh, 2006, Hosking & Shield, 2001; Pfannkuch, 2001; Watt, 2005).

Conceptualisations of learning, and the implications for teaching and assessment, are complex, and theoretical perspectives often overlap when trying to understand how we learn new things. It must be emphasised that approaches to learning, like forms of understanding, are situational; they describe responses of the moment, not enduring characteristics of students. Furthermore, a synthesis of learning theories must be cognisant of both individual and collective learning (Anthony & Walshaw, 2007).

The following discussion examines three broad clusters of learning theories. The literature from the United States labels these clusters as 'behaviourist', 'cognitive' and 'situated', however in the United Kingdom they are more commonly referred to as 'behaviourist', 'constructivist', and 'sociocultural' or 'activist'. The discussion examines how these theories of learning manifest themselves in classroom practice, the perceptions of mathematics learning inherent in each, and the nature and role of assessment.

2.3.1 Behaviourist theories of learning.

Behaviourist theories of learning emerged strongly in the 1930s and remained a dominant theoretical perspective into the 1960s and 1970s. Behaviourist theories construe learning as the conditioned response to external stimuli. These theories do not take into account human consciousness to explain learning; they are only interested in observable behaviour (Skinner, 1976; Thorndike, 1924). Within the New Zealand context, pedagogical approaches in mathematics classrooms remained heavily influenced by behaviourist views of learning into the late 1980s (Begg, 1998).

Within this paradigm, school mathematics is regarded as an absolute body of knowledge without a cultural context, that is to be transmitted to students (Leder & Forgasz, 1992). Mathematical knowledge is conceptualised as a collection of hierarchically ordered facts with complex knowledge conceived to be made up of individual elements. The teacher's role is to train students to respond to instruction correctly and quickly. Correct responses are encouraged by the provision of rewards, and incorrect responses ignored so the response extinguishes. The teacher's focus is to present correct procedures and provide opportunities for practice. Accordingly,

students are not offered the opportunities to explore or construct alternative mathematical procedures.

> A basic assumption is that any practice of a wrong association tends to strengthen it. Therefore, it is essential to prevent students from making mistakes or being exposed to errors made by peers. Consequently, students rarely interact or collaborate with each other (Even & Tirosh, 2002, p. 231).

Assessment under behaviourist theories focuses on the examination of observable behaviour as a means of providing evidence that some learning objective has been met. Assessment tasks that concentrate on knowledge of mathematical definitions and procedures, rather than conceptual understanding, act as a tool for checking whether the information has been received, absorbed and memorised (Begg, 1998; Gipps, 1994). Checking or marking answers is focused on interpretation of correctness or incorrectness. As such, feedback is often limited to an indication of an incorrect response, or sometimes the provision of a correct answer with little indication of how to progress from the student's answer to the required answer (Neyland, 1996). Positive feedback, often in the form of non-specific, ego-oriented praise by teachers, is often used as part of the rewards system within the classroom (James, 2006).

In this model, classroom formative assessment practices are conceived largely as a process by which the teacher monitors and controls students' learning by adapting future teaching of the content, often in subsequent years to a different group of students. Learners are seen as passive participants—they are not in a position to either interpret assessment information, or to use it to inform the next steps in the learning process (James, 2006).

2.3.2 Cognitive, constructivist theories of learning.

Constructivist theories of learning gained prominence in the 1960s and received extensive attention through the 1990s for their implications for teaching and assessment (von Glaserfeld, 1995). Constructivism builds on Piaget's ideas about the development and construction of knowledge. A basic assumption is that knowledge is gained by an active process of construction, rather than a passive assimilation of information or rote memorisation (Even & Tirosh, 2002). Thus knowledge is not a

representation of an independently existing real world but is made up of conceptual structures which are adapted to the individual's range of experiences (Anthony & Walshaw, 2003). Learning is understood as a process of conceptual growth and growth in general cognitive abilities, such as problem solving strategies and metacognitive processes.

Metacognition is considered to have two main components. The first of these—self-monitoring—refers to students' awareness of their current understandings. The second element identifies self-regulating mechanisms which involve planning what to do next, checking the outcomes of strategies employed, and evaluating and revising strategies (Black, McCormick, James, & Pedder, 2006). Successful mathematics learners monitor their own understandings (Anthony, 1994). As 'intentional' learners (Bereiter & Scardamalia, 1989) they take responsibility for their learning, consciously choosing the strategy and direction for their learning.

Under these theories, conceptions of mathematics changed from being an absolute body of knowledge that could be transmitted to "a relativist perspective where individuals construct their own mathematical schemas to help them gain other ways of viewing their experiences and making sense of their world" (Begg, 1998, p. 8). This move to what has been termed a 'fallibilist' philopophy (Ernest, 2004)—where mathematical knowledge is understood to be fallible and eternally open to revision—accepts that mathematics is made up of overlapping structures that are open to adaptation and change. Rather than being thought of as something that is either gained or not gained, mathematical knowledge can be more accurately characterised as 'partial'. It is something that develops as concepts become better understood and stronger links with other mathematical concepts are formed (Sfard, 2003).

Significantly, constructivism challenged traditional teaching and assessment processes. The role of the teacher, within a constructivist learning environment, changed to helping students move from being 'novices' to being 'experts' at organising mathematical knowledge into structures that make it more meaningful and more useful. Accordingly, curricula advocated problem solving and situations requiring inductive and deductive reasoning as suitable contexts for constructing mathematical knowledge.

Assessment moved to reflect the complex, and often partial nature, of students' mathematical knowledge. The focus shifted towards determining "what students know within their complex webs of ideas" (Begg, 1998, p. 9), and providing them credit for what they can do rather than assessing what they can not do. When assessments are viewed as "key contributors to students' awareness of their own learning and to increasing their ability towards reflective abstraction" (Confrey & Kazak, 2006, p. 329) the role of feedback changes. In addition to evaluating students' current understanding, it is designed to scaffold students' future learning.

Within the mathematics education environment, identification with constructivist principles saw changes in the nature of assessment tasks, focusing on increased understanding of students' reasoning. Multiple choice and short answer questions designed to assess the acquisition of specific low level mathematical knowledge were complemented by a range of diverse assessment types designed to assess a deeper level of thinking, for example: diagnostic tasks, investigative activities, portfolios, practical assessments, and small group projects (Ruthven, 2002).

Early constructivist perspectives saw learning as an individual cognitive process. During the 1990s these theories were expanded into social learning theories that emphasised the role of social processes in the construction of knowledge (Sfard, 1998). This expansion was consistent with "a clear shift away from viewing mathematics learning as *acquisition* toward understanding mathematics learning as *participation* in the discursive and cultural practices of a community" (Goos, 2004, pp. 261-262, emphasis in original).

2.3.3 Sociocultural theories of learning.

Sociocultural theories of learning turn the focus away from individual attributes towards broader communities of practice (Boaler, 1999; Sfard, 1998). Social learning theories focus on the intrinsically social and situated nature of cognition. Rather than asking what kinds of cognitive processes and conceptual structures are involved in learning, sociocultural perspectives ask "what kinds of social engagements provide the proper context for learning to take place" (Even & Tirosh, 2002, p. 232).

According to Anthony and Walshaw (2003), sociocultural theories of learning are actually theories of social practice, of which learning is one of many dimensions.

Jean Lave has been instrumental in the critical attention that sociocultural theories have received in recent years (see Lave, 1988; Lave & Wenger, 1991). Central to Lave's theory is the notion of a 'community of practice'.

> A community of practice is an intrinsic condition for the existence of knowledge not least because it provides the interpretive support necessary for making sense of its heritage. Thus, participation in the cultural practice in which any knowledge exists is an epistemological principle of learning. The social structure of this practice, its power relations, and its conditions for legitimacy define possibilities for learning. (Lave & Wenger, 1991, p. 98)

From a sociocultural perspective, the learner is both shaping and being shaped by the 'community of practice' (Lave & Wenger, 1991). Learning, by definition, is a "social and collaborative activity in which people develop their thinking together" (James, 2006, p. 57). This does not mean that other theories (e.g., social constructivism) have ignored the social factors in learning, rather sociocultural theories conceive thinking and reasoning as social activities that extend the idea that social interaction stimulates an individual's internal cognitive activity. Learning is perceived as "something that takes place in a participation framework, not in an isolated individual's mind" (Even & Tirosh, 2002, p. 232).

Assessment activities under these theories are embedded in learning, which is embedded in the sociocultural activities of the classroom. This means that assessment is a dynamic process that provides a prospective measure of performance including abilities that are developing, and is predictive of how the child might perform independently in the future (Palincsar, 1998). As such, formative assessment is a central feature of conceptions of assessment within sociocultural learning theories.

Sociocultural theories of learning argue that participation in collaborative activities, discourse and social interaction are central to the development of metacognitive processes critical to student engagement with formative assessment practices (Wood, 1998). Wood asserts that effective group interactions encourage students to think

about their understandings, requiring students to plan, organise, and evaluate their learning, both collectively and individually.

The notion that effective engagement with formative assessment is socially situated advances our understanding beyond treating assessment as a one-sided process controlled by the teacher to a consideration of teacher-student and student-student dynamics (Ross, Rolheister, & Hogaboam-Gray, 2002). Within a sociocultural learning environment the teacher co-participates as a learner in a community of learners, and so the relationships between teacher and students are developed in less hierarchical ways. Accordingly, information gained from assessments is used interactively between student and teacher and between peers, helping learners to become aware of, and in control of, their own learning (Torrance & Pryor, 1998). Teachers' assessments of students' understanding sit alongside peer- and self-assessment as central parts of the social processes "that mediate the development of intellectual abilities, construction of knowledge and formation of students' identities" (Shepard, 2000, p. 4).

Central to sociocultural theories of learning, Vygotsky's (1978) notion of the Zone of Proximal Development (ZPD) has implications for the nature of formative feedback in a sociocultural learning environment. The ZPD draws attention to the gap between what the learner can achieve without help and what may be achieved with suitable help. Wood, Bruner, and Ross's (1976) concept of 'scaffolding' and Rogoff's (1990) broader notion of 'guided participation' enhance the ZPD model, and emphasise the teacher's role in setting learning goals that are achievable by the learner with appropriate help, and then providing such help through effective pedagogies and formative feedback practices (Black, 2000). The implication of this view is that, rather than supplying students with predetermined solution paths, effective formative feedback should move students into the ZPD and encourage them to actively engage with the feedback.

Sociocultural theories of learning promote the increased use of alternative assessment practices that take account of the social and cultural context in which learning occurs. Despite recognition of practical implementation challenges (Watson, 2006), the use of self- and peer-assessment, observations, portfolios, practical assessments,

investigations, and small group projects have all been highlighted as valid methods of gathering information about students' achievement (Bourke, 2000, Pfannkuch, 2001; Ruthven, 2002; Watson, 2004; Wiliam et al., 2004).

To date, assessment practices are slow to align with changing views about how we learn. Reynolds and Trehan (2000) point to the way in which contemporary approaches to learning often use 'traditional' methods of assessment without thinking through the contradiction between the assumptions behind such methods and the espoused learning objectives. Egan (1997) concludes a discussion on perspectives of contemporary learning theories with the statement: "We will, of course, want evidence that students have learned the content that has made up the lesson or unit of study; this can be achieved through traditional techniques" (p. 272).

Mathematics learning.

Increasingly, mathematical learning and knowledge is understood relationally, between people and settings, and is seen as situated within, and as a product of, those social situations (Lerman, 2000). For example, Lave (1988) found that adults did not make use of their school-learned mathematics in shopping situations, and indeed did not see the similarity between the mathematics used for shopping and the mathematics they learned in school. She concluded that the mathematics used was more dependent on the situation or context than on the actual mathematical demands within the tasks. This, in part, led her to use the phrase 'situated learning' to describe how learning is linked to the situation or context in which it takes place. Similarly, Nunes and Bryant (1996) examined the mathematical problem solving skills of Brazilian children acting as street corner vendors. Despite limited formal schooling, these children developed situated knowledge of effective mathematical strategies. The researchers concluded that the traditional mathematics learned in schools does not seem to transfer readily to out-of-school tasks, and a reliance on standard algorithms may not be useful in solving practical mathematical activities.

Greater recognition that traditional school mathematics taught within an acquisition model of learning inadequately prepares students for out-of-school mathematical situations has led to calls for significant changes in the way that mathematics is taught

and assessed in schools (Goos, 2004). Curriculum documents from around the world (e.g., Curriculum and Evaluation Standards for School Mathematics (NCTM, 1989); Mathematics in the New Zealand Curriculum, (MoE, 1992); Mathematics: A Curriculum Profile for Australian schools (Australian Education Council, 1994); Principles and Standards for School Mathematics (NCTM, 2000)) all promote an increased emphasis on the processes of problem solving, reasoning, communicating mathematical ideas, and "allowing student to experience the actual process through which mathematics develops" (Goos, 2004, p. 258). Sociocultural perspectives of learning have provided a theoretical rationale and framework for suggested reforms in mathematics education (Forman, 2003).

From a sociocultural perspective, mathematics learning can be theorised as a social and communicative activity encouraging "progress to more complete participation in the practices, beliefs, conventions and values of communities of practitioners" (Goos, Galbraith, & Renshaw, 2004, p. 91). This perspective acknowledges and builds on the notion that students' mathematical knowledge is partial and developing; a notion central to the design of the current research project. Additionally, this perspective implies a change in emphasis from the traditional content orientation of mathematics learning towards a more process oriented approach focused on mathematical practices as well as mathematical concepts and skills (Even & Tirosh, 2002). Making connections, hypothesising, conjecturing, and refutation are key processes of mathematics practitioners (Burton, 1999). Accordingly "the act of making connections is not something students need to *know*, it is something they need to *do*" (Boaler, 2002, p.12, emphasis in original).

Pedagogical practices, including assessment practices, consistent with sociocultural theories of learning create classroom environments where students are actively involved in all aspects of the mathematics learning process (Black et al., 2006). Black et al. report that eliciting contributions from students and encouraging them to listen to the contributions of others is a key pedagogical practice that encourages students' engagement and autonomy in learning. By "encouraging teachers to frame their questioning so that it explores key features of learning, and to encourage all pupils to contribute and share ideas whether or not they are confident that they are correct" (p. 128) helps legitimise and build on students' existing knowledge. This practice is

underpinned by the principle that conceptual change must evolve from the learners' existing understandings, that students must be actively involved in the learning, and that such involvement should take place in social and community discourse.

A study by Hiebert et al. (1997) described how a mathematics teacher created an environment where all students' contributions were supported and encouraged by both the teacher and the other students. This 'caring' culture promoted mathematics learning for all of the members in the group. Similarly, a study by Watson (2002) with low attaining secondary students identified that teachers perceive that students prefer to learn in a 'togetherness' environment in which teaching fosters an awareness of learning encouraging and supporting students to take risks. Within such a supportive environment, as students interact and question one another, the focus can move from the teacher to the students.

Reform approaches to teaching mathematics, focused on inquiry communities of practice, can encourage flexible mathematical thinkers who can apply the mathematics they know to a range of unfamiliar situations Boaler (1998). Boaler conducted a three year case study looking at two secondary schools in England. One school (Phoenix Park) used a process-based 'reform' approach for the teaching of mathematics, emphasising higher-order thinking, and students' responsibility for their own learning. The other school (Amber Hill) used a 'traditional' content-based approach emphasising practice of test items. Notably, in contrast to the students at Amber Hill, students at Phoenix Park were not required to work on their mathematics for a set period of time each day but judged for themselves when they understood the mathematics involved before moving on to other activities or subjects. This resulted in less class time being spent on mathematics at Phoenix Park.

Although matched in terms of prior achievement, students at Phoenix Park outperformed those at Amber Hill in the national school-leaving examination (the General Certificate of Secondary Education, or GCSE) by, on average, one third of a grade (equivalent to a standardized effect size of 0.21). This success is despite the fact that the students at Phoenix Park had not encountered all of the areas of mathematics assessed in the examination, and that the GCSE examination was markedly different from anything they were accustomed to at Phoenix Park. Students

at Amber Hill reported that they had problems interpreting the unfamiliar questions in the examinations and could not see how the procedures they had learned in their mathematics classes could be applied to these questions. In contrast, Phoenix park students felt their success in the examinations was due to their ability to think about unfamiliar questions and determine how the mathematics they knew applied to those situations.

The significance of Boaler's (1998) study is that it was the way students learned mathematics at Phoenix Park that helped them develop as flexible mathematical thinkers. Indeed, Boaler argued that the students at Phoenix Park did not know *more* mathematics than the students at Amber Hill, rather they were able to apply the mathematics they knew more effectively to a wider range of mathematical problems. Students at Amber Hill developed procedural knowledge that was little use to them in unfamiliar situations. In contrast, students at Phoenix Park developed a strong conceptual understanding which advantaged them in a range of assessments and a variety of familiar and unfamiliar situations.

In addition to performing well in the 'traditional' examination, the Phoenix Park students performed significantly better on a series of 'applied' questions they were given during the course of the research. These questions required the students to synthesise and apply the mathematics they knew to unfamiliar situations. Boaler (1998) concluded that; "a traditional textbook approach that emphasises computational rules and procedures at the expense of depth of understanding, is disadvantageous to students primarily because it encourages learning that is inflexible, school-bound and of limited use" (p. 60).

Assessment practices in mathematics

As noted earlier, in addition to being thought of as a route to further knowledge, the act of drawing connections is a key aspect of mathematical work in itself (Burton, 1999). In fact, many researchers (e.g., Boaler, 2002; Ernest, 2004; Hiebert et al., 1997) claim that it is through the process of making connections that students derive mathematical meaning. As such, formative feedback in mathematics should be designed to allow students to collaboratively make these connections themselves,

rather than being shown how to answer the question by their teacher. In effect, 'working it out' is a key part of students' mathematical learning, and formative feedback designed to place the student in the ZPD has a critical part to play in 'working it out'.

Teachers' content knowledge affects their ability to help students make mathematical connections (Askew, Brown, Rhodes, Wiliam, & Johnson, 1997). In a study by Khisty and Chval (2002), teachers with good knowledge and awareness of conceptual connections and who focused on mathematical connections and meaning significantly developed their students' mathematical reasoning. Conversely, in a study of statistics knowledge, Burgess (2007) found that teachers without a sound conceptual understanding of statistics missed opportunities to enhance their students' learning through making connections. These findings have implications for teachers' abilities to provide effective formative feedback for two reasons. Firstly, teachers need to be able to decipher students' current understanding and thinking from their assessed work. Secondly, teachers must then provide students with appropriate targeted feedback that places the student in the ZPD. Sound content knowledge enhances their ability to undertake these actions effectively.

Sociocultural theories of learning promote the use of alternative assessment methods, however the use of alternative assessment in senior secondary school mathematics classrooms is not common (Pfannkuch, 2001). Investigating New Zealand teachers' perceptions and practices of assessment in school mathematics in the mid 1990s, Pfannkuch found that although secondary teachers reported using a variety of alternate assessment strategies, these were mostly limited to their non-examination classes. Traditional examinations mirroring the look of the external summative national examinations were still the most common assessment types used in senior secondary school mathematics classes. Pfannkuch commented that, rather than providing evidence of individual student's progress, "the secondary school's attention is focussed on providing assessment information that indicates how a student will fare in a national examination" (p. 199). Despite awareness that a single mark did not adequately describe the mathematical performance of a student, Pfannkuch found that the format of many school reports to parents and students appeared to undermine

attempts to gather a wide range of assessment information. Pfannkuch's study was conducted in 1995 and, to date there has been no follow up study conducted.

A recent study by Brown (2006) investigated the relationship between students' conceptions of assessment and mathematics achievement using self-report inventories with four New Zealand secondary schools. Brown identified that the conceptions that assessment makes students accountable and was beneficial for students loaded positively on achievement, while the conceptions that assessment is fun and assessment is ignored had negative loadings on achievement. Brown concluded that students' use of assessment formatively to take responsibility for their learning positively affects students' mathematical learning.

Also investigating teachers' perceptions of methods of assessment in mathematics, Watt (2005) surveyed 60 teachers from 11 secondary schools in metropolitan Sydney. The teachers reported using traditional mathematics tests as their main assessment method; they perceive this to be a valid measure of students' ability, particularly at the senior school level. When the results were analysed based upon how long a teacher had been teaching, those teachers with 20+ years of teaching were more likely to prefer traditional tests and more likely to consider alternate methods of assessment as subjective than those teachers with less than 10 years teaching experience. Teachers with less than 10 years experience suggested that the explicit introduction of pedagogical strategies such as group work and practical work might provide opportunities for greater use of alternate assessment formats. This group of teachers did, however, express concern about subjectivity in alternate assessment strategies. Moreover, all teachers perceived that the nature of the senior curriculum was less conducive to using alternate assessment strategies.

Lewis's (2005) study involving senior mathematics teachers in New Zealand noted that teachers' felt that only some content topics (geometry, trigonometry, measurement and statistics and probability) could be suitably assessed using SBA. Teachers argued that the rationale for a SBA component is to assess those components of the curriculum which are not appropriately, adequately, and validly assessed in a traditional examination format.

These studies suggest that despite curriculum and nationally mandated changes to assessment formats, teachers' beliefs in senior mathematics classrooms remain relatively traditional in nature.

2.4 Impact of high-stakes assessment

High-stakes summative assessment impacts on teachers' pedagogy, curriculum and students' learning (Zepke et al., 2006). A range of studies show that the impact can be positive or negative, depending on the extent of the high-stakes pressure exerted on school systems, and the nature of competition between students inherent in the design of the assessment system.

High-stakes summative assessment can negatively impact on teachers' formative practices (Black & Wiliam, 1998a; Crooks, 1988; Harlen & Deakin Crick, 2003; Johnston & McClune, 2000). In a study by Pollard et al. (2000), investigating the introduction of high-stakes national testing in English primary schools, the researchers found that the impact of high-stakes assessment on teachers was to drive out formative practices. Interviews with students indicated that they were only aware of assessment as a summative activity; there was no evidence from students that teachers were communicating any formative or diagnostic assessment to them. Although teachers intended their assessment interactions to be essentially formative, subtle changes in their discourse with the students indicated a summative performance-related stance, which was implicitly communicated to the students. Students reported test anxiety, uncertainty, tension, and a reduction in confidence in their own self-assessment. As a result, students were more likely to attribute success or failure to innate characteristics, such as luck, and were less positive about assessment information that revealed their weaknesses.

Test anxiety and negative consequences for self-esteem have been observed across several studies. A study by Davies and Brember (1998) found that the introduction of National Curriculum testing in English primary schools coincided with a change in the relationship between students' self-esteem and achievement as measured on standardized tests in maths and reading. Prior to the introduction of national testing

there appeared to be no correlation between these variables. Post introduction revealed a positive correlation between self-esteem and achievement.

Primary school students' experiences during the month prior to sitting the National Curriculum tests in England were also investigated by Reay and Wiliam (1999). The researchers described the classes they observed as being at 'fever pitch' due to the impending tests. Despite the fact that the assessments were ostensibly low-stake for the students, they were high-stakes for the school. The schools were held accountable for the attainment of their students, and charged with raising achievement scores in subsequent years. Teachers were observed to be anxious, and berated children for poor performance in practice tests. Even though students recognised that the testing regime was primarily concerned with the performance of the teachers and the school, they reported being worried about their own performance and the possible consequences for their future. Additionally, nearly all the students "indicated a sense of unease and feelings of discomfort about what the SATs might reveal about themselves as learners" (p. 346). The researchers suggest that preparing for the National Curriculum tests created highly anxious, performance oriented students, and argued that "the narrowing of the focus of assessment, together with an emphasis on achieving the highest scores possible, produces a situation in which unjustifiable educational practices are not only possible, but encouraged" (pp. 352-353).

Extreme test anxiety was also observed in a study by Leonard and Davey (2001). Investigating students' perceptions of the national '11-plus' test in Northern Ireland, the researchers found that students identified the tests as the cause of extreme anxiety with, at times, significantly detrimental effects to the self-esteem of those students who did not meet their own, or others, expectations. Students also reported that they tended to make value judgements about other students' intelligence based solely on the grades they received. The authors noted that this negative feature was all the more regrettable given that the measures were known to be unreliable.

The effects of high-stakes assessment on curriculum have been noted in a number of studies. Reay and Wiliam's (1999) study into year 6 pupils leading up to Key Stage 2 National Curriculum tests observed a significant narrowing of the curriculum. The classroom content emphasised only the material covered by the tests, to the detriment

of school subjects that were not covered by the tests (e.g., geography). Likewise a study by McNeil and Valenzuela (2000) investigating the impact of the secondary school high-stakes Texas Assessment of Academic Skills (TAAS) identified widespread test-preparation activities at the expense of the official curriculum. They noted that these practices were more widespread in those schools where administrator pay was tied to test scores and where test scores had been historically low.

High-stakes assessment has been shown to negatively impact on teachers' pedagogy. In Reay and Wiliam's study (1999), the researchers observed that "it was the emphasis on more individualised, competitive ways of working, which were increasingly displacing the mutually supportive, collaborative group work to which the children were accustomed …that caused the most disquiet" (p. 351). Similarly, Johnson and McClune (2000), investigating perceptions of pedagogical practices during preparation for the '11-plus' national science examination in Northern Ireland[3], found that, due to the perceived nature of the selection tests, teachers felt compelled to teach through highly structured activities and through the transmission of information. The researchers observed that the teachers' transmissive stance was often in tension to the students reported preference for hands-on experiences, problem-solving tasks, and a willingness to take risks and be creative. This discriminated and demoralised those students who preferred to learn in other ways.

In Pollard et al.'s study (2000), the researchers concluded that when teacher accountability unduly emphasises students' achievement (rather than effective teaching), teachers spend a lot of time preparing students for assessments at the expense of quality learning. Similarly, in a case study of a senior mathematics teacher Hosking and Shield (2001) noted that the teacher's pedagogical practices revolved around the process of students developing skill mastery and then reproducing it for assessment. The mathematics teacher acknowledged that "the students did not develop an understanding of the topic but rather rote learnt the techniques needed to pass" (p. 293). As such, the students in the research classroom were confined to passive

[3] The results of these national tests have traditionally been used for the highly competitive selection into grammar schools.

involvement in both the learning and assessment process by the teacher's transmissive pedagogical stance. The teacher, however, perceived that his actions were in the best interest of the students.

Teachers' perceptions of high-stakes assessment impact on teachers' feedback practices. In the study by Hosking and Shield (2001), the research teacher's focus on marks, correct solutions and examination strategies determined his feedback practices. Feedback from formal assessments consisted of ticks and crosses indicating correct and incorrect solutions. The mathematics teacher justified his feedback practices arguing that students were only interested in whether they got a question right or wrong and hence "[m]any of them would probably just look at their marks and not be too worried about how they lost marks" (p. 294). The research teacher conceded that the more able students might want more feedback but perceived that the less able students would just reject it, so there was "no point offering anything more to them" (p. 294). The teacher saw the purpose of returning assessed work as an opportunity for students to check the marking and try to "get more marks" (p. 295). Oral whole-class feedback consisted of the provision of correct solutions with public recognition for those who received high results. According to the mathematics teacher, once the assessment had been sat and the marks returned to the student, the assessment process was complete.

The tensions appear to be an inherent challenge of high-stakes assessment systems. As Harlen (2006) opined:

> It is natural for students and teachers to aim for high performance, but when this is measured by external tests and when the results are accompanied by penalties for low performance, the aim becomes to perform well in the tests and this is often not the same as to learn well (p. 80).

High stakes assessment does not, however, inevitably lead to negative impacts on teaching and learning. Where close alignment exists between curriculum and assessment systems, high-stakes assessment can focus teachers' and students' attentions on key goals of learning (Biggs, 1996, 1998; Birmingham, 2001; Clune, 2001; Porter & Smithson, 2001). Preece and Skinner (1999) investigated the assessment of National Curriculum science in England and Wales at key stage 3

(14-year-old pupils). Despite some expressed concerns that an overcrowded curriculum encouraged transmissive modes of teaching with less depth of coverage of some topics, 60% of the teachers reported that a structured curriculum and assessment system based on clear objectives gave a much better focus to their teaching, resulting in it being more organised and systematic.

A number of studies have indicated that high-stakes assessment systems based on SBA principles can exhibit close alignment between curriculum and assessment, resulting in a positive impact on teachers' pedagogy and students' learning. In a case study investigation of nine state education systems in the United States, Clune (2001) concluded that those systems exhibiting high levels of alignment produced "widespread and substantial gains in the quality of teaching and learning for all students" (p. 14). Similarly, Birmingham (2001), researching the introduction of New York State mathematics Standards in secondary schools, identified that a stronger alignment between assessment and teaching resulted in higher success rates and allowed for more student-focused teaching. A report from The Education Commission of the States (2002) found that the clear goals for learning and teaching afforded by SBA are beneficial in the shaping of the performance of teachers, learners and school systems. In practice, however, alignment depends on teachers' ability to understand expectations set out in the Standards, the ability to obtain resources and expertise to help students meet those Standards, and teacher knowledge (Linn & Herman, 1997).

Wilson and Floden (2001), investigating the introduction of SBA in 23 school districts in eight states in the United States, found "little evidence to support critics' fears that Standards would lead to an over or under emphasis on some learning outcomes" (p. 195), often referred to as 'teaching to the test'. Although many teachers reported that tests affected instruction, classroom observations confirmed that assessment requirements were not predominantly driving the enacted curriculum. Teachers in one state said that they were aware of the debate about teaching to the test and, while some rejected the claim, others felt "it was a non-issue because the test helped teachers to focus and raise expectations for all students" (p. 198).

A range of studies report the improvement in professional dialogue between teachers as a significant positive consequence of Standards-Based reform (e.g., Clune, 2001; James, 2000; OECD, 2005; Wilson & Floden 2001). For example, teachers in Wilson and Floden's study reported that conversations with peers were "the most productive and meaningful professional development in recent times" (p. 197). The researchers reported that the clarity of assessment Standards created a catalyst for teachers using their professional judgement to create "a more coherent teaching practice" (p. 214).

The New Zealand Context.

There is a limited research literature base on high-stakes assessment systems in New Zealand, particularly in the area of secondary school mathematics. As reported earlier, Pfannkuch (2001) investigated teachers' perceptions and practices relating to assessment of school mathematics. In addition to the findings already reported, Pfannkuch investigated perceptions relating to the high-stakes nature of secondary school assessment practices. The teachers interviewed perceived that schools were accountable to the Education Review Office, the MoE, politicians, and parents, based on their national assessment results. She concluded that this perception of accountability determined the teachers' assessment and pedagogical practices, subverting the original aims of the mathematics curriculum to provide an integrated mathematical experience encouraging students' higher order thinking skills.

Loretz (2002) investigated the changes in mathematics assessment tasks and teachers' perceptions of their practice due to the introduction of NCEA. Although this research was conducted during the initial stages of implementation, and only covered the external examinations, it does provide some insight into teachers' general perceptions of the impact of assessment on their pedagogy. Loretz reported that teachers felt the introduction of NCEA had constrained their teaching practice. Teachers reported that prior to NCEA they taught mathematics with an holistic approach, integrating topics as appropriate. They perceived that this approach mirrored the types of questions students would face in the end-of-year School Certificate examination. In contrast, subsequent to the introduction of NCEA, teachers perceived that their teaching was more atomised and closely adhered to the assessment criteria specified in the

Standards. The majority of teachers were unhappy with this change to their practice, which they saw as being imposed on them by the changes in the assessment system.

In order to contextualise the teachers' perceptions, Loretz (2002) went on to compare the nature of the mathematics questions asked in the School Certificate and Level 1 NCEA examinations. He identified that within NCEA, assessment tasks covered a broader range of thinking skills including the higher level skills notably absent in the old system. Additionally, the contexts of questions were more authentic within NCEA. In contrast to the teachers' perceptions, Loretz noted that topic integration in the assessments was rare in both systems.

In contrast to the constricting effect of assessment on pedagogy reported in Pfannkuch's (2001) study, two assessment research studies in secondary schools by Bushnell (drama, 1992) and Eng (English, 1992) found that the use of grade-related assessment criteria had a number of positive effects on pedagogy and learning. Teachers in both of these studies reported an increased awareness of their teaching practice, more effective reporting to students and parents about progress and achievement, and a positive impact on their own teaching at other levels of the school. Additionally, teachers in the study by Eng reported a shift towards a more student-centred approach to their teaching.

More recently, a report by Barrington (2004) looked at the perceived benefits of the introduction of NCEA. A discussion group with principals revealed that with the introduction of SBA, teachers were more focussed on higher order thinking skills and more concerned with teaching for learning rather than teaching for high-stakes. The principals also reported that the introduction of NCEA had resulted in an increased use of formative assessment and greater collaboration between students and teachers. It was generally felt that the new SBA system provided schools with a set of enhanced tools and increased flexibility to meet the learning needs of their students. Similarly, teachers in a study by Meyer, McClure, Walkley, McKenzie, and Weir (2006) reported that the internal assessment under SBA had "sharpened [teachers'] awareness of the effectiveness of their teaching and focused their teaching on those issues that were seen as most important and/or relating to the various [A]chievement

[S]tandards" (p. 5). (Additional findings from Meyer et al.'s research into the impact of NCEA on student motivation are reported in section 2.6.3.)

The New Zealand Council for Educational Research's (NZCER) 'Learning Curves' project has documented changes in the subject and assessment choices offered to senior students in six medium-sized New Zealand secondary schools during the implementation of NCEA. A number of reports have been generated from this research project (e.g., Hipkins, 2004; Hipkins, Vaughan, Beals, & Ferral, 2004; Hipkins & Vaughan, 2005). In the most recent report, Hipkins et al. (2007) identified a number of significant findings associated with the Learning Curves project.

Firstly, critics of NCEA have often argued that it creates 'credit seekers'; students whose engagement in learning is determined by whether credits are attached to the learning activity. Access to a wider range of assessment Standards under the new assessment regime has resulted in over-assessment of students, with "content acquisition" (Hipkins et al., 2007, p. 6) becoming the main outcome of learning in the senior secondary school. Hipkins et al. (2007) noted that "teachers have always used the prospect of gaining qualifications as a means of creating a sense of purpose for learning" (p. 9). With the change to increased opportunities to secure 'credits' towards qualifications during the year, many teachers quickly came to believe that they would not be able to interest students in any learning if there were no credits attached (Hipkins, et al., 2004).

Results from the Learning Curves project, however, challenge the belief that students are solely extrinsically motivated (Hipkins et al., 2007). Although students initially cited gaining credits as an extrinsic motivational factor for learning, a deeper investigation of their views revealed that students are also intrinsically motivated. The researchers noted, however, that the pressure of over-assessment identified by the students may well prevent them from "seeing the other types of engagement that are possible, if they are too busy trying to keep on top of constant assessment pressures" (Hipkins et al., 2007, p. 17).

> [I]t hardly seems fair to complain that the NCEA per se has made students into 'credit seekers' rather than fostering a genuine love of learning. Many students, from across the ability spectrum, have become happy collectors of credits … but this is what has frequently been encouraged by their teachers. (Hipkins et al., 2007, p. 9)

Secondly, Hipkins et al. (2007) reported that mathematics, English and science teachers were very positive about the flexibility offered by NCEA to provide alternative courses designed to meet the learning needs of an increasingly diverse group of students. The teachers identified these new courses as important means of motivating low ability students (Hipkins et al., 2004). Similar results were reported by Meyer et al. (2006), who concluded that teachers perceive that NCEA was more motivating for low ability students.

Lastly, "student engagement in learning is not solely attributable to their engagement in standards-based assessment" (Hipkins et al., 2007, p. 21). Their longitudinal data supports the claim that students who, prior to their involvement with NCEA, had good levels of engagement and motivation, and had gauged their learning progress through intrinsic factors, were the most likely to show good levels of engagement with school in Y11. Students with high levels of school engagement showed good engagement with NCEA, and had good results. Furthermore, students who rated their current learning environment positively showed higher engagement in learning. Accordingly, Hipkins et al. argue that:

> If approaches to learning and the learning environment itself are more important to current engagement in learning than is engagement with assessment, there is an opportunity to challenge the dominant teacher view that the extrinsic reward of assessment credits is the key way to give purpose and meaning to learning. (p. 21)

As such Hipkins et al. remain cautiously optimistic that the potential benefits of the new system will be realised in time, if deeply entrenched traditional attitudes to what counts as indications of quality learning can be overcome" (p. 26).

2.5 Assessment as a continuum.

A number of researchers have argued that the summative and formative roles of assessment are not always compatible and institutional demands mean that the summative role often takes precedence (Boud, 1995; Crooks, 1988; Ritter, 2000).

In-class assessments by teachers often emulate summative external tests in the assumption that this represents good assessment practice (Crooks, 1988; Harlen & Deakin Crick, 2003; Pollard et al., 2000). However, Black and Wiliam (1998a) claim that over-reliance on summative type assessment teaches the lower-achieving students that they lack ability, resulting in students losing confidence in their ability to learn and therefore de-motivating them.

Increasingly, research is providing evidence that the formative and summative roles of assessment need not be mutually exclusive (Biggs, 1998; Black, 2000; Black & Wiliam, 1998a; Brookhart, 2004; Harlen, 2006; Sebatane, 1998). We know that successful students often learn to use feedback from assessment events, designed to provide summative information, for their own formative purposes, by reviewing work and engaging in self-assessment as a regular ongoing process (Anthony, 1994; Tanner & Jones, 2003). For example, Anthony's (1994) study of Year 12 students learning strategies identified that successful students are simultaneously able to monitor and control their learning using strong metacognitive strategies. Similarly, in interviews with secondary school students, Brookhart (2001) found that successful students had developed a "natural searching for the formative" (p. 166). This observation is in accord with Sebatane (1998), who argued that any theory about formative assessment may need to include some references to summative assessment, and vice versa, in order to describe more adequately the cyclical process successful students apparently construct for themselves. As Biggs (1998) opined:

> There is a powerful interaction between FA [formative assessment] and SA [summative assessment] that could usefully be incorporated in an overall synthesis so that both … are conceptualised within the same framework. (p. 106)

Wiliam and Black (1996) suggest that, rather than a dichotomy, summative and formative assessments can be viewed as the ends of a continuum along which all

assessment can be located. Similarly, the National Curriculum Task Group on Assessment and Testing in England (DES/WO, 1988) took the view that a single assessment system could serve both summative and formative functions, provided the formative function was the foundation of the system. This notion is based on the premise that it is easier to ascribe a summative grade to an assessment task designed primarily as a formative tool, than visa versa (Harlen, 2006).

In order to conceptualise a single high-stakes assessment system that can serve both formative and summative roles, it is important to make a distinction between the gathering of assessment evidence and the interpretation of that evidence (Harlen & James, 1997). The use of the terms 'summative' and 'formative' assessment can give the impression that these are different types of assessment events (Harlen, 2006). Assessment gathers information about current levels of achievement for a student. It is the interpretation of that information, and the subsequent action taken, that moves the assessment task into the formative realm. As long as a distinction is established between the gathering and the interpretation of assessment evidence, the formative function of assessment tasks need not be incompatible with high-stakes summative assessment (Wiliam & Black, 1996).

2.6 A theoretical way forward

The following section draws upon a theoretical framework for formative assessment proposed by Sadler to examine the extent to which an SBA system can satisfy both formative and summative purposes of assessment. Sadler (1989) proposed three key learner-based components to formative assessment, namely:

> [T]he learner has to (a) possess a concept of the *standard* (or goal, or reference level) being aimed for, (b) compare the *actual* (or current) *level of performance* with the standard, and (c) engage in appropriate *action* which leads to some closure of the gap. (p. 121, emphasis in original)

Notably, Sadler (1989) stated these components from the students' perspective. Theorising from this perspective reflects the position adopted in the current project that ultimately the responsibility for taking any necessary actions to enhance future learning rests with the student. There is, however, an inherent assumption in this

position: that students' commitment to learning is likely to be strengthened when they take more responsibility for monitoring their own strengths and weaknesses, and for devising consequent strategies for improving their learning outcomes. Similar theoretical perspectives have also been suggested by Gipps (1994) and Black and Wiliam (1998a).

The following three sub-sections discuss the research literature pertinent to Sadler's (1989) theoretical position and are organised around the three key components of formative assessment identified above. It is recognised that these three components intersect within any learning environment and it is difficult to discuss one without reference to the others. Many of the research studies reported cross the boundaries between the three sections detailed below.

2.6.1 Understanding of assessment criteria.

Sadler (1989) theorised that a shared understanding of the assessment criteria between teacher and student is important if formative assessment practices are to be effective. The provision of criteria alone does not, however, ensure a common shared understanding. Sadler argued that it is necessary for teachers to share some of the responsibility for assessment with the students through a process of students becoming 'apprentices in judgement'.

> [P]roviding guided but direct and authentic evaluative experience for students enables them to develop their evaluative knowledge, thereby bringing them within the guild of people who are able to determine quality using multiple criteria. It also enables transfer of some of the responsibility for making decisions from teacher to learner. In this way, students are gradually exposed to the full set of criteria and the rules for using them and so build up a body of evaluative knowledge. It also makes them aware of the difficulties which even teachers face of making such assessments; they become insiders rather than consumers. (Sadler, 1989, p. 135)

A key premise of Sadler's (1989) argument is that students develop a concept of the goal or learning target which originally is the teacher's, but which ideally the student will internalise through the processes of self- and peer-assessment. Students create deeper understandings of the explicit assessment criteria as they think about and apply criteria to their own work and the work of others (Gipps, 1994; Sadler, 1989). Self-

and peer-assessment also encourage metacognition and autonomy as students act in the role of assessor and judge of quality, thereby contributing to a shift in the conceptions of the roles of the teacher and student in the learning process (Black et al., 2006; Brookhart, 2001).

The positive effects of self-assessment are not limited to the learning that occurs at school. Crooks (2001) observed that self-monitoring is a key component of the work of all professionals and argued that:

> [I]f we want our students to become professional learners and professionals in their fields we should actively promote self-assessment. If students are asked and encouraged to critically examine and comment on their own work, assessment can become more dialogue than monologue, and can contribute powerfully to the educational development of students. (p. 5)

Research within the mathematics domain has highlighted the effectiveness of student self-assessment in raising achievement. Fontana and Fernandes (1994) conducted a study involving 25 mathematics teachers and over 350 students (aged 8–14 years), focused on regular, mainly daily, self-assessment by the students. Comparisons revealed that the experimental group had twice the mean gain in achievement of the control group. Similarly, Schunk's (1996) study exploring self regulatory processes of 9–10 year old mathematics students identified that achievement scores for students engaged in self-assessment were significantly higher than those who were not involved in self-assessment.

A number of research studies in other subject domains further illustrate the effectiveness of self-assessment in raising student achievement levels. A study on oral reading rates of elementary school students by McCurdy and Shapiro (1992) investigated formative feedback provided by one of three methods. Over a period of nine weeks, feedback was provided by the teacher or through peer- or self-assessment, and comparisons were made with a control group who received no formative feedback. Although all three experimental groups experienced higher achievement gains than the control group, the largest gains between pre- and post-test scores were achieved by the group involved in self-assessment.

Similar findings were repeated in a study by Frederiksen and White (1997). The researchers reported that 30 middle school science students trained in reflective self-assessment understood the assessment process and performed better than the control group who received no such training. Peer-assessment also has also been shown to have a positive impact on students' achievement. Gibbs (1999), investigating the introduction of peer-assessment at the university level engineering and pharmacy course, found that it resulted in significant achievement gains for students.

A study by McDonald and Boud (2003) sought to address the paucity of studies that examined the influence of self-assessment training on externally judged student achievement using conventional examinations. This project involved a large-scale intervention in the West Indies focusing on all curriculum areas in the final year of high school. Teachers from 10 secondary schools were trained in how to develop self-assessment skills in their students. The schools involved covered a range of high, middle and low achieving academic schools as measured against national examinations. Post-intervention, the achievement levels of the 256 students in the experimental group were significantly higher than the 259 students in the control group, and this advantage persisted over a range of curriculum areas. Additionally, the students were surveyed about their reaction to the self-assessment training. They indicated that the training had been beneficial to them and not an additional burden in their examinations year. Ninety-five percent of respondents commented that they appreciated being able to judge their level of preparedness prior to the end of year examinations with sufficient time to address any shortcomings. Furthermore, students commented that not only was the self-assessment training relevant for preparation for external examinations, but it also had wider impacts on their perceptions of future learning opportunities and possible careers.

A range of studies have demonstrated the iterative nature of self- and peer-assessment. Klenowski (1995) used a case study approach to examine self-assessment practices in an Australian high school and an English 'future education' college; both chosen for their student centred teaching and learning approach. She found that students who participated in self-assessment practices became more interested in the assessment criteria and feedback provided, rather than in the grade given to their assessed work. This, in turn, strengthened their ability to accurately self-assess. In reporting on the

King's-Medway-Oxfordshire Formative Assessment Programme (KMOFAP), Black, Harrison, Lee, Marshal, and Wiliam, (2003) found that peer-assessment turned out to be an important stimulus for self-assessment. The authors reported that students appeared to find it easier to make sense of their own work if they examined other students' work alongside their own.

Ross, Radnor, Mitchell, and Bierton (1993) sought to bring students into the assessment process as reflective practitioners by encouraging students to reflect on their work and to articulate their reflections through 'assessment conversations'. Their study found that students "are capable of rich and sophisticated responses to and understandings of their own work ... in collaboration with their conversation partner" (p. 161). This in turn led to improved self-assessment practices.

In a study of 50 students in high school English and anatomy classes, Brookhart (2001) found that high achieving students engaged in self-assessment as a regular, ongoing process and actively tried to fit new information about their learning into their careers as students. High achieving students did not make "neat distinctions between formative and summative assessment, but used assessment in a variety of integrated ways. This [was] consistent with their outlook on learning, which they reported viewing as one of their important life processes" (p. 153).

Many of the research studies into the effectiveness of self-assessment techniques have investigated situations where students mark their own responses to preset questions. Carroll (1994) found improvements in secondary mathematics students' achievement when they marked teacher generated solutions to algebra problems using the supplied assessment schedules. Similarly, Koch and Shulamith (1991) investigated a group of college physics students who were taught to generate their own assessment questions and marking criteria. These students achieved better learning gains than those who used the questions supplied by the teacher for self-assessment. Even greater gains were demonstrated when the questions generated by student were shared, answered, and critiqued amongst peers.

The use of student generated exemplars also benefits student understanding of assessment criteria. Orsmond, Merry, and Reiling (2002) investigated first-year

undergraduate biology students as they designed an exemplar poster for a content unit. The students were subsequently taught the principles of self- and peer-assessment and applied these to their own posters, as well as the other posters produced by the class. In addition to benefiting students' understanding of the assessment criteria, the researchers concluded that the use of exemplars had provided a focus for meaningful formative assessment discussion. Notably, students were able to make "more objective judgement as a result of peer-assessment compared to self-assessment" (p. 309).

In addition to helping students' learning, self- and peer-assessment provides accurate assessment information, highlighting their validity as alternative assessment strategies. In a study investigating the reliabilities of self- and peer-assessment in college biology students, Stefani (1994) found correlations with teachers' assessments of 0.71 for self- and 0.89 for peer-assessments. All students reported that the use of self- and peer- assessment had made them think more, and 85% said that it made them learn more. Similar results had been found by Hughes and Large (1993), who reported a correlation for peer-assessment of 0.83 between the mean ratings of a group of final year undergraduate pharmacology students and those of a group of staff.

Despite evidence that a focus on self- and peer-assessment is beneficial to students' learning, research has identified that it is not common practice, even amongst those teachers who take assessment seriously (Black et al., 2003). Daws and Singh (1996) identified that only about a third of the UK secondary science teachers in their study involved pupils directly in their own assessment. Tiknaz and Sutton (2006), investigating the role of key stage 3 geography assessment tasks in promoting formative assessment, found that despite recognising their value, teachers in general "did not make extensive use of self-assessment and their use of peer assessment was even rarer" (p. 337). The researchers report that the majority of teachers were not prepared to "share their power with pupils as assessors of themselves and their peers" (p.337). According to Sadler (1989), students are typically given few opportunities to make appropriate qualitative judgements. Sadler questioned whether this suggests an underlying assumption that only teachers have the skills and expertise to evaluate student work, and that these skills are not transferable to students.

Recent indications suggest, however, that the use of self- and peer- assessment is becoming a more established part of some classrooms. A report from the OECD (2005), summarising case studies from eight countries (including New Zealand), identified that peer- and self-assessment was a frequent practice in the secondary schools visited. Similarly, the KMOFAP (Black et al., 2003), investigating the development of secondary mathematics and science teachers' strategies for implementing formative assessment, found that almost all the teachers in their study mentioned some form of self-assessment in their plans.

In the New Zealand Education Review Office's (2004) evaluative report of NCEA, both teachers and students reported that students were now receiving more information about assessment criteria and expectations than previously, which helped students to evaluate their own learning. The report identified two major approaches to self-assessment that were being utilised. Firstly, the use of tracking sheets so students could monitor their progress toward achieving components of an assessment Standard. Secondly, increasing students' self-assessment skills by encouraging them to reflect on their performance, set goals and discuss these with peers and teachers.

Despite apparent improvements in practice, the OECD (2005) report noted that the use of self- and peer-assessment remains "one of the most challenging aspects of teaching in the formative assessment mode" (p. 65). The report identified that focusing students' attention on specifics relating to the criteria for a high quality piece of work is a key challenge for implementing self-assessment practices. The report also noted that students need to be coached and provided with opportunities to practise their skills if they are to provide useful assessments of either their own work, or the work of their peers.

2.6.2 Feedback.

The second condition theorised by Sadler as necessary for effective formative assessment is that students must be able to "compare the *actual* (or current) *level of performance* with the standard" (p. 121, emphasis in original). Implicit in this condition is the provision of feedback on current student achievement. Findings of early meta-analyses identified the quality of feedback provided from assessment tasks

is a key feature in any procedure for effective formative assessment (Black &Wiliam, 1998a; Crooks, 1988; Natriello, 1987).

In a recent review of meta-analyses, Hattie and Timperley (2007) highlighted feedback as a powerful moderator that enhances learning and achievement with an overall effect size of 0.79, based on almost 7000 effect sizes. Hattie places this figure in context by reporting that "it fell in the top 5 to 10 highest influences on achievement in [his] (1999) synthesis, along with direct instruction (0.93), reciprocal teaching (0.86) [and] students' prior cognitive ability (0.71)" (p. 83). In another significant review of the impact of feedback on learning, Kluger and DeNisi (1996) concluded that overall, the use of feedback has a positive effect on student performance with an average effect size of 0.4, the equivalent of raising the average student to the 65th percentile. While Hattie and Timperley, and Kluger and De Nisi acknowledge that the impact of feedback can be either positive or negative depending upon specific features of the feedback, both sets of research conclude that feedback is an important component of any formative assessment system.

Early behaviourist conceptions of feedback identified it with 'knowledge of results', whereby the reinforcement of correct knowledge and skills was theorised to lead to motivation to improve the given grade (Sadler, 1998). Alternative conceptions of feedback centred on ego-oriented or 'praise for effort' feedback, assuming that such praise would promote a positive self-image, leading to higher motivation and achievement, and perhaps even a love of learning (Sadler, 1998). In recent years, the assumptions upon which these traditional conceptions are based have been questioned. Increasingly, research has focused on the quality of the feedback and, in particular, how it relates to the task in hand.

Task- and ego-involving feedback

In a major review on the impact of feedback on learning, Kluger and DeNisi (1996) noted a high variability in reported effect sizes. In the studies they reviewed, approximately 40% of the feedback interventions had a negative effect on performance. In those cases where feedback was not helpful, the feedback was merely a judgement or grading with no indication of how to improve. Kluger and

DeNisi proposed a Feedback Intervention Theory, arguing that feedback changes the locus of attention amongst three "general and hierarchically organised levels of control: task-learning, task-motivation and meta-tasks (including self-related) processes" (p. 254). When faced with negative cues from feedback, students invoke a mid level task-motivation strategy and try to utilise universal strategies that work on most tasks (for example expending more effort, persistence, and focusing attention on the task). If this is not effective, students may migrate to more effective task-learning activities, searching for a task specific strategy. In some cases, negative feedback may cause student to move towards less effective meta-task activities concentrating on the 'self', and away from the details of the task. Kluger and DeNisi concluded that feedback that provides cues to support learning—highlighting feedback at the task-learning level—and avoids cues that direct attention to the meta-task level, can have a significantly positive effect on achievement (effect size greater than 1).

Similar conclusions were drawn by Hattie and Timperley (2007). In their review they noted that: "Those studies showing the highest effect sizes involved students receiving information feedback about a task and how to do it more effectively. Lower effect sizes were related to praise, rewards and punishment" (p. 84). This finding is in accord with the 'feedback typology' suggested by Tunstall and Gipps (1996) who noted that although ego-involving feedback that draws attention to the self can initially increase students' interest towards a task, it has little if any effect on achievement.

A landmark study on the effect of task- and ego-involving feedback on interest and performance was conducted by Butler (1988). The 132 fifth and sixth grade Israeli participants were divided into three groups and given either ego-involving numerical grades, task-involving individual comments, or both. Butler found that students given comments only scored significantly higher in the follow-up testing than students offered other forms of feedback, and maintained their high levels of task-involved orientation and interest. Furthermore, the comment-only group improved their subsequent performances primarily on those components of the initial task targeted by the comments.

Students who were given both grades and tailored comments concentrated on the grade, at the expense of the comments, and showed less relative improvement than those students who received only tailored comments. Butler observed that "the maintenance of ego-involved motivation seems to depend on the continued availability of opportunities for social comparison, so that both effort and interest wane when these are withdrawn" (p. 13). Butler also found that the inclusion of ego-involving feedback, such as normative grades, either by themselves or in addition to comments, caused a shift from an initial task-involved orientation to an ego-involved orientation in the follow up learning sessions. Butler suggested that task-involved intrinsically-motivated students actively incorporate cues which help with the formation of internal standards for guiding and assessing performance. She concluded that interest and performance of students is best achieved by promoting task-involvement through giving task-related, non-ego-involving feedback (see also Brookhart & DeVoge, 1999; Higgins, Hartley, & Skelton, 2002).

Recognition of the value of task-oriented comment-only feedback can also be found in the results of the KMOFAP project (Black et al., 2003; Black & Wiliam, 2006). The KMOFAP action research project looked at mathematics and science teachers' strategies for implementing formative assessment in their classrooms. At the commencement of the study nearly half of the teachers specifically mentioned wanting to make more use of comment-only marking. Initial concerns, expressed by some teachers, that reduced use of marks and grades would not be well received by the school, or the parents, proved unfounded. Experience showed that the provision of task-oriented comment only feedback on written work stimulated teachers to improve their scaffolding of students' learning, and created a focus by students, and their parents, on the learning issues rather than on trying to interpret a grade. Likewise, parents from several of the case study schools in the OECD report (2005) endorsed task-oriented comment-only marking claiming that "they had a better idea of what their children were doing and how they might be able to help them with their school work" (p. 59). The OECD report did note, however, that 'dropping marks' can be problematic in cases where parents and students focused on social comparisons between students rather than focusing on students' individual progress.

Seemingly in contrast with the research cited above, Smith and Gorard (2005) conducted a small intervention study examining the value of comment-only marking with 104 Year 7 students in a Welsh comprehensive school. In this study, a group of 26 students were provided with comments-only on their assessed work and compared with three other groups, each of 26 students, who received grades with minimal comments. The researchers concluded that, based on the results of their study, the approach of providing comment-only marking was ineffective in raising student achievement and "somewhat unpopular with the students" (p. 37). Significantly though, the researchers noted that most of the comments provided to the experimental group "appeared to focus upon enhancing self-esteem or self-image rather than on what needed to be done to improve and how students might go about making any improvements" (p. 36). These findings are partially explained by the observation cited earlier that ego-involving feedback has little effect on student achievement (Tunstall and Gipps, 1996).

Collectively these studies demonstrate that feedback is most effective when it is task-oriented, providing scaffolded responses to student errors, rather than simply indicating whether an answer is right or wrong (Black & Wiliam, 1998a; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Sadler, 1989; Tunstall & Gipps, 1996). Feedback should provide an indication of progress towards desired learning outcomes, and should encourage the view that mistakes are a part of learning (Kluger & DeNisi, 1996). "Such emphasis on personal progress enhances self-efficacy, encourages effort attributions, and reduces attention to social comparison" (Crooks, 1988, p. 55). In mathematics, a domain previously dominated by the perception that mathematical ability is 'fixed', effective feedback should be designed to lead students to believe that their mathematical ability is incremental (Wiliam, 1999). Such an approach has been found to encourage deep rather than surface learning (Black & Wiliam, 1998a).

Recent developments

Hattie and Timperley (2007) have recently expanded on the notion of task- and ego-oriented feedback by proposing a model that distinguishes between four main levels of feedback: the task, the process, the regulatory and the self levels. Feedback at the task level identifies whether work is correct or incorrect, and "may include directions

to acquire more, different or correct information" (p. 90). Feedback at the process-level is concerned with the processing of information or necessary actions to complete the task. For example, "[y]ou need to edit this piece of writing by attending to the descriptors you have used so the reader is able to understand the nuances of your meaning" (p. 90). Feedback at the regulatory level is designed to improve students' ability to assess their current product, or give them confidence to engage further on a task. Regulatory feedback is designed to influence students' "self-efficacy, self-regulatory proficiencies and self-beliefs about [themselves] as learners" (p. 90). For example, "[y]ou already know the key feature of the opening of an argument. Check to see whether you have incorporated them in your first paragraph" (p. 90). The fourth level suggested by Hattie and Timperley concerns feedback directed to the 'self', which they argue is too often unrelated to the students' performance on a task and is rarely effective. This definition of 'self' feedback is consistent with what is more commonly referred to as ego-oriented feedback.

Students' perceptions of feedback

Several studies report that older students value feedback that encourages them to engage with the subject material in deep ways (Higgins et al., 2002). Investigating perceptions of feedback for students in higher education, Higgins et al. found that 97% of students indicated that they usually 'read' the written feedback they receive and 82% of the students claimed to 'pay close attention' to the feedback. Many students went on to argue that they expected feedback as recognition of the effort they have made in producing the assessed work. Similarly, Hyland's (2000) study of university history students found that 90% of the participants believed that feedback could help them identify their strengths and weaknesses, engender a sense of achievement, and raise their marks in future work. Hyland also commented that students never seem to lose faith in the potential value of feedback, despite any problems they may encounter when attempting to use it. In a recent study involving secondary school students in New Zealand, Meyer et al. (2006) reported that receiving feedback on assessed work was appreciated by students of all ability levels.

Students' perceptions of preferred feedback are affected by their learning goals (Cowie, 2005). Investigating Y7 to Y10 New Zealand science students, Cowie found

that those students with learning goals viewed formative assessment as a joint responsibility between the teacher and the pupil. As such, they "expressed a preference for teacher feedback in the form of suggestions because these maintained an active role for them in making sense of ideas" (p. 137). In contrast, performance oriented students viewed "assessment as a teacher's sole responsibility" and "preferred feedback on how to complete tasks" (p. 137).

2.6.3 Engagement with feedback

The third key component of formative assessment proposed by Sadler (1989) is that students should "engage in appropriate action which leads to some closure of the gap" (p. 121). Feedback, however detailed, will not lead to improvement unless students are motivated and empowered to engage with it, and it cannot be assumed that students will independently act on the feedback offered (Sadler, 1998, Tunstall & Gipps, 1996). Students often fail to recognise the formative nature of feedback, and how it applies to their work (Black & Wiliam, 1998a; Sadler, 1989).

In addition to finding out about their knowledge and skills, students see assessment as "an event that reveals something intrinsic about them as individuals" *(*Reay & Wiliam, 1999, p. 343). As such, students' past assessment experiences, self-efficacy and motivational states all affect their engagement in assessment practices (Crooks, 1988). The following sections examine the research literature from a number of interrelated components that influence students' propensity to effectively engage with formative feedback.

Motivation.

Assessment and motivation for learning are integrally related constructs. Major reviews of classroom assessment practices all discuss assessment in terms of its relationship to student motivation (Black & Wiliam, 1998a; Crooks, 1988; Natriello, 1987). Students' motivational states will affect the way they engage with assessment processes, as well as being affected by that engagement. As such, motivation for learning is not only seen as a necessary input into education, it is also seen as an important outcome of education (Harlen & Deakin Crick, 2003).

Motivation is a multifaceted construct (Harlen & Deakin Crick, 2003). A distinction can be made between intrinsic motivation—where students find interest and satisfaction in what they learn and in the learning process itself, and extrinsic motivation—where students engage in learning because it is a means to an end, but have limited interest in the content that is learned (Harlen & Deakin Crick, 2003). Furthermore, any discussion of motivation should consider the attributions people make for their success and failure, self-efficacy, and the impact of goal orientation as interconnected components of motivation for learning (Kellaghan, Madaus, & Raczek, 1996).

Intrinsic, rather than extrinsic, motivation has been identified as a key component for life long learning (Black & Wiliam, 1998a; Harlen & Deakin Crick, 2003). Intrinsic motivation encourages levels of engagement that lead to development of conceptual understanding and higher level thinking skills (Kellaghan et al., 1996). In contrast, extrinsic motivation can lead to performance goals and shallow learning (Crooks, 1988; Kellaghan et al., 1996). Benmansour's (1999) study of secondary school mathematics students found that students with strong extrinsic motivation had high levels of test anxiety and made greater use of passive rather than active learning strategies. In contrast, students with a stronger intrinsic motivation had lower test anxiety, higher self-efficacy, and used a wider range of active learning strategies. However, Hidi and Harackiewicz (2000) caution not to oversimplify the inter-connectedness of these two components of motivation for learning. Hidi and Harackiewicz contend that "a combination of intrinsic rewards inherent in interesting activities and external rewards, particularly those that provide performance feedback, may be required to maintain individuals' engagement across complex and often difficult—perhaps painful—periods of learning" (p. 159).

Weiner (1979) proposed an attributional model for motivation. In this model, Weiner argued that there are four dominant attributions people make for their successes and failures: their natural ability, the amount of effort they put in to the task, an element of 'luck' in the questions they are asked, and the influence of other people such as the teacher. The attributions people make affect their emotional and motivational reaction to their success and failure. Central to Weiner's theory are the sub-dimension of 'locus of control' and 'stability of cause'. Locus of control refers to

students' attribution of their success or failure as either internal or external. Students with an internal locus of control will attribute their success to factors they have some element of control over, such as their ability, or the amount of effort they put in to preparing for assessment events. In contrast, students with an external locus of control will attribute their success to external factors, such as the teacher or 'they just got a little lucky with the questions'. Stability of cause encapsulates the extent to which students perceive certain causes, such as ability or effort, to be "stable" (p. 6) and therefore hard to change, or "unstable" (p. 6) and easier for the student to have influence over.

In particular, students' perception of their ability as either stable or unstable will affect the amount of effort they are willing to expend on a learning task (Dweck & Leggett, 1988). Students who see their ability as stable—and therefore cannot be changed—will see success at learning tasks as a confirmation of their ability, but failure as a challenge to their view of their ability. Accordingly, the risk of failure must be minimised and so students will avoid challenges when their confidence of success is low. Students who perceive their ability as unstable—as an incremental set of skills that can be expanded—will persist in the face of difficulty and will not be deterred by failures on learning activities. Students who see their ability as external and stable may fail to appreciate that assessment feedback can provide them with a way of improving their achievement, since it is beyond their power to do so. Students who see their ability as internal and unstable are most likely to respond to appropriate feedback by increasing their effort (Dweck & Leggett, 1988).

Recent research into the impact of NCEA on student motivation was undertaken by Meyer and colleagues (2006). This research distinguished between two main student motivations identified as 'doing my best' and 'doing just enough'. Doing my best can be characterised as:

> [S]tudents valuing work that leads to Merit or Excellence (e.g., "I expect to get [excellence] or at least Merit when I try"), getting a good education (e.g., "I aim to get a good education, not just completing tasks to get credits"), and doing their best (e.g., "I strive for Merit or Excellence even when I don't need this to achieve my goals"). (p. 26)

Doing just enough can be characterised as students doing the minimum work required to achieve goals:

> [D]oing the minimum work required to achieve goals (e.g., "I work for the number of credits I need at each level, no more"), friends' opinions (e.g., "What my friends think influences whether I work in school"), and paid work outside school (e.g., "The subject interferes with part-time work commitments"). (p. 26)

Meyer et al. (2006) report that "the strongest predictors of high academic achievement and higher grades were a high motivation orientation towards *Doing My Best* and a low motivation orientation towards *Doing Just Enough*" (p. 2, emphasis in original). Students motivated by 'doing my best' were most likely to select subjects based on interest and career goals and were "high on the *Getting Feedback* and *Excellence* and low on the *Work Avoidance* factors" (p. 2). Students motivated by 'doing just enough' chose subjects because of external factors such as the influence of friends and were "high on *Work Avoidance* and also somewhat high on *Getting Feedback"* (p. 2, emphasis in original).

The researchers claimed that the structure of the NCEA encouraged some students to adopt a minimalist approach. They found that many students in their study agreed that it was hard to be motivated to do more than the minimum 80 credits, and many indicated that there was little motivation to aim for the higher grades of Merit and Excellence when these carried no extra credits. However, the results suggest a negative relationship between the motive to do just enough and the number of credits achieved:

> This may mean that many of them will not obtain enough credits to actually get by, because people do not always achieve exactly what they aim for. So students aiming to do just enough may actually fail to achieve their goal, not because they lack the required ability but because their motivation orientation leads them to achieve less than they are capable of. If these same students are motivated to do their best, they are more likely to pass the required number of credits, and also obtain Merit and Excellence grades. (p. 2)

Self-efficacy

Students' sense of self-efficacy can influence their approach to assessment tasks (Kluger & DeNisi, 1996). Self-efficacy refers to how capable the learner feels of succeeding in a particular task, or type of task (Harlen, 2006). Research with grade three students undertaken in the United States by Brookhart and DeVoge (1999)

demonstrated that teachers' implicit and explicit assessment practices impact on students' self-efficacy. In particular, the researchers emphasised that feedback from past assessed work, either by the teacher or through self or peer-assessment, helps students decide whether they are capable of undertaking future work successfully. This self-efficacy affects the amount of effort students are willing to invest in a particular task. Highly efficacious mathematics students in Benmansour's (1999) study showed higher intrinsic goal orientation, lower test anxiety, and used a wider range of active self-regulated learning strategies, than students with low self-efficacy.

Empowerment

The role that the nature of assessment plays in students' sense of autonomy and empowerment has been highlighted by Anderson (1993):

> When assessment is perceived exclusively as the teacher's domain, students willingly wait for the teacher to judge their success or failure. In contrast, when children continually participate in the assessment process, they learn to recognise their own expertise. As active assessors, they necessarily exercise a more autonomous and decision making role in their learning. Consequently, instead of being used to gain power over a child, assessment empowers that child. (p. 103)

Recent research in New Zealand suggests that the introduction of SBA in secondary schools has resulted in students becoming more empowered and self-regulated (Barrington, 2004; Hipkins et al., 2005; 2007; Meyer et al., 2006). For example, Hipkins et al. (2005) reported that teachers have observed an increase in student confidence. Similarly, the principals in Barrington's study perceived that students were more empowered under NCEA through greater knowledge of the impact of assessment on their learning, more transparency of the grading system, and a clearer alignment between curriculum and assessment.

Findings from an OECD study (2005) attributed greater student empowerment and engagement in learning to an increased use of formative assessment. Teachers in the study reported stronger relationships with their students, and increased contact with parents. In one case study school, students commented that "instead of just getting grades, they felt they were involved in a process with their teachers" (p. 72). Students across the case study schools were identified as taking "more responsibility for their

learning and taking more pride in their work" (p. 72). Similarly, research into SBA from Canada and the USA found that the transparency of assessment processes has resulted in students gaining more knowledge of the learning process itself and applying higher order thinking skills to their learning (Boss, Endorf, & Duckendahl, 2001; Brush, 1997; O'Donnovan, Price, & Rust, 2000).

Performance vs. learning orientation

Common to many theories on motivation is a reference to goal orientation. Students' learning goals influence their direction and distribution of effort towards learning tasks (Dweck & Leggett, 1988). Dweck and Leggett proposed a framework involving 'learning' and 'performance' goal orientations. Learning oriented students are characterised by a desire to increase their own knowledge and understanding of a topic. They persist in the face of difficulties, and achieve a sense of mastery based on self-referenced standards. Performance oriented students, on the other hand, are characterised by a desire to publicly demonstrate their competence and do better than others. Performance oriented students do not necessarily desire to improve their understanding of the material, and so may seek the easiest way to achieve goals and consider ability to be more important than effort (Ames, 1992).

Performance oriented practices of comparing students' achievements with one another, and of emphasising competition rather than personal improvement, have a negative effect on students' learning processes (Black & Wiliam, 1998a). Such a competitive environment threatens a student's sense of worth. Students come to perceive themselves only as worthy as their ability to achieve competitively and, since only a few can win at this game, the majority of students may feel incompetent and hence unworthy (Ames, 1992). Research also suggests that performance orientation can have a constricting effect on curriculum. In classrooms where social comparison information becomes emphasised, students may learn that what is not assessed is not worth learning (Ames, 1992). As such, performance oriented students display shallow learning strategies, emphasising memorisation and rote learning (Dweck & Leggett, 1988).

A range of studies suggests that a learning orientation promotes long-term, high-quality involvement in learning (Ames, 1992; Dweck, 1992; Roderick & Engel, 2001; Schraw, Horn, Thorndike-Christ, & Bruning, 1995; Schunk, 1996). Students with learning goals demonstrate stronger evidence of superior learning strategies, higher self-efficacy, and have more positive views and interest in their school work. A learning orientation positively affects achievement and the development of metacognition (Schraw et al., 1995; Schunk, 1996). Furthermore, students with learning goals are willing to expend more effort than those with performance goals (Harlen & Deakin Crick, 2003).

Students' individual goal orientations are heavily influenced by the predominant goal orientation of their teachers and school (Ames, 1992). Concerns raised over school systems that are too heavily performance oriented at the expense of a learning orientation are summed up by Covington and Müeller (2001):

> If high grades become increasingly important as students grow older, not only as indications of their personal worth but also as passports to prestigious occupations, then what becomes of the intrinsic value of learning? (p. 159)

Implications for classrooms.

The manner in which students are assessed is one of the most important factors promoting, or inhibiting, motivation and engagement in learning (Ames, 1992; Harlen & Deakin Crick, 2003; Stiggins, 2001). How the assessment environment impacts on the learning environment is largely determined by the teacher (Hosking & Shield, 2001; Stiggins, 2001; Watson, 1998). Stiggins (2001) has claimed that teachers can "enhance or destroy students' desire to learn more quickly and more permanently through assessment, than through any other tools at their disposal" (p. 36). In a review on the impact of high-stakes assessment on motivation, Harlen and Deakin Crick (2003) identified a number of pedagogical strategies for "ensuring that the benefits of summative assessment can be had without the negative impacts on students' motivation for learning" (p. 203). Key amongst these strategies were: the development of students' understanding of assessment criteria, improving students' self-assessment skills, emphasising learning goals rather than performance goals, and providing feedback to students in relation to those goals. Harlen and Deakin Crick

highlight the importance of a supportive school wide ethos that uses "assessment to convey a sense of learning progress to students" (p. 202) for these strategies to be effective .

However, effective translation of research findings into classroom practice is a complex issue (Black & Wiliam, 2003). In the publications resulting from their 1998 review, Black and Wiliam argued that it was neither advisable not possible to "lay out what formative assessment would look like in practice" (Black & Wiliam, 2003, p. 629).

> Thus the improvement of formative assessment cannot be a simple matter. There is no 'quick fix' that can be added to existing practice with promise of rapid reward. On the contrary, if the substantial rewards of which the evidence holds out promise are to be secured, this will only come about if each teacher finds his or her own ways of incorporating the lessons and ideas that are set out above into her or his own patterns of classroom work. This can only happen relatively slowly, and through sustained programmes of professional development and support. (Black & Wiliam, 1998b, p. 15, emphasis in original)

Consistent with their espoused view, Black and Wiliam conducted research investigating how teachers could be supported to affect change in their formative assessment practices. KMOFAP involved 24 mathematics and science teachers from six schools in exploring how formative assessment might be put into practice. Rather than imposing a model of "good formative assessment" (p. 630) on teachers, the project focused on supporting teachers in developing their own professional practice. This approach, consistent with a sociocultural perspective on learning, recognised that the teachers involved in the project were "engaged in a process of knowledge creation, albeit of a distinct kind, and possibly relevant only in the settings in which they work" (Black & Wiliam, 2003, p. 631). The researchers noted that as the teachers explored the relevance of formative assessment for their own practice, they transformed ideas from other teachers into new ideas, strategies and techniques, and these were in turn communicated to other teachers, creating a 'snowball' effect. Subsequent research (Black et al., 2006) reinforced the importance of co-constructing pedagogical strategies supportive of formative assessment principles.

As such, although research can identify key principles that will affect students' use of formative assessment strategies, the implementation of these strategies will need to be

negotiated between the participants in the process. This will initially be between classroom teachers, but will subsequently involve the students as active participants in the classroom community of practice.

2.7 Conclusion

Research conclusively demonstrates that improving formative assessment practices can positively impact on students' learning. There is debate, however, about whether a single assessment system can practically serve both high-stakes summative purposes as well as formative purposes of assessment. There is an extensive body of research that indicates that an overemphasis on high-stakes summative assessment: reduces formative practices, has a negative impact on motivations for learning, restricts curriculum, and encourages teachers to adopt transmissive pedagogical models. However, there is a growing body of research that indicates that SBA high-stakes assessment focuses teachers' and students' attentions on key aspects of learning, aligns curriculum and assessment, encourages a learning orientation in students, and helps students become active participants in their own assessment.

Much of the research on the impact of high-stakes assessment on pedagogy and learning has been conducted in environments where the results of assessment are very high-stakes for either the student or the school. There is a limited research base on the impact of high-stakes SBA assessment on students in New Zealand secondary schools. It remains to be seen whether the design of NCEA will support effective formative assessment practices in the classroom, or whether the summative purpose of NCEA will dominate.

Research has provided us with a clear picture on a number of important formative assessment practices. Self- and peer-assessment increase students' metacognition, understanding of assessment criteria, and engagement with assessment. However research has also demonstrated that self- and peer-assessment are not common formative assessment practices. Research has shown that task-oriented feedback, rather than ego-oriented feedback, is the most effective formative feedback practice. However, much of the research on effective feedback practices assumes that the

feedback will prompt the individual student to address learning deficiencies. There is a paucity of research into the range of ways that students engage with feedback, especially within a sociocultural learning environment.

While sociocultural perspectives of assessment promote the use of a wide range of alternative assessment practices, research indicates that senior secondary school mathematics teachers are reluctant to use alternative assessments strategies, preferring to use traditional pen and paper tests. Given mathematics teachers' predilection to use traditional assessment practices, there is a limited research base on how these practices can be used formatively in a sociocultural framework. In particular, given that realistic mathematical inquiry is increasingly acknowledged as a key component of mathematics learning, there is little research on how students independently and collectively use assessment feedback as a prompt for future inquiry based learning.

Although there are a number of studies that have investigated students' perceptions of assessment events there is a paucity of studies that have investigated students' perceptions of the formative potential of SBA, and none of these has been conducted in New Zealand. Given the definition of formative assessment adopted in the current research, the students' voice is a critical addition to understanding the complex nature of the interplay between assessment and learning. Unfortunately, the students' voice has been underrepresented in educational research, especially in areas that directly affect them (Fielding, 2006; Vaughan, 2003). In particular, little is known about *how* students develop knowledge of effective formative assessment strategies, and their understandings and preferences for engaging with feedback.

3 METHODOLOGY

3.1 Introduction

The methodology for this research project is informed by the contemporary sociocultural theory of learning as outlined in chapter 2. According to this theory, learning and understanding are socially constructed in a community of practice (Lave & Wenger, 1991). In particular, this project examines how students develop understanding of how assessment can be used formatively to positively impact on their learning. It is theorised that the social norms within the classroom assessment environment play a major role in the development of these understandings (Sadler, 1989). These understandings cannot be separated and studied without cognisance of the environment within which they were formed. Since students' perspectives are socially situated, it is necessary to use a methodology that contextually examines these perceptions. Accordingly, a case study methodology was adopted for this research project.

This chapter is divided into two main sections. The first of these sections discusses the theoretical conceptualisation of case study methodology. It starts by examining the nature of reality adopted in this research, and how this influenced the decision to use case study methodology. The section then discusses the philosophical and theoretical characteristics of case study methodology in general, including a discussion of the nature of quality criteria such as validity and reliability. Following this is a discussion of the importance of the student voice in research on education. An analysis of the characteristics of the data generation methods is provided, and finally, ethical issues concerning participation in educational research are considered.

The second section examines the methodology in action. It details the selection and characteristics of the case study school, and the case study participants. The section then outlines the data generating procedures, and concludes with a discussion of the framework used to analyse the data.

3.2 The nature of reality

In striving to make sense of the world, researchers must first consider perceptions of the nature of reality. Two of the more common conceptualisations of reality are the 'positivist' and 'interpretivist' paradigms (Bassey, 1999). From a 'positivist' view, reality is something that is out there waiting to be discovered. It exists whether someone observes it or not and is independent of who observes it. In other words, two independent observers would both discover identical realities. In contrast to this, the 'interpretivist' view of reality argues that reality is not something that exists waiting to be discovered; rather it is a socially constructed concept. It is something that is constructed in the minds of the observer and those involved in the situation, and is therefore consistent with a sociocultural perspective of learning. From an interpretivist perspective, multiple realities may exist for a given situation. People perceive, and therefore construct, the world in similar but rarely identical ways.

Positivists would argue that, since reality is independent of the observer, given time and patience, it should be possible to discover and explain the realities in human behaviour and relationships, and make factual statements about them. Also, because positivists do not see themselves as a significant variable in the research, another observer would discover and describe the same reality, and make the same factual statements. In contrast, an interpretivist would argue that the researcher is a potential variable in the investigation of reality. By asking questions or observing the researcher may change the situation they are trying to observe. Indeed reality is, to a certain extent, filtered through the researcher's eyes. Any statements made would be interpretive to some degree. Thus it is unlikely that another researcher, researching the same situation, would discover a reality that was the same. An implication of this view is that consideration of the researcher's experience in the chosen educational field must be made, in order to maximise the validity of interpretations of the data. The present research investigates students' perceptions. It is argued below that these perceptions are socially and contextually situated within the classroom. Accordingly, this research is guided by an interpretivist position, and the details of the researcher's experience in the secondary school sector are outlined in section 3.4.2

3.3 Case study methodology

"Human behaviour, thoughts and feelings are partly determined by their context. If you want to understand people in real life you have to study them in their context and the way they operate" (Gillham, 2000, p. 11). A case study methodology is particularly appropriate to study human behaviour in the real world 'as it happens' (Stake, 1995). It is the preferred methodology in situations where it is very difficult to separate the phenomenon under investigation from the surrounding influences (Yin, 2003). A distinguishing feature of case study methodology is the belief that human systems develop a characteristic wholeness or integrity, and are not simply a loose collection of traits that can be investigated without concern for the relationship between them (Sturman, 1999).

The nature of case study in education has been debated in the research community for more that 30 years. Historically, case study has its roots in the disciplines of sociology, anthropology, history, psychology, law and medicine. Perhaps because of this broad base, there seems to be little agreement about what case study is. Stake (1995) recognised the complexity of defining case study, arguing that: "We cannot make precise definitions of cases or case studies because practices already exist for case study in many disciplines" (p. 2). At the simplest level, case study has been characterised as "a sort of catch-all category for research that is not a survey or an experiment and is not statistical in nature" (Merriam, 1998, pp. 18 – 19).

A number of key educational researchers have contributed to debate over definitions of case study research in education. Robert Yin (2003), one of the leading proponents of case study research in the United States, defines it as "an empirical enquiry that investigates a contemporary phenomenon within its real-life context especially when the boundaries between phenomenon and context are not clearly evident" (p. 13). Anchored in real life, case studies can give a rich holistic account of a phenomenon. It is this holistic view of the interrelationship of contributing variables that requires an in-depth study from within the educational environment, as opposed to the contrived contexts of experiments or survey. "Case studies offer a means of investigating complex social units consisting of multiple variables of potential importance in

understanding the phenomenon" (Merriam, 1998, p. 41). Cohen, Manion, and, Morrison (2000) expanded the discussion, taking the position that:

> Unlike the experimenter who manipulates variables to determine their causal significance or the surveyor who asks standardised questions of large, representative samples of individuals, the case study researcher typically observes the characteristics of an individual unit—a child, a clique, a class, a school or a community. The purpose of such observation is to probe deeply and to analyse intensively the multifarious phenomena that constitutes the life cycle of the unit with a view to establishing generalisations about the wider population to which that unit belongs. (p. 185)

The strength of case study research is that it is conducted 'in situ' rather than in a laboratory, and, as such, it can provide a rich description of the research problem from within the surrounding context. Contexts are unique and dynamic; therefore case studies investigate the complex dynamic and unfolding interactions of events and people in a unique setting (Cohen et al., 2000). This is especially true in educational settings such as the one in the current research, where the interplay between the school, the classroom environment, the teachers, the students, the curriculum, and assessment forms a complex web.

Within an educational environment, Bassey (1999) describes case studies as:

> [A]n empirical enquiry which is: conducted within a localised boundary of space and time (i.e. a singularity); into interesting aspects of an educational activity, or programme or institution, or system; mainly in its natural context and within an ethic of respect for persons; in such a way that sufficient data are collected for the researcher to be able to explore significant features of the case, to create plausible interpretations of what is found, to test for the trustworthiness of these interpretations, to relate the argument or story to any relevant research in the literature, to convey convincingly to an audience this argument or story. (p. 58)

Case studies are distinguished more by the singularities they investigate than by the methods they employ (Hitchcock & Hughes, 1995). Case studies rely on triangulating multiple sources of data, and, although qualitative methods form the bulk of data generating techniques, the use of quantitative methods is not precluded. The data generating methods used are informed by the prior development of theoretical propositions (Yin, 2003).

The current research is informed by the theoretical framework developed by Sadler (1989). Key to this framework is the notion that assessment information is only formative if students use it to positively impact on their own learning. Therefore, students' perceptions of the formative feedback they receive, and their use of this feedback, is central to this project. Accordingly, this research used data generation methods designed to capture and contextualise the student voice. These methods are discussed in more detail in section 3.7

3.3.1 Categories of case studies

A number of researchers have developed terminology distinguishing various categorisations of case study methodologies. Yin (1993) identified three sub-categories of case studies in terms of their outcomes. He defined these as descriptive, explanatory and exploratory case studies.

Descriptive case studies present "a complete description of a phenomenon within its context" (Yin, 1993, p. 5), often via narrative accounts. This is similar to Bassey's (1999) 'story-telling and picture-drawing' categorisation, and consistent with Stake's (1995) 'intrinsic' case studies. In a descriptive case study, the researcher looks into a particular situation for its own sake, irrespective of outside concerns. They are interested in it, not because by studying it we learn about other cases or about some general issue or problems, but because we need to learn about that particular case. In other words we have an intrinsic interest in the case.

The second of Yin's (1993) categories, explanatory case studies, "present data on cause effect relationships—explaining which causes produced which effects" (p. 5). Bassey (1999) described this category as 'theory-testing', designed to gather evidence to confirm, or refute, accepted theory. It is Yin's (1993) exploratory case study that is most relevant to the current research. An exploratory case study is designed to investigate one or more particular situations in order to try to understand an outside concern. The case is chosen because it is thought to be typical of something more general and we may gain insight into the overriding question by studying the particular case. The focus is on the issue rather than the case and the researcher will start and end with the issues dominant. In this regard the case study is 'instrumental'

(Stake, 1995) to accomplish something other than an understanding of the particular case. This paradigm seeks to understand, describe and explain what is happening within a particular case without making value judgements or trying to induce change.

Often exploratory case studies are undertaken because there is a lack of theory, or existing theories fail to adequately explain a phenomenon (Yin, 2003). In this regard, exploratory case studies 'seek' information with the aim of formulating theory, and are consistent with Bassey's (1999) 'theory-seeking' case study. From an interpretivist perspective, theory is something researchers create, rather than something that is there and established. This does not necessarily mean that case study research is entered into without some general theoretical stance, but this stance is merely a starting point. The evidence looked for is dictated initially by the study's broad aims, but increasingly the study is directed by the successively revised theories or explanations of what has been found. The current research was informed initially by the broad base of literature surrounding formative assessment, and, in particular, Sadler's (1989) theory of formative assessment. This general theoretical stance informed the data generating phase of the pilot project, and the subsequent main project. During early data generation, it emerged that students' development of formative assessment practices appeared to be socially situated. Accordingly a sociocultural perspective incorporating communities of practice was used to inform later data generating phases. This resulted in a closer examination of the social interactions that help students develop knowledge of how to use assessment information formatively.

This exploratory perspective to case study methodology is consistent with Nachmias and Nachmias's (1992) concept of 'research before theory', where research may pose new problems for theory, require theoretical innovation, refine existing theories, or serve to vary past theoretical assumptions. Exploratory case studies are often used as pilot studies to subsequent studies, by defining questions and hypotheses that may be tested using subsequent case studies, or alternate methodologies such as large scale surveys or experiments. However, the use of case studies as a preliminary to other studies does not reduce the legitimacy of case studies as a significant research method in its own right (Cohen et al., 2000).

3.4 Quality criteria: validity and reliability

In any research study the reader must judge the degree to which the research is both valid and reliable. In essence they must be convinced that the project has trustworthiness (Gillham, 2000). But the concepts of validity and reliability are more contentious for qualitative than for quantitative research. Indeed, the research community is still developing a consensus as to the appropriate criteria for assessing validity and reliability in qualitative research (Merriam, 1998).

3.4.1 External Validity

External validity is the extent to which the findings of a given study can be generalised and applied to other situations. A criticism often levelled at case study research is that it lacks external validity since the results are not able to be generalised. Classical science research, based along positivist lines, has traditionally been aimed at finding generalisations based on large samples of data generated from carefully controlled experiments with only a very few variables (Gillham, 2000). Indeed, when most people think of 'research evidence' they think of this 'scientific evidence'. Social scientists, on the other hand, generally study smaller samples and have more variables to deal with. Scientific generalisations, therefore, are often not appropriate for summarising social findings because of the sheer complexity of social events (Bassey, 1999). Generalisation from one group of people, or one institution, to another, may be compromised because there are too many elements that are specific to that group or institution (Gillham).

Adopting a case study methodology, however, does not preclude an interest in generalisations. Generalisations are "matters of judgement rather than calculation and the task of case study is to produce ordered reports of experience which invite judgement and offer evidence to which judgement can appeal" (Stenhouse 1988, p. 49). In other words, it is up to the reader to decide whether the researcher has presented enough evidence to support any general statements that might have been made.

Generalisations made in case study research are retrospective in nature (Stenhouse, 1988). They provide us with an analytical look from a historical perspective. We can

look back over the evidence and make generalisations about what *has* happened and, perhaps, what *might* happen in the future. To add to this discussion, Bassey (1999), leaning heavily on the concept of 'fuzzy logic', introduced the notion of 'fuzzy propositions' (more tentative) and 'fuzzy' generalisations (less tentative). Essentially a fuzzy generalisation asserts that something *may* happen without any measure of its probability. It is a qualified generalisation, carrying the idea of possibility, but not certainty. Instead of 'do x, instead of y, and your pupils *will* learn more' we have 'do x, instead of y, and your pupils *may* learn more" (Bassey, 1999, p. 46, emphasis in original). This viewpoint is consistent with Berg's (2004) argument that, if case studies are undertaken properly, they should not only fit the specific individual, group or event studied, but also generally provide understanding about similar individual, groups and events.

To add to our understanding of the nature of generalisations made from case study research, Stake (1995) introduced the hierarchical terms 'petites' generalisations and 'grandes' generalisations. The first—petites generalisations—refers to general statements made within a study and are relevant only to that study. The latter—grandes generalisations—concerns generalisations about issues with reference to the wider population, of which the case is one example. Stake concludes by reminding us that, above all else, it is important that researchers make clear the speculative and tentative nature of their assertions.

To increase the validity of generalisations, Tripp (1985) argued for a cumulative process in bringing case studies together. In essence, this is what is done in arenas such as psychology or the law where case studies have historically been the prevalent research methodology. Tripp saw this process as a means to provide 'qualitative generalisations' in which individuals, meeting the facts of a new case, apply them to their knowledge of similar cases, and so develop a personal understanding. This view is consistent with Stake (1995) who used the term 'generalisation' to refer to the learning process through which we individually acquire concepts and information and steadily generalise them to other situations as we learn more. Accordingly, through careful replication of case studies over time, we should be able to extend our 'petite' generalisations to 'grandes' generalisations.

3.4.2 Internal validity

Internal validity addresses the match between research findings and reality (Merriam, 1998). Is the researcher's analysis and report of the evidence a 'true and correct' representation of respondents' perceptions? To address this issue we must first consider two presuppositions. Firstly, that a true and unique interpretation of the data collected exists, and secondly, that this interpretation may be determined by standard applicable technical procedures.

If we consider the first presupposition, that a true and unique interpretation of the data exists, an interpretivist view of reality argues that it is socially situated and, as such, multiple realities exist (Bassey, 1999). Merriam (1998) argues that if reality is viewed in this manner, internal validity is a definite strength of case study methodology. The critical issue is not the determination of one singular and absolute truth but rather the assessment of the relative plausibility of an interpretation when compared with alternative plausible interpretations (Yin, 2003). Is the researcher's interpretation of the data 'trustworthy'? "[T]he concern is with meaning, understanding and description, rather than discovering an objective truth" (Bourke, 2000, p. 66).

If we consider the second presupposition, since multiple realities exist, this implies that a standard applicable approach is invalid in analysing the data as it will not identify multiple realities. From an interpretivist perspective, meaning and interpretations are contextually and culturally grounded. As such, naturally occurring discourse, such as interview narrative and the interpretation of observational data, rely on the shared cultural understanding between the researcher and the case study participants. It is this shared cultural knowledge base that allows for the analysis of the interview material by a process that Mishler (1986) calls 'expansion'. In this process, researchers bring together all the shared knowledge and information available to help with the interpretation of the narrative. To accomplish this expansion of meaning researchers must use their 'best understanding', making references to other material as well as to the shared knowledge between the participants. They must introduce factual material from other parts of the interview as well as from general knowledge of the situation. Not all statements can be taken at face value and must be judged within the context of the broader discussion. As Gillham (2000) argues:

"Interpreting research data is more than a matter of good intentions. It requires discipline and concentration to present a 'true' picture: anything that gets in the way of that threatens the validity of your research" (p. 25).

In the current research project, the accuracy of the interpretation of the data is enhanced by the researcher's background. As the researcher, I was a secondary school teacher for 18 years and, in my last position, I was responsible for the implementation of NCEA throughout the school. Additionally, I was a facilitator involved in the professional development of teachers during the implementation of NCEA. This background provided me with an in-depth knowledge of the philosophical underpinnings, as well as the organisational structure, inherent in the NCEA. As such, I am closer to the research topic through shared understandings with the participants.

3.4.3 Reliability

Reliability has traditionally been concerned with the ability to replicate a given research project; if the study was repeated would the same results occur? This perspective is based on the assumption that there is a single reality and that studying it repeatedly will yield the same results. However this traditional view of reliability has been identified as a difficult area in social sciences research due to the fluid nature of human behaviour and perceptions. Since individuals' experiences will influence their interpretations of reality, there are no benchmarks to judge reliability in the traditional sense of quantitative research (Merriam, 1998).

Lincoln and Guba (2000) suggest that rather than using the term 'reliability' in the traditional sense it would be more beneficial to consider the 'dependability' or 'consistency' of the results obtained. Rather than requiring that an outside researcher obtain the same results by replicating the study it is more important that, given the data collected, the results make sense within the context of the research situation.

Within this research project there are two main reliability issues: how consistent the data from different participants was, and how likely is it that similar data and conclusions would emerge from replication with other samples of teachers or students

in other schools? Considering the first point, the researcher must be cognisant of the variability within the data, and report the consistency or contrasting nature of the data as reliably as possible. Recognition is also made that student' perceptions are likely to develop over the course of the research project, due partly to their involvement in the research and partly due to their expanding experience base. As such the researcher must reliably report the consistency or changing nature of individual student's perception over time.

The second reliability issue relates to whether similar data and findings would emerge from studying a different sample of students and teachers. The choice of case study methodology recognises that perceptions and experiences are contextually situated. As such, the data and conclusions generated from this research project are specific to the research participants and school. Secondary schools in New Zealand are legislated to respond and adapt to their community of learners, and the wider community within which the school is situated. Accordingly, it is reasonable to expect a range of teachers' assessment practices and students' perceptions of these practices. As such, it is acknowledged that the findings of this project represent a subset of the range of possible perceptions of secondary school students and teachers. Further research will need to be undertaken to more fully describe the complexity of the interplay between teachers' practices and students' perceptions' regarding assessment experiences. Additional issues regarding the reliability of the various methods used in this research are discussed in subsequent sections.

3.4.4 Triangulation

Data triangulation involves the use of two or more methods of data generation with the aim of explaining the richness and complexity of human behaviour by examining it from more than one perspective. Reliance on a single method of data generation may bias, or distort, the researcher's picture of events and relationships. However, triangulation should not simply combine different kinds of data, it should attempt to relate them, so as to counteract the threats to validity identified in each (Cohen et al., 2000). The use of multiple sources of evidence, each with its own strengths and weaknesses, is a key characteristic of case study research (Gillham, 2000).

Traditionally, data generation in case study research has centred on the qualitative techniques of interviewing, observation and document analysis (Merriam, 1998). However case study methodology does not preclude the use of quantitative methods such as the Assessment Experience Questionnaire (AEQ, Gibbs & Simpson, 2003) used in the current research. The AEQ will be discussed in more detail in a later section, but it is pertinent to note a number of key points as they relate to qualitative data and triangulation. Firstly, the AEQ seeks a quantitative measure of students' perceptions of influences on assessment. These perceptions are socially situated and, as such, are a qualitative dimension of their assessment experiences. Secondly, the quantitative measure from the AEQ can be used to compare the perceptions of the nine case study students with the larger school population of Y12 Traditional Mathematics students. If it can be established that there is no significant difference between the two groups, then this increases the trustworthiness of the data generated from the case study students as being representative of the Y12 Trad maths cohort. In this respect, the AEQ serves as a valuable source of triangulation.

This research project is interested in the 'realities' students have constructed and how they make sense of the world they live in; realities that are constructed by individuals interacting with their social worlds (Merriam, 1998). From a sociocultural perspective these realities are personal, holistic, multidimensional, and ever-changing. With respect to this project, students' perceptions of the formative use of assessment information are determined by their individual and collective experiences. Socially situated in time and space, students' perceptions of reality will be fluid, rather than static. As such, we could reasonably expect to have a degree of variability in perceptions, both between students and within students over time and social and cognitive contexts. In accordance with this view, triangulation not only includes generating data using different methods, it also involves using the same method over a period of time. Accordingly, this project generated data over a school year.

3.5 Selection of the singularity

A case study is a study of a singularity conducted in depth in natural settings (Bassey, 1999). From an educational perspective, a singularity could be either an individual or

groups of schools, teachers, or students. The key feature of a singularity is that there is a boundary around the group being considered. The choice of a singularity is determined according to the needs of the research study, rather than external criteria such as a random selection. "Since generalisation, in a statistical sense, is not a goal of qualitative research, representative sampling is not necessary or even justifiable" (Merriam, 1998, p. 61). Rather the case should be representative of the general issue we are trying to investigate.

This research project utilised 'purposeful' sampling (Patton, 1990). Purposeful sampling is based on the assumption that the investigator wants to discover, understand and gain insight into a particular issue and therefore must select a singularity from which the most can be learned. In selecting a singularity, the researcher should decide what selection criteria are essential when choosing the people or sites, and then use his or her special knowledge or expertise about the group to select subjects who represent this population (Merriam, 1998). In order to investigate students' perceptions and autonomous use of formative assessment within an SBA mathematics education environment it was important to choose a school that publicly supports student autonomy in learning, and is proactive in its support and use of SBA. From within the research school, one of the Y12 mathematics courses (Y12 Trad Maths[4]) was selected as the singularity for the case study. Since this research seeks to explore the 'student voice', a group of Y12 Trad Maths students was selected (The selection process is discussed in detail in section 3.9.2). The student participants were representative of the same experience or knowledge base (i.e. NCEA), and so had homogeneity, but with sufficient variation to allow for contrasting opinions (Krueger & Casey, 2000). Indeed Mishler (1986) argues that a small group of well-informed acute observers brought together as a discussion group is many times more valuable than any statistically random sample.

[4] Y12 Trad Maths is the common name for the Y 12 Traditional Mathematics course. This course is designed to prepare students for the full range of Y13 mathematics courses.

3.6 The students' voice

A significant feature of Sadler's (1989) theory is that assessment is only formative if students use the feedback to positively impact on their learning. The implication of this position is that researchers need to be cognisant of the understandings students have of how assessment information can be used formatively. To achieve this, the current research included the 'students' voice'.

Educational researchers are increasingly interested in understanding students' perspectives of learning, teaching and assessment. It has been argued that if we want to have a positive impact on students' lives, we need to include their perspectives as a central feature of research and policy development (Vaughan, 2003). This is best achieved through research that consults students directly about aspects of their schooling. As Gipps and Tunstall (1998) contend, "the pupil's voice has come to be seen as crucially important to understanding the complexities of learning in school. Naturalistic research approaches which are grounded in pupils' accounts are needed to probe this" (p. 149). A number of recent studies within the New Zealand context highlight the value of the 'students' voice' (e.g. Alton-Lee & Nuthall, 1990, 1998; Alton-Lee, Nuthall, & Patrick, 1993; Anthony, 1994; Bishop, 2003; Bishop Berryman, Tiakiwai, & Richardson, 2003; Bourke, 2000; Kane & Maw, 2005, Nuthall & Alton-Lee, 1993; Nuthall, 1999a, 1999b, 2002; Rawlins, 2000). For example, recent research by Kane and Maw (2005) began with the premise that "to improve classroom practice in secondary schools we need to ask for, and attend to, the needs and views of students" (p. 311).

Listening to what students have to say about their school experiences is part of a growing movement worldwide, largely driven by changing perspectives of teaching and learning (Fielding, 2006). Students' understanding of their own role in the learning process has been highlighted as an area for prospective research by a number of influential researchers in education (Tunstall & Gipps, 1996; Pollard et al., 2000). Contemporary perspectives of learning re-conceptualise the roles of the teacher and the student. Increasingly, children are seen as consumers of education, making

decisions about what and how they would like to learn and how they are going to be assessed.

Indeed, research in mathematics education has identified that students and teachers often hold differing views of what is, in essence, a shared experience (Anthony, 2000, Brown, 2002; Rawlins, 2000). For example, Brown (2002) identified tensions with teachers and students continuing to 'talk past each other in terms of their conceptions of learning' (p. 72). Given the sociocultural learning perspective adopted in the current research, it is reasonable to expect that the 'reality' of the classroom appears differently to these two groups who bring different learning, social and cultural histories to the classroom. Research into students' perceptions of the classroom is complicated by the interdependence of teachers' practices and students' views of learning. Although research has recognised that effective teachers' practices must be aligned with the ways students learn (Alton-Lee & Nuthall, 1998), students' views of learning are heavily influenced by the nature of the activities that they are exposed to in the classroom, and the perceived values these activities convey (Brown, 2002). In other words, students construct views of learning in ways they have been socialised to do so through their perceptions of what teachers value, perceptions which may be at odds with the actual values held by teachers. Research must be cognisant, therefore, of how students' learning is shaped by their classroom experiences, including their assessment experiences (Kane & Maw, 2005). Researching only teachers' perceptions provides us with a partial, and perhaps biased, picture of the impact their pedagogical practices have on students' learning. The students' voice is critical; it provides us with an opportunity to understand a different perspective of how teachers' activities affect students' learning (Nuthall, 2002).

Involving students in research often has surprising results, highlighting issues that teachers had not thought important. Additionally, students often benefit from such research in indirect ways. Fielding (2006) claims that because they are respected, listened to and taken seriously, students become more confident and more positive about learning and about school. "Once you start to have faith in young people and give them the opportunity to have a say, teaching and learning can be a joint activity" (Fielding, 2006, p. 12).

Research involving students' perceptions potentially raises a number of validity issues with regard to data generation and interpretation that must be addressed. For example, student participants may mould their reported perceptions to meet what they consider to be the goals of the research. Within this research, two reasons for this behaviour can be proposed. Firstly, the students may want to satisfy the researcher and may pick up on subtle clues as to what the researcher is hoping to find out. This concern is especially valid during the focus group interviews and measures taken to counteract this threat are discussed in more depth in section 3.7.1. Secondly, students may be concerned that some of their views will not be well received by the school, and may choose not to put them forward. This is an ethical consideration and is discussed in more depth in section 3.8.2.

A second validity issue concerns the trustworthiness of students' reports of particular classroom events. This does not imply that students created 'realities' are necessarily inaccurate or wrong. Indeed the interpretivist perspective adopted in this research validates their perceptions and memories of particular classroom events. But it is useful to contextualise their perceptions through the use of classroom observations and interviews with the teachers.

3.7 Data generating tools

Case study methodology uses a wide variety of data generating techniques. Although this research focused on students' perspectives it also generated some data from the classroom and the classroom teacher, in order to provide a deeper understanding of the social and contextual influences on the learning environment. The main data generating tools used in this project were: focus group interviews, written response questionnaires, classroom observations, and a quantitative questionnaire. The theoretical aspects of these data generating techniques are discussed in more depth in the next sections.

3.7.1 Focus group interviews

Focus group interviews are a potentially productive data generating tool, but must be used with care if the limitations of the method are to be minimised. Cohen et al.

(2000) argue that focus groups provide a setting where individuals feel comfortable in self-disclosure, and the group dynamics can provide a more exhaustive view of the issues or topics. Focus group interviews not only gather the views of individual respondents but also provide: "a situation where the synergy of the group adds to the depth and insight" (Anderson, 1990, p. 241). As such, focus groups maximise the advantages of the group dynamics of the situation and provide stimulus for elaboration and expression. However the potential for peer group pressure amongst the group must also be recognised.

Although a collective group culture can produce a wealth of material, focus group interviews are often criticized for creating conformity of views (Morgan & Krueger, 1993). Individuals' behaviour can be subject to group influence and participants may espouse views that they do not genuinely hold. Several possible reasons for this phenomenon can be hypothesised. Subtle cues from the interviewer may make students respond in ways they believe are expected by the interviewer. The interviewer may provide, albeit subconsciously, immediate reward for favourable comments made by the participant, for example the nod of a head or verbal suggestions indicating agreement. These cues may encourage the participants to simply 'give the researcher what they want'. At other times a participant may put forward a viewpoint that they hope will carry favour with another member of the group that the respondent admires. (Albrecht, Johnson, Walther, & Krueger, 1993). To accommodate these limitations, the current research adopted a triangulation approach by revisiting specific topics from different angles, both within a given focus group interview, and across different focus group interviews. This allowed the robustness of the emerging findings to be tested. Additionally, every effort was made to ensure that subtle clues were not conveyed to the research students by the researcher. Individual written response questionnaires were also used subsequent to the focus group interviews to override the influence of peer group pressure (see section 3.7.2).

Focus group interviews typically vary from a structured interview format (Mishler, 1986). Rather than asking narrow focussed questions, focus group interviews aim to encourage participants to talk around the research issue by asking open-ended questions that allow participants to engage in a 'narrative', relating their experiences

to the research questions. Mishler refers to such narratives as the most internally consistent understanding and interpretation of past and present experiences. Indeed, one of the primary ways that people make sense of their experiences is to cast them in a narrative form.

Mishler (1986) further argues that because we all live out narratives in our own lives, and because we understand our own lives in terms of these narratives, a narrative form is appropriate for understanding the actions and perceptions of others. To encourage the open discourse of narratives this project adopted a semi-structured interview format that allowed participants to control the flow and introduction of topics and encouraged them to extend their responses. In this respect the interview was more akin to a conversation or discussion between the researcher and the participants (Anderson, 1990).

A possible limitation of semi-structured interviews is that the researcher may have less control over the data that are generated. Accordingly, particular attention was given to chairing the group without being too obtrusive. This helped to ensure that the group continued to move forward through the list of desired topics while still allowing the participants to freely and openly discuss issues of concern to them.

From a sociocultural perspective, focus group methodology has a degree of external validity because focus groups are grounded in the human tendency to discuss issues and ideas in groups (Albrecht, et al., 1993). Opinions about a variety of issues are generally determined not by individual information gathering and deliberation but through communication with others. “Personal opinion might be more appropriately described as derived from social rather than personal processes” (Albrecht, et al., 1993, p. 54). Albrecht et al. also comment that one of the advantages of opinions gathered through focus groups, rather than individual interviews, is the degree of isomorphism between group opinions and those of the population at large.

Focus group interviews can offer advantages over individual interviews. Vaughn, Shay Schumm, and Sinaugub (1996) argue that when exploring perceptions, experiences, and beliefs, focus group interviews allow a greater degree of anonymity for participants. The participants have the security of being among others who share

many of their feelings and experiences, and this gives them a basis for sharing their views and perceptions. Unlike most structured interviews or surveys, focus group interviews provide a 'loosening effect' promoting candour and participation that allows students the opportunity to clarify or extend their point of view, and to provide examples to support this view. Individual participants have the freedom to choose whether they want to respond to individual questions or not. Focus group interviews, when done well, generate data that "would not come out in either the participants' own casual conversation or in response to the researcher's preconceived questions" (Morgan, 1988, p. 21).

In summary, focus group interviews are compatible with the major assumptions of the qualitative research paradigm as discussed earlier. Firstly, multiple views of reality exist. This is a strength of focus group interviews since individuals' diverse opinions and perspectives are desired and encouraged. Secondly, there is the potential influence of the 'inquirer and respondent relationship': "The interaction between the moderator and the respondents, and the interaction between the respondents themselves are recognised as having the potential to add depth and dimension to the knowledge gained" (Vaughn et al., 1996, p. 16). Lastly, the nature of truth statements is such that truth is influenced by perspective and by context.

3.7.2 Individual written response questionnaires

To complement and triangulate the data generated from focus group interviews, a series of short individual written response questions were used. These provided participants with the opportunity to comment further on any of the questions or concepts touched upon in the focus group interview. It also allowed participants to write down any views, or perceptions, that they did not feel comfortable expressing in the focus group interview, did not express because the discussion moved away from that point and the opportunity was lost, or did not think of until after the focus group interview had finished.

3.7.3 Classroom observational data

Classroom observational data are a valuable addition to case study research. Observations make it possible to record behaviours and actions as they happen, and

provide a firsthand account of the phenomenon of interest—albeit from the researcher's perspective in this project—rather than a second hand account obtained from interviews. With respect to this research, classroom observations provided an opportunity for me, as the researcher, to develop a contextual awareness of the classroom culture. It helped to develop partial understandings of the social structures and norms within the classroom, the nature of the interactions between the students and their teacher, and between the students themselves. This enculturation process was important in aiding a valid analysis of the focus group, and other data. It provided a lens through which the data was viewed with a more informed eye. Observational data were also used to inform and guide future data generation. In particular, observations of specific incidents, actions, or behaviours were used as reference points for discussion in subsequent focus group interviews.

As the researcher, I was aware of the need for classroom observations to be systematically planned and conducted (Cohen et al., 2000). An important component of the observational process was the taking of comprehensive field notes. In addition to the factual description of events, these notes included observer commentary including the researcher's feelings, reactions, hunches, initial interpretations and working hypotheses (Merriam, 1998). In reality, field notes also record informal interviews and conversations which are interwoven with observational data.

3.7.4 Assessment Experience Questionnaire

The Assessment Experience Questionnaire (AEQ) (see Appendix I) was developed by Gibbs and Simpson (2003) at the Open University, primarily as a diagnostic tool to help teachers identify how well the assessment in their course supports students' learning. Equally it can add to our understanding of students' perceptions of assessment experiences related to NCEA, within their specific course of study. Gibbs and Simpson state that their philosophical starting point was that assessment is seen as having a profound influence on students' learning. It affects "what students focus their attention on, on how much they study, on their quality of engagement with learning tasks, and, through feedback, on their understanding and future learning" (p. 2).

The development of the AEQ was based on a review of the literature of ways in which assessment influences students' learning, and a series of open-ended interviews conducted with students involved in science courses taught by the Open University (Gibbs & Simpson, 2004). Subsequently, a 36-item questionnaire, organised into six sections with six items in each section, was developed. The sections were organised around the grouping of the conditions for supporting students' learning and included specific sections for assignments (section 2 on the AEQ) and examinations (section 6 on the AEQ). Each of the 36 items uses a five point Likert scale to examine the assessment experiences of students. The items are designed so that higher scores represent a view of assessment that is consistent with assessment for learning as detailed in chapter 2. A number of the items use reverse scoring and are identified with an asterisk.

The prototype version of the AEQ was trialled by Gibbs and Simpson (2003) with 731 students. A Principle Component Analysis with Varimax Rotation and Kaiser Normalisation was conducted resulting in six different factors being identified, accounting for 50% of the variance.

Factor 1: Quality of feedback: A high score indicates that students perceive that they get plenty of timely feedback. The feedback is understandable and useful, explaining grades, misunderstandings and improvements. A low score indicates that students perceive the feedback to be insufficient to support their learning and too late to be useful, emphasising social comparison rather than personal progress.

Factor 2: Use of feedback: A high score indicates that students use the feedback to guide follow-up learning, to tackle subsequent assignments differently, and to revise. A low score indicates that the feedback has little impact on subsequent studying and learning.

Factor 3: Focus on assignments: A high score indicates that the students see assignment requirements as both challenging and requiring understanding. A low score indicates that assessment demands are perceived as unchallenging and as not requiring understanding.

Factor 4: Learning from the exam: A high score indicates that the perceived exam demands had a positive influence on the quality of learning undertaken for the exam, and that the exam itself was a positive experience. A low score indicates that students' perceive that they did not learn anything new for the exam, and do not have a better, or more thorough understanding after the exam.

Factor 5: Distribution of effort: A high score indicates that students distribute their study effort evenly through the course and across topics, and feel this is necessary to do well. A low score indicates that study effort is allocated narrowly to assessed topics, and concentrated in those weeks where assessment takes place. Students feel that they can get away with this and still do well.

Factor 6: Approach to the exam: A high score indicates that students feel that a sound understanding is necessary to secure high marks in the exam. A low score indicates that the perceived exam demands encourage memorisation and subsequent forgetting of course material

In this study, the AEQ was administered to all students in the Y12 Trad Maths cohort.

3.8 Ethical considerations

As with all educational research studies involving human participants, ethical concerns regarding: informing and selecting participants, maintaining confidentiality and anonymity, collecting of data, and disseminating of results are central. For research to be considered comprehensive and valuable it is imperative that an ethically sound relationship be established between the researcher, the teachers, the students, and their parents.

3.8.1 Informed consent

This relationship starts with informed consent, a fundamental to ethical research (Cardno, 2003). To facilitate informed consent, potential participant teachers and students were provided with full information about the aims of the study, the role they would be expected to take, their commitments to the research process, and the procedures designed to minimise harm. A central tenet of informed consent was an understanding that their participation was voluntary and independent of any school

assessment procedures or appraisal systems. This information was provided in a form designed to be easily comprehensible to the respective participants. Separate information sheets were prepared for the teachers and the students. Potential participants were also given the opportunity to discuss the details of the research and ask any relevant questions. Copies of the relevant documentation are included in the appendices (see Appendices A, B, C, & D).

3.8.2 Confidentiality and anonymity

Assurances of confidentiality and anonymity are important ethical considerations when conducting research. This is especially true for students, who must be made to feel that they can be open and honest with no fear of retribution. However, when any research is conducted it must be recognised that there is always a risk of a breach of confidentiality, and that the researcher can only give an assurance of confidentiality and anonymity to the extent allowed by law. It is always possible that one of the participants will pass on confidential information to a party outside the research process. To mitigate against this risk, a culture of confidentiality was established amongst those taking part in the current research. This involved establishing a clear expectation that all participants, including the researcher, would respect any information shared through the research process and would treat it with confidentiality.

Establishing a culture of confidentiality can be problematic. The utilisation of focus groups meant that it was inevitable that those involved in the focus group interviews would have their identities revealed to the other participants of the focus group. Indeed, maintaining the anonymity of one student focus group from the other was difficult to achieve. Therefore research participants were continually informed of the clear expectations regarding confidentiality and anonymity.

Procedures for protecting the anonymity of the students from the teachers were also developed and are discussed in more detail in the following sections. Additionally, neither the school nor any individuals are identified either directly or indirectly in verbal or written form. Where direct quotes from the interview tapes or written

correspondence are used in subsequent publications, pseudonyms are assigned to maintain anonymity.

3.9 Methodology in action

This section details the methods that were used to undertake the current research project. It initially describes the case study school and the reasons for choosing this particular school. It then discusses the procedures used to select the case study students. The data generating tools are explained, including relevant issues associated with their use. The section concludes with a discussion of the relevant ethical issues associated with this research.

3.9.1 The setting

Euclid College (pseudonym) is a medium to large (~1300 students), decile 7[5] urban secondary school in a provincial city. One of the key factors influencing the decision to approach Euclid College was its pro-active stance towards SBA and NCEA. The college had been involved in the early trials of Unit Standards during the mid 1990s and had subsequently replaced their Sixth Form Certificate courses solely with Unit Standards based courses in 1997. These experiences increase the likelihood that the school has begun to develop pedagogical and organisational structures consistent with the philosophical underpinnings of NCEA. Additionally, the school espouses a strong ethos of developing student self autonomy with "students encouraged to take responsibility for their learning" (school website promotion). However it should also be recognised that choosing a school with a strong history of support for SBA may provide an atypical set of data. As such, this project seeks to investigate students' perceptions from within the singularity described, and does not claim that these perceptions are indicative of all students' perceptions from all schools. It is up to the

[5] The Ministry of Education uses a decile rating system for school funding purposes. Each decile contains approximately 10% of schools. Schools in decile 1 have the highest proportion of students from low socio-economic backgrounds. Schools in decile 10 have the lowest proportions of students from low socio-economic backgrounds.

reader to decide the extent to which the findings of this project can be applied to other educational settings.

Prior to the study it was acknowledged that the research may find areas of the school's assessment practice that need improvement. It was therefore deemed to be important that the selected school be supportive of the research aims, and aware of any possible implications of the research findings. Initial approaches to the Head of Department of Mathematics (HOD) confirmed the school's proactive stance towards being involved in educational research. The HOD expressed a strong desire to improve student autonomy in learning and felt that the project would provide a good opportunity for the mathematics department, and the wider school, to examine particular aspects of its assessment practices, namely the effect of teacher feedback practices on students' learning. He put it succinctly by saying that: "if there is something that we can do better, we want to know".

The school was subsequently approached officially, and agreed for a pilot study to be conducted in 2004. The pilot study was undertaken with one teacher in their Y11 mathematics class. Data generation was via a focus group interview, class observations and an interview with the teacher. A Y11 class was chosen as it is the first year most students are involved in NCEA. It was hoped this would provide an opportunity to examine students' formative use of high stakes assessment during the development of these dispositions. Analysis of the pilot study data revealed that students' understandings of the potentially formative features of assessment events were tentative and underdeveloped. To more fully address the research questions it was subsequently decided to conduct the main part of the research project with Y12 students. This decision was based on the argument that Y12 students would have had a year to develop their formative practices, and this year group would contain a wider cross section of ability levels than would be expected in Y13.

The 'Traditional Maths' (Trad Maths) course was selected as the case study course from the three Y12 mathematics courses offered at the school. This was mainly for pragmatic reasons; it was the course that offered the largest number of students from which to select the case study participants. Additionally the students in this course have all the requirements, identified by Morse (1994), necessary for being good

informants; namely that they have the knowledge and experience that the researcher requires, the ability to reflect on and articulate their experiences, and the time and willingness to be interviewed.

The results from the pilot study were shared with the school and were positively received. The school was interested to continue with the main research project and official confirmation was sought and given.

3.9.2 The case study participants

At the start of 2005 the teachers of the three Y12 Trad Maths classes were approached to take part in the research. The teachers expressed a genuine interest in being involved, and this interest was maintained throughout the research. Each of the Trad Maths classes were visited to explain the research project, provide them with the information sheet and consent form (See Appendices A & B), and answer any questions.

The information sheet invited students to be involved in the focus group interviews and follow up written questionnaires throughout the year (See Appendix H). Students were provided with blank envelopes and asked to return their signed consent forms to the main school office. Bypassing the teacher from the consent process helped to ensure the confidentiality of the student participants.

Ten students across the three classes agreed to take part in the research; one student subsequently failed to complete any of the data generating phases and another female student completed only one phase. The remaining eight case study students consisted of six females and two males, a gender balance that reflects the gender ratio within the cohort of Y12 students. Ideally a larger sample of students would have been preferred as this would have provided a wider range of experiences and perceptions from which to draw conclusions. As such, the small sample size is recognised as a limitation of this research.

A potential limitation of any self selecting sample is that it may attract students who are more interested in either the research process, or the research questions. With

respect to this project, students electing to take part in the research may hold stronger views about the formative potential of NCEA than those held by the wider Y12 Trad Maths cohort. A self selecting sample may also attract the more able students who may be more comfortable having their achievement and assessment levels examined in more depth. An analysis of the end-of-year results of the nine case study students (See Appendix L) reveals that the research students can be characterised as moderate to high achievers, with all students passing more than 2/3 of the standards they were entered for. Four of the students passed all of the standards they were entered for, two of these with very good grades. Anthony's (1994) study of Y12 students' mathematics learning strategies found that high achieving students typically use more effective strategies, especially with regard to assessment. As such the limitation of the sample is noted as it restricts the generalisations that can be made from the current research. Accordingly, subsequent follow-up research will need to consider a wider sample of students covering a broader range of ability levels.

To mitigate the limitations identified above and to examine the representative nature of the sample, the AEQ (Gibbs & Simpson, 2003) was administered to the entire cohort. The results of this procedure are discussed in more depth in the results chapter, but essentially, analysis revealed no significant variance between the expressed views of the case study students and the wider cohort of student taking Y12 Trad Maths.

One of the data generation procedures involved whole-class observations of students. Accordingly, students were offered the opportunity in writing to decline to have observational data recorded for them. As part of the University ethics application it was determined that a nil return would imply consent for this part of the research procedure. No students declined to have observational data recorded for them.

3.9.3 Communication with students

The initial intention was that communication with the students would primarily be via email, although surface mail was requested by one student. In recognition of the potential concern that guardians may have about someone contacting their teenagers via the internet, copies of all correspondence were sent to guardians. As it transpired,

communication via email was not without its problems. As the researcher, I had initially thought that the students might appreciate the use of technology, and be more likely to respond to written questionnaires emailed to them. However, many of the students were inconsistent about checking their email boxes, and, on several occasions, emails bounced back because the student's 'inbox' was full. Additionally, when two of the guardians saw the initial student questionnaire, there was some confusion about whether they should also fill out the questionnaire.

In an effort to streamline communication, students agreed to the use of 'text' messages to arrange meetings, or inform them that an email had been sent. This proved to be a highly successful strategy for several reasons. Firstly, texting to multiple recipients provided a quick and convenient way of organising, and subsequently reminding students of, meetings times and venues. Secondly, by using what might be considered a 'young person's' technology, students responded in a very casual and open way which helped to bridge any potential barriers due to the age difference between the students and myself. Additionally, students were able to inform me if they were ill or had been delayed on the day of our meeting. A potential improvement to this strategy could be to ask guardians if they would like to be sent copies of the texts so that they could monitor the communication between their son/daughter and myself.

3.9.4 Focus group interviews

The nine case study students were divided into two convenient groups of four and five students each. One group consisted of four students from one of the classes, the other group consisted of two students from one class and three from the remaining class. A series of three pairs of focus group interviews were conducted through the year; the first pair towards the end of term one after the students had completed their first assessment for the course. The second pair of focus group interviews were designed to generate data midway through the year, and were completed towards the end of the term two. Students sat their mock exams in the last week of the term three , and so the last set of interviews were undertaken early in term four, once students had received their assessed work. Students generally attended their scheduled focus group interview. On one occasion, one of the students was ill during her scheduled

interview, but was able to attend the paired interview with the other participants' consent.

The focus groups followed a semi-structured interview format with a series of prompt questions used to guide the discussions (see Appendix G). At the start of each focus group interview students were encouraged to talk openly, either directly to the question or to other related issues. It was also stressed that there were no right or wrong answers. It was reaffirmed that the school was very interested in their open and honest views. They appeared to appreciate that their views were being sought, and enthusiastically took part in the interview process. It was also stressed that the individual views expressed during the focus group interviews should remain confidential to the participants. Special effort was made to create an atmosphere in which each person felt free to share his or her point of view. This was achieved by emphasising the desire for a range of different experiences and feelings to be heard. All members of the group were encouraged to participate and, at times, individuals were asked if they had anything to contribute.

The focus group interviews were conducted at times convenient to the students, usually either lunchtime or directly after school. The school had a suitable interview room away from the general staff area which helped to ensure anonymity for participants. Each of the interviews lasted for approximately 45 minutes to one hour and was audio recorded.

3.9.5 Individual written questionnaires

Following each focus group interview a set of questions was emailed to the students. These questions required short written answers, and were designed to be consistent with the prompt questions used in the focus group interview (see Appendix I).

The return rate of the emailed questionnaires for the first term was 100%. The return rate dropped off, however, during the year and it is not known whether this was because students felt that they had nothing to add to their interview data, or if there were other reasons.

3.9.6 Class observations

To provide a backdrop against which the focus group students' comments could be reviewed, a series of non-participant classroom observations were made. Each of the mathematics class was observed three times in term one. A further three observations were conducted in term two, with a final three being conducted bridging the mock exams in terms three and four. For each set of class observations the first observation was conducted during the typical teaching of the unit of work, the second during the period immediately prior to the students sitting the assessment task, and the final observation when the students received their assessed work back. Written field notes were taken on a 'Palm Pilot' during the observations and edited immediately afterwards. This provided an immediate electronic record of the observations.

In addition to providing background information, the observations provided opportunities for informal discussions with the class teachers and, in some instances, the research students at their instigation. Classroom observations were valued as they helped to familiarise the researcher with the culture of the Y12 mathematics classrooms, as well as helping the students to become comfortable having the researcher around.

3.9.7 Assessment experience questionnaire

The AEQ, developed by Gibbs and Simpson (2003), was administered to all three mathematics classes in term four. The purpose of this aspect of the data generating procedure was twofold. Firstly, it provided a quantitative tool to gauge the extent to which the case study students were representative of the wider cohort. Secondly, the AEQ provided some valuable data about the wider cohort's perceptions of their assessment experiences in Y12 Trad Maths.

The AEQ was not added to the data generating procedures until midway through the year and so was administered only once. A suggested improvement in the methodological design would be to administer the AEQ at the start and end of the data generating phase of the research. Since the AEQ had not been part of the original project design, additional approval from both the University ethics committee, the

school board of trustees, and the students was sought, and granted (see Appendices E & F).

Each class was visited in turn and the purpose of the AEQ and each student's right to not participate was explained. Students were given 15 minutes to complete the questionnaire and could either hand it back at the end of the 15 minutes or take it away to complete and hand in to the main school office. The case study students had been asked, via a text message, to write their name on the bottom of the AEQ so that they could be separated from the rest of the cohort. Students who elected not to complete the questionnaire were given alternative study material by the teacher. The response rate for the three Trad Maths classes was 78%.

3.9.8 Teacher focus group interview

A one hour focus group interview with the three teachers of the Trad Maths classes was conducted in term three. The purpose of this interview was to provide an insight into the assessment culture of the three classrooms and to examine the teachers' perceptions of the role of assessment in students' learning. Additionally it provided a backdrop against which the comments of the students could be situated. The focus group interview adopted a semi-structured format (see Appendix G) and was conducted at a convenient time for the teachers. An individual interview with the HOD was also conducted using a semi-structured format. This interview, conducted in term one, provided background data on the school, the Trad Maths classes and the relevant assessment policies.

The teachers maintained a high level of interest throughout the project and relevant field notes from additional informal conversations with the teachers were also taken. Towards the end of the year the teachers asked me to come in to informally discuss the provisional findings of the research. This session reinforced the teachers' commitment to the research process and their genuine interest in improving their assessment practices.

3.9.9 Student assessed work

Throughout the year student assessed work was stored in individual folders in each teacher's classroom. I had negotiated with the students and the school to have access to these folders and arranged a time when I could temporarily remove all the folders to maintain the confidentiality of the research participants. These folders were examined midway through the year and at the end of the year. Copies of examples of assessment feedback were electronically scanned and appear in the results chapter. These provide triangulated evidence of the feedback practices conducted by the teachers, as commented on by the students.

It is acknowledged that participation in this research study might have led the participant teachers to change their feedback practices during the course of the year. To mitigate against this possibility this concern was raised with the teachers at the start of the research process, and a review of the students' folders revealed no notable changes to feedback practices over the timeframe of the research.

3.10 Data Analysis

This research project generated both quantitative and qualitative data. The quantitative data were analysed using a statistical package (Statistical Package for the Social Sciences, or SPSS) with relevant statistical procedures applied. These results are both reported separately and incorporated into the qualitative data analysis. The qualitative data for this project were generated from a number of sources as detailed in the 'methodology in action' section of this chapter. Although class observations and an examination of students' assessed work were conducted, the main sources of qualitative data were in the focus group interviews and individual written response questionnaires. The narrative data generated in these last two methods are evidence of the participants' perceptions of their experiences and are consistent with the sociological tradition which treats narrative as a window into human experience (Ryan & Bernard, 2003). Consistent with an interpretivist view of reality, analysis of narrative treats participants' perceptions as contextually situated accounts of reality, thereby generating plausible accounts of students' experiences (Silverman, 2003). The data are analysed through aggregation of instances until common themes start to emerge (Stake, 1995).

This project analysed the data using a four stage framework adapted from the work of Vaughn et al. (1996). It should be noted that this framework is applied in a cyclic rather than a linear form. The aim is to learn from the data, by revisiting it until a clear understanding of the patterns and explanations emerges. In order to establish a comprehensive analysis of the data, an iterative process through the various data analysis stages was followed.

Stage 1: Reviewing the data

Shortly after each pair of student focus group interviews, the digital audio recording of the interviews were imported into a computer software program called Annotape. This program allowed the researcher to create 'sound bites' that were coded for further analysis and annotations. These audio recordings were listened to a number of times, with special attention paid to the participants' words, views, frequency of expressed ideas, and intensity of responses. One of the strengths of Annotape, over

typed transcripts, is that it maintained the relevant section as an audio recording, allowing the timbre and emotion of the students' voices to be captured. This helped to develop an awareness of the extent of feeling towards a particular theme, and whether it was important to a high proportion of the participants, or just a few. Additional data were also considered from the students' written response questionnaires and the classroom observations. The process of reviewing the data helped the researcher to gain an understanding of the social norms within the mathematics classroom. The review of the data also helped to focus subsequent stages of the research; the data from the first and second term informing the nature of investigations in later terms. This review process was repeated numerous times during the analysis stage of the research project in order to extract maximum meaning from the data.

Stage 2: Developing coding categories

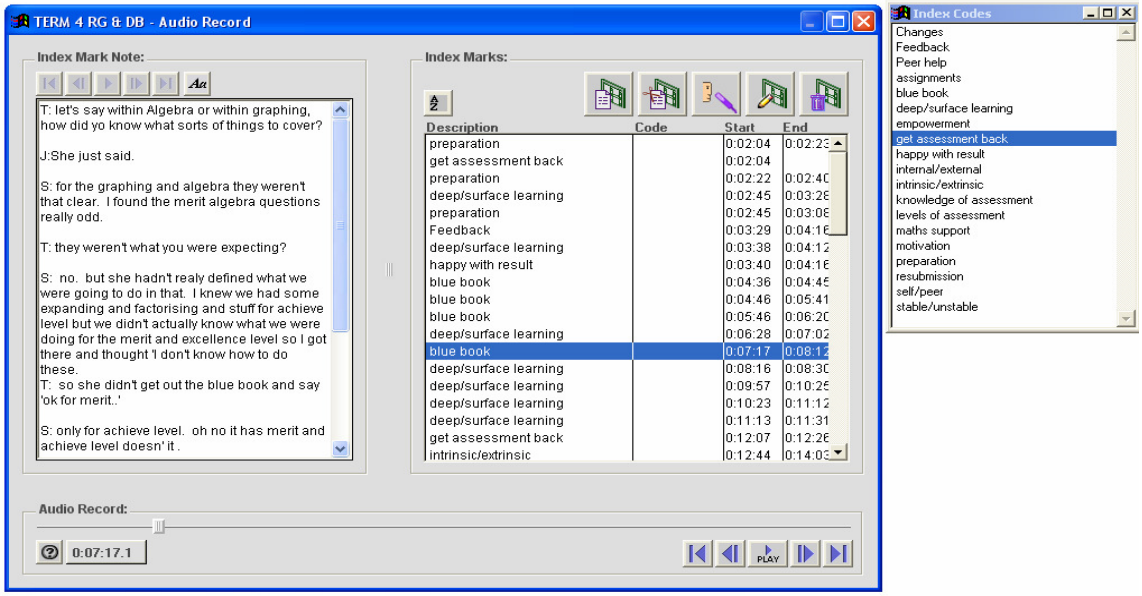
During the review stage, a set of preliminary codes were developed to aid the analysis of the data using a multi-step process identified by Bogdan and Biklen (2007). Initially the data were trawled for regularities and patterns as well as for topics. Words and phrases which represent these topics and patterns were then identified; these formed the coding category labels. These coding categories allowed for sorting of the data so that the material bearing on a given topic could be separated from other data. The labels of these coding categories were considered fluid during the data collection phase of the research project. Bogden and Biklen identified the development of a list of coding categories as a crucial step in data analysis.

Stage 3: Coding of data

Once the data generation phase of the project was completed, the data on the digital tapes were coded using Annotape. As the digital recording was listened to, index marks were added to the recording and codes allocated. In some instances, particular sections of the recording were given multiple codes. Once the coding process had been completed the recordings were then sorted, so all instances of a particular coding category could be listened to sequentially. This categorisation of common material aided in the recognition and development of emerging themes.

An example of the working screen and index codes is given below. At the top of the screen is the code identifying the setting of the particular focus group interview. The right hand box contains the index codes used. The centre box contains a time log of the audio recording with the index codes attached. The left hand box allows the researcher to transcribe particular sections of the audio recording and to make annotated notes. This procedure was followed for all of the digital recordings. Individual questionnaires were also coded using the same coding categories and stored in a 'word' table.

Figure 3.1 Annotape working screen



Stage 4: Identifying and reporting of themes

Once the coding had been completed the data were repeatedly examined to identify the emerging themes. These themes provided an organisational structure for summarising the data and writing up the findings. Summaries and explanations of these themes were written with supporting units of data (in the form of direct quotes and field notes) used to provide the reader with a context for the summaries. These results then form the basis of the discussion chapter of the research project.

3.11 Summary.

This chapter has detailed the justification for, and use of, case study as a methodology for the current research project. This research is informed by the belief that students' development of the formative use of assessment information is contextually and socially situated. Accordingly it was argued that a case study was the preferred methodology, as it provides a means of investigating students' perceptions from within the complex contextual and social setting of the mathematics classroom. Additionally this research adopts an interpretivist perspective of reality. This perspective informed the choice of data generating methods, particularly the use of focus group interviews to investigate students' conceptions of the 'realities' they create. This chapter also detailed the range of data generating methods used to provide a contextual grounding for the students' perceptions. Lastly, the method of organising and analysing the data was outlined. The results of this analysis are examined in the next chapter.

4 RESULTS

4.1 Introduction

This chapter presents the results from the current study divided into five sections. The results are presented here with limited discussion; a fuller discussion is provided in chapter 5. The first section (4.2) provides a brief general description of the structure of the Y12 mathematics classes involved in this study. The purpose of this section is to situate the reader into the current study and provide a lens through which to view the data. Section 4.3 presents the quantitative data from the Assessment Experience Questionnaire (AEQ). This questionnaire was primarily used to establish whether the perceptions of the research students were representative of the wider school cohort of Y12 mathematics students, therefore establishing the trustworthiness of the data presented here.

The final three sections of the chapter organise the data in accord with the three conditions of Sadler's (1989) theory of formative assessment. Section 4.4 examines the data concerning the development of a shared understanding of the assessment criteria students are working towards. Section 4.5 reports on the written and oral feedback received by students on assessed work. Section 4.6 examines students' engagement with the feedback given. The findings presented in these three sections include data from the student focus group interviews and the written questionnaires completed subsequent to the focus group interviews. The findings also include data from the focus group interview and additional casual conversations with the research teachers. Within each of the sections the analyses are supported by quotes derived from a variety of sources. The quotes selected to illustrate a given observation or finding are representative of a common viewpoint expressed by the participants. The degree of conformity to this viewpoint is commented on in the accompanying text. During the research process, a number of students showed themselves to be more articulate at expressing a commonly held view. This can be observed in the relative proportions of quotes from the research students cited in this chapter. However, one of the advantages of using Annotape to code the digital recording directly was that the

researcher can gauge the level of agreement with a particular viewpoint by the background discussions and expressions that may not be transcribed by normal methods. Table 4.1 below explains of the notation used to identify the sources of the quotes.

For ease of reference the nine students have been allocated the pseudonyms of Tracy, Sally, Kate, Raewyn, Hannah, Megan, Chelsea, Andrew, and James. The three teachers in the study have been called Ms Brown, Ms Clarke and Mr Smith.

Table 4.1 Source of quotes

Example of Quote	*Source*
And they probably wouldn't try them so if they got one of the Achieve ones wrong they couldn't use these other ones to fill in. (Tracy, FG2)	Verbal comment made by the student Tracy during Focus Group interview 2 held during the 2nd term.
If there was an example question to go with each requirement, but that would be a bit difficult. (Andrew, Q1)	Written comment made by the student Andrew in the written questionnaire emailed out subsequent to the first focus group interview.
Ms Clarke: I find I am writing the same comments. Ms Brown: Same here. Ms Clarke: So you may as well talk to them as a class.	Example of a conversation during one of the focus group interviews.

4.2 The Y12 mathematics classroom

In order to help situate the reader, a brief description of the data from the classroom observations is given in this section. This general description is derived from the

researcher's field notes from the classroom observations of the three mathematics classes.

The three classes observed exhibited similarities with respect to pedagogy, student activities, class size, and students' achievement levels. It is not intended that the description presented here portrays any one individual teacher's sequence of lessons. Rather the description is intended to be representative of the three classrooms, and provides the reader with a sense of the structure of the lessons, and the roles and actions of the participants. As outlined in the methodology chapter (section 3.9.6), three sets of three observations (triplets) were made of each mathematics class; the first triplet of observations in the first term, the second set in term two, and the third set bridging the mock examinations at the end of term three and the start of term four.

4.2.1 A typical teaching session

The first classroom observation of each triplet was conducted during the teaching of the mathematics unit. The purpose of this observation was to gain a sense of the structure of a typical mathematics period, its learning activities and the inter-relationship between the students and their teacher.

All three classrooms share a similar organisational layout. The desks were arranged in rows facing the whiteboard at the front of the room. At the start of the period students informally entered the room, and proceeded to find a seat and get their books and other material out. In all of the classes observed the teacher was always in the room at the start of the period and informally interacted with the students as they came in. The teachers all appeared to have good relationships with their students, and the atmosphere was lively and typically good-natured.

The structure of the lessons followed a consistent pattern within all three classes. The mathematics period started with the teacher reminding students of organisational details of assignments, assessments and 'Support Maths' sessions. A small amount of time was then set aside for any questions concerning the homework, although very few students took the opportunity to ask questions. The teacher would then provide an introduction to the day's lesson, supported by appropriate notes written on the

whiteboard, which the students copied down into their note books. During the 'teaching' of the material the teacher would generally ask closed questions, either to specific individuals or to the whole class. These episodes of teacher centred activity lasted between 10 and 20 minutes. After the introduction, students were usually requested to engage with a set of exercises from the issued textbook or a worksheet supplied by the teacher. Most of these activities were completed individually, although students generally asked their peers or the teacher for help when needed. During this time the teacher circulated the room checking student progress and answering questions. Occasionally, the teacher would move to the whiteboard to clarify a general concern to a larger group of students, allowing others to continue working. Students typically worked on the set activities until the teacher issued homework and clarified organisational details immediately prior to the end of the mathematics period.

4.2.2 Prior to the assessment

The second of each triplet of classroom observation was conducted during the mathematics period immediately prior to the students sitting the unit assessment. In most instances this period acted as a revision session in which the teacher provided a synopsis of the mathematical content of the unit and addressed any students' concerns. During the revision session, the teacher's emphasis was normally on ensuring that students understood the mathematical content rather than specifying the technical requirements to pass the Standard (e.g., you need to get 4 out of the 5 Achieve questions right). The teacher typically emphasised the types and levels of mathematical problem that the students might reasonably expect in the assessment. For example, in the observation of Mr Smith's class during the first term algebra unit, he emphasised that students could reasonably expect to solve an expression such as $x^2+8x=400$ to achieve at the Merit level. Notably, in this instance he gave an example of a specific mathematical problem but did not highlight the Merit level assessment criteria that this problem is an example of, namely that it requires the use of the quadratic formula.

Observations confirmed teachers' expressed goals that the mathematical content they present in these revision sessions is designed to maximise the pass rate for the class,

rather than maximising individual student's grades. There was an emphasis on mathematical problems at the Achieve and Merit levels, rather than the Excellence level. During these revision sessions the teacher would often hand out revision material for the students to use as a study resource, complementing any other material that may have been handed out at an earlier stage. In some instances, due to time constraints, the teachers needed to complete teaching the course material during this period, and so could not offer an in-class revision period.

4.2.3 Return of students' scripts

The third classroom observation in each triplet was conducted when the assessed work was returned to the students. Work was usually returned within a week of the assessment having been sat by the students but, at times, this could be delayed due to school moderation requirements. The structure of the class period when the assessed work was returned was influenced by whether any students were required to complete a 'resubmission' of their assessed work. Resubmission[6] opportunities are designed to allow students who are close to securing a higher grade the opportunity to find and correct errors in their work. The original script was returned to the student with no marking or feedback on the particular question under consideration. The student was then given a fixed time period, usually 10 or 15 minutes, to find and correct any errors. Whether or not a student was offered a resubmission opportunity was at the professional discretion of the teacher. During this time, the other students were engaged in set work for the new mathematics unit.

If any students needed to undertake a resubmission then this was done in the normal classroom, typically during the first 15 minutes of the period. If no students were required to undertake a resubmission, or after these had been completed, all assessed work was returned. Work was returned on an individual basis with no grades read aloud. In some instances the teacher made private individual comments to students who had done well, or had narrowly missed securing the next grade.

---

[6] A resubmission re-assesses part of the student's original assessment task. A 'resit' is a second opportunity for the student to be assessed against the entire Standard on a new assessment task.

Once the scripts had been returned the teacher then provided oral whole-class feedback. This feedback usually started shortly after the last script had been returned, with students being given limited time to look through their scripts. The teacher typically highlighted common mistakes made by the student cohort on selected questions. At times the teacher provided a partial, or full, model answer, at other times, the teacher provided just the final answer to the mathematical problem. The teachers later confirmed the researcher's observation that the purpose of this feedback session was to help students understand errors, rather than check the teacher's marking.

Following the teacher directed review of questions, the teachers then invited students to ask questions about specific concerns that had not been addressed in either the written or whole-class oral feedback. Teachers used these student initiated concerns to either further direct the whole-class sessions, or interact with students individually or in small groups. Seeking clarification from the teacher was not common, with typically less than 3 or 4 students in a class asking questions during each of the observed sessions.

During these classroom sessions a lack of focus by some students was noted. Some students were clearly off-task and not looking at their returned work, while others, although focussed on their assessment scripts, were discussing them with their peers, rather than attending to the teacher's oral feedback. The nature of these student discussions was explored in the student focus group interviews and is addressed in a subsequent section of this results chapter (section 4.6.1).

Once the whole-class oral feedback session had been completed, and any questions answered, the assessment scripts were collected in, and stored in students' individual assessment folders held by the teacher. Assessment scripts are retained by the school for audit and aegrotat purposes, although students can ask for access to these folders during specified school times.

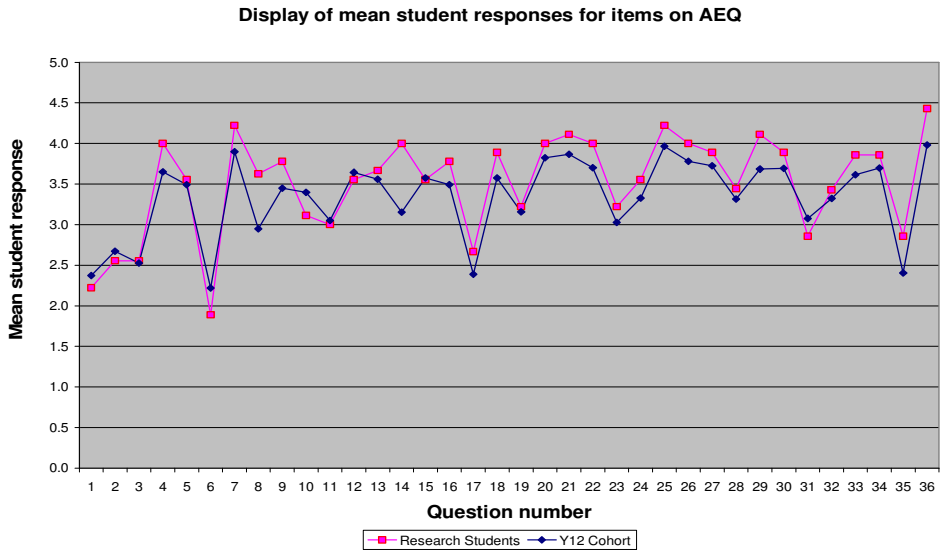
4.3 Assessment Experience Questionnaire.

As outlined in the methodology chapter, the AEQ (Gibbs & Simpson, 2003) was administered to all students (n=76) taking the Y12 Traditional Maths (Trad Maths) course. The AEQ served two purposes. In the first instance it provided a mechanism to examine whether the expressed perceptions of the research students were representative of the larger Year 12 Trad Maths cohort. In the second instance, it provided additional data to complement the data generated from the focus group interviews, questionnaires, and class observations.

4.3.1 Representativeness of research students

The questionnaire consisted of 36 items (I), each scored on a five point Likert scale showing the students' agreement or disagreement with the given item (see Appendix I). A number of the items are reverse scored. A high score is consistent with 'assessment for learning' principles (Gibbs & Simpson, 2003, see section 3.7.4). In addition to the nine research students, 50 of the 67 non-research students returned the questionnaire, giving an overall response rate of 78%. The following graph displays the mean item score for the research students (red) and the Y12 cohort (blue).

Figure 4.1 Display of mean student responses for items of AEQ



On most of the 36 items, the mean response of the research student (R) closely shadows the mean response of the Y12 cohort (Y12) (see Appendix J). The research students scored higher on 27 of the items, lower on 6 items, are the same as the Y12 cohort on 3 items. Where differences exist these are generally small with the exception of three items where the differences are statistically significant. On I14[7] (The feedback comes back very quickly: R, 4.0; Y12, 3.2) both groups agree with this item. On I35 (I'll probably forget most of it after the exam: R, 2.9; Y12, 2.4) both groups disagree with this item. On I8 (I learn more from doing the assignment than from doing the course material: R, 3.6; Y12, 2.9) the two groups bridge the neutral value of 3.0 although it should be noted that the Y12 score is not statistically distinct from the neutral value.

On average the research students scored 0.18 higher over the 36 items. Two interpretations for this difference can be suggested. Firstly, the self selecting sample of research students may be more inclined to pay attention to teachers' feedback and may be more likely to give a positive account of the use of formative assessment. Secondly, theses higher AEQ scores might have arisen from the research students' involvement in repeated questioning about formative assessment prior to the administration of the AEQ. However, with these differences noted, for the purposes of the current research the differences between the two groups are sufficiently minor to suggest that the research students can be considered a representative sample of the Y12 cohort.

4.3.2 Factor scores

Factor scores, in accord with the factors identified by Gibbs and Simpson (2003) were calculated for the 59 students in the Y12 cohort. A full listing of the individual items for each factor, and the mean response from the students, is detailed in table 4.3, along with an indication of whether the students generally agreed or disagreed with the individual items given. A discussion of the individual factor scores is conducted below.

---

[7] I14 refers to Item 14 on the AEQ.

The first factor relates to students' perceptions of the quality of feedback received. Strong responses indicated students believed the feedback they received showed them how to do better next time (I21, 3.9)[8], helped them to understand things better (I20, 3.8) and explained why they received the mark they did (I22, 3.7). Students also believed that they received plenty of feedback (I13, 3.6; I15, 3.6), and, on balance, this feedback came back quickly enough to be useful (I18, 3.6; I14, 3.2). One interesting response within the factor is that students indicated that they agreed with the item: "I would learn more if I received more feedback" (I17, 2.4). Whilst it should be noted that the students perceived that they received plenty of feedback, and they found the feedback valuable, taken together these responses could be interpreted as students emphasising the importance they placed on quality feedback. In essence, the more feedback they received, the better.

The second factor relates to how students used feedback. Students' responses to these items indicated that they actively engaged with the feedback given to them. They read feedback provided, tried to make sense of it (I25, 4.0) and were not just reading the grade (I30, 3.7). They used the feedback to go over what they had done in the assignment (I26, 3.8) and to help them in subsequent assignment work (I27, 3.7). Interestingly, students also stated that the feedback was useful for revision purposes (I29, 3.7) although this appears to be contrary to some of the data generated in the focus group interviews. This is discussed in more depth in section 4.6.2.

The third factor, Focus on Assignments, investigates the extent to which students were selective about what to study in relation to assignment demands. Whilst the 95% confidence interval for this factor score indicates an overall neutral response, responses to individual items were not neutral. For I2 (2.7) and I3 (2.5), responses indicated that students were selective about what they studied. For I4 (3.6) and I5 (3.5) students appeared to recognise that they needed to study regularly to do well in the course. Although, there appears to be a contradiction in these responses, it is conceded that it is possible for students to regularly study selected material and still receive good grades (see deep-achieving approach as described by Biggs, 1987).

---

[8] (I21, 3.9) indicates that item 21 received a mean score for all Y12 Trad Maths students of 3.9

Factor four investigates whether students saw the 'practice'[9] examination as a positive learning experience. AEQ responses suggest that students did learn from the examination. They agreed that it consolidated their understanding of content (I34, 3.7; I32, 3.3) and that it prompted them to learn new material, not previously mastered (I33, 3.6).

The distribution of effort during the course, specifically towards the assignments, is examined in factor five. The overall neutral value for this factor masks a variety of student responses to the individual items. Although students generally saw the assignments as challenging and thought provoking (I12, 3.6; I7, 3.9), they did not alter the amount of work conducted in weeks when assignments were due (I1, 2.4; I6, 2.2).

The last factor relates to students' approach to the 'practice' examination. Once again the overall neutral value for the factor score disguises an interesting pattern of students' responses. In I36 (4.0) students indicated that they perceived that a strong understanding of the course material was the key to securing high grades in the examination. In contrast to this, students felt that they would not retain the material subsequent to the examination (I2.4, 2.4). Combined these responses appear to indicate an ideal (I36) deep approach to learning and an actual (I24) surface approach to learning emphasising memorisation rather than understanding, especially for those students who did not get high results. Notably, students' responses to I31 (Preparing for the exam was mainly a matter of memorisation) were neutral.

The distributions of individual student scores for each factor are displayed figure 4.2. on the next page.

[9] The 'practice' examination, sat in term three consisted of: two practice attempts at External Achievement Standards sat in the end of year examinations, a first opportunity to sit an internal Unit Standard, two optional resit opportunities of internal Achievement and Unit Standards

Figure 4.2 Individual students' responses for factor scores

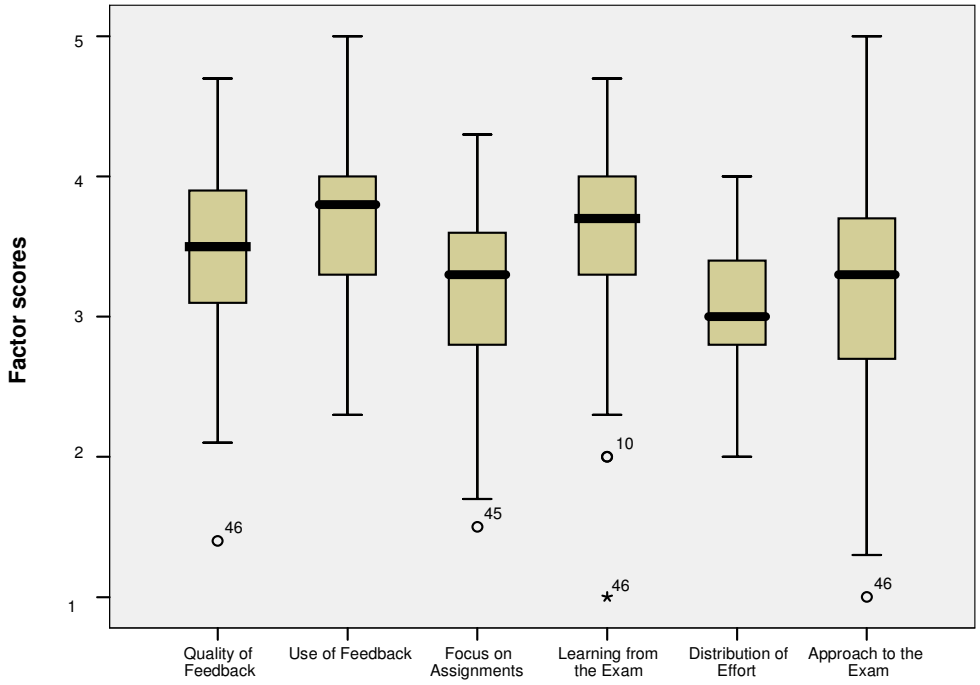


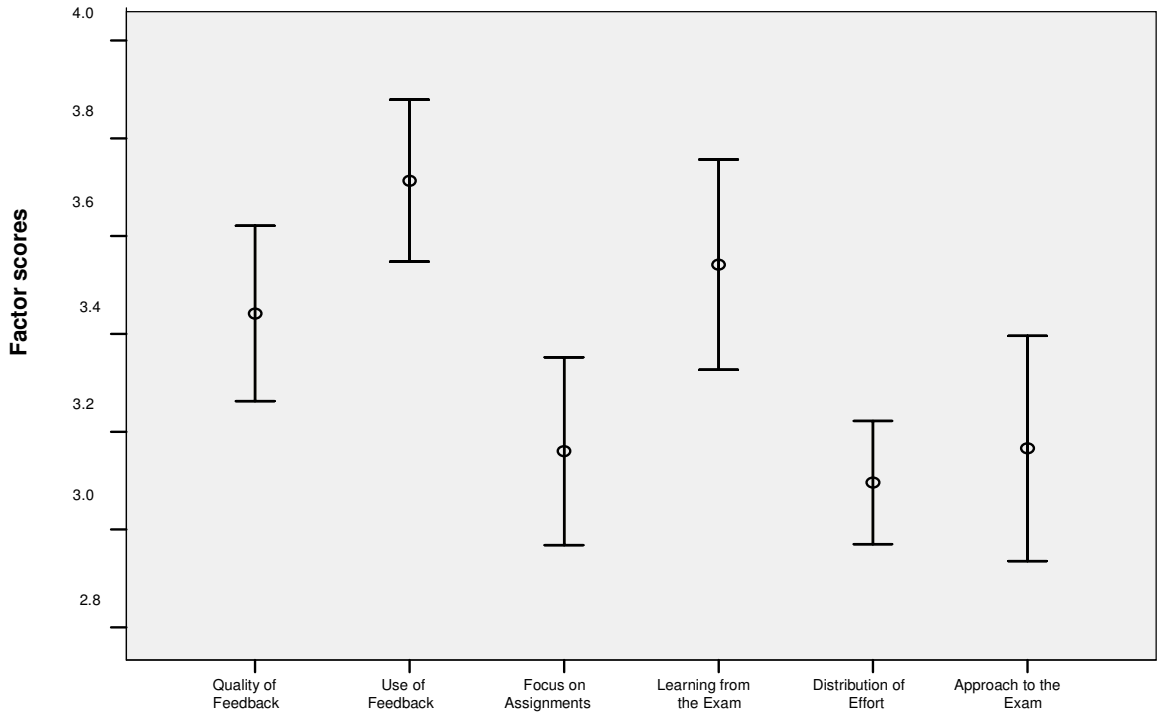
Table 4.2 Statistics for factor scores

	Quality of Feedback	*Use of Feedback*	*Focus on Assignments*	*Learning from the Exam*	*Distribution of Effort*	*Approach to the Exam*
Mean	3.41	3.69	3.09	3.54	3.07	3.17
Median	3.40	3.80	3.30	3.70	3.00	3.30
Standard Error of Mean	.08	.08	.09	.11	.06	.11
Std Deviation	.64	.59	.72	.78	.44	.84

An examination of the graphs in figure 4.2 reveals that all of the distributions are reasonably symmetrical in shape with the bulk of values centred close to the mean, indicating an approximately Normal distribution of individual factor scores. Notably, over 75% of the individual student scores for factors one, two and four, are above the neutral value of 3.

To examine the statistical significance for each of the six factors, 95% confidence intervals for the distribution of sample means were calculated and are displayed in figure 4.3.

Figure 4.3 95% confidence intervals for the mean factor scores



An examination of the 95% confidence intervals for the mean factor scores illustrates that three of the Factors, (F3, Focus on Assignments; F5, Distribution of Effort; F6, Approach to the Exam) include the neutral value of 3.0 from the Likert scale. We can conclude from this that there is no statistical evidence to suggest that the students either agreed or disagreed with the items provided in the questionnaire for these three factors. The other three confidence intervals (F1, Quality of Feedback; F2, Use of Feedback; F3, Learning from the Exam) do not include the value 3.0, indicating that there is strong statistical evidence that the mean scores for these factors are greater than the neutral value.

Table 4.3 Y12 Cohort mean responses to individual items

	Factor and question (N.B. questions with an '' are reverse scored)*	*Mean*	*Agree/ disagree*
Factor 1: Quality of feedback (mean = 3.4, median = 3.4)			
13	On this course I get plenty of feedback about how I am doing	3.6	agree
14	The feedback comes very quickly	3.2	agree
15	There is hardly any feedback on my assignments when I get them back*	3.6	disagree
16	When I get things wrong or misunderstand them I don't receive much guidance in what to do about it*	3.5	disagree
17	I would learn more if I received more feedback*	2.4	agree
18	Whatever feedback I get comes too late to be useful*	3.6	disagree
20	The feedback helps me to understand things better	3.8	agree
21	The feedback shows me how to do better next time	3.9	agree
22	Once I have read the feedback I understand why I got the marks I did	3.7	agree
23	I don't understand some of the feedback*	3.0	neutral
24	I can seldom see from the feedback what I need to do to improve*	3.3	disagree
Factor 2: Use of feedback (mean = 3.7, median = 3.8)			
25	I read the feedback carefully and try and understand what the feedback is saying	4.0	agree
26	I use the feedback to go back over what I have done in the assignment	3.8	agree
27	The feedback does not help me with subsequent assignments*	3.7	disagree
28	The feedback prompts me to go back over material covered earlier in the course	3.3	agree
29	I do not use the feedback for revising*	3.7	disagree
30	I tend to only read the marks*	3.7	disagree
Factor 3: Focus on assignments (mean = 3.1, median = 3.3)			
2	I can be quite selective about what I study and still do well*	2.7	agree
3	I only study things that are going to be covered in the assignments*	2.5	agree
4	I have to study regularly if I want to do well on the course	3.6	agree
5	On this course it is possible to do quite well without studying much*	3.5	disagree
Factor 4: Learning from the exam (mean = 3.5, median = 3.7)			
32	Doing the exam brought things together for me	3.3	agree
33	I learnt new things while preparing for the exam	3.6	agree
34	I understand things better as a result of the exam	3.7	agree
Factor 5: Distribution of effort (mean = 3.1, median = 3.0)			
1	I do the same amount of studying each week regardless of whether an assignment is due*	2.4	agree
6	In weeks when assignments are due I put in many more hours	2.2	disagree
7	Tackling the assignments really makes me think	3.9	agree
8	I learn more from doing the assignments than from studying the course material	2.9	neutral
12	The assignments are not very challenging*	3.6	disagree
Factor 6: Approach to the exam (mean = 3.2, median = 3.3)			
31	Preparing for the exam was mainly a matter of memorising*	3.1	neutral
35	I'll probably forget most of it after the exam*	2.4	agree
36	In the exam you can get away with not understanding and still get good marks*	4.0	disagree

4.4 Understanding assessment criteria

A key aspect of Sadler's (1989) theory of formative assessment is that effective assessment requires that students and teachers must develop a shared understanding of the assessment criteria they are working towards. This section reports on two ways that teachers attempted to develop students' understanding of the assessment criteria: through the structuring of the units of work, and the provision of a set of 'I can do' sheets that specify the assessment criteria in student oriented language.

Noticeably, despite the fact that Unit Standards are included in the Y12 Trad Maths course, in nearly all of the conversations where the assessments were mentioned, the research students and teachers talked almost exclusively about Achievement Standards.

4.4.1 The structure of units of work

Evidence from the focus group interview with the three teachers indicates that the teaching of the units of mathematical content are structured around the levels of the Achievement Standard. Instructional discourse and activities were, for the most part, focussed on providing students with opportunities to engage with tasks associated with the Achieve and Merit level assessment criteria, with less emphasis placed on the assessment criteria at the Excellence level. In the following extract from the teachers' focus group interview the structure and rationale behind the design of the units of learning is outlined.

> Ms Brown: We start off with Achieve, and then we say "Ok you have done all of the skills for Achieve so now we can move onto Merit", and then one or two periods on Excellence, usually only one. So they should know really what the level is. You are allowed to do that aren't you? That's good? Or is that bad?
>
> Mr Smith: I think [the structure] is good because it allows them to recognise which questions are which, because those who are just going for Achieve it is important that they focus on the Achieve very well. There is no point in spreading themselves too thin, almost getting Achieve or almost getting Merit so I think there is a lot of benefit for those kids in saying "right these are the Achieve questions, these are the Merit questions, you've got to get these right." Maybe just spend a little bit of time on the Excellence stuff.

Ms Brown: If I know that kids can't get Excellence then I'll say "we are not doing the Excellence". I'll say "have a go at it but if you don't like it then go back to these other things".

Ms Clarke: Sometimes I will just do the Achieve and Merit and I will say "look exercise such and such is covering the Excellence. Sorry we don't have time so you are going to have to do it at home". So if they really want to get Excellence then they are going to have to get off their bum and do it at home.

Mr Smith: I'll say "this is inverse normal; this is Excellence. I know that some of you are still struggling with the Achieve questions. I suggest that you carry on with exercise 14.2".

When asked, the teachers reported that although they would like to spend more time helping the low achieving students master the lower level skills they are often constrained by the time available to teach the unit of work.

In the above extract we note that Mr Smith aligns particular mathematical problems with a specific level of the Standard, for example, Achieve questions or Merit questions. During the classroom observations it was noted that teachers typically identified the level of a mathematical problem, but tended not to make explicit links between the assessment criteria and individual mathematical question or content. For example, one of the teachers identified a complex quadratic as being a Merit question in algebra ($x^2+8x=400$). Students may mistakenly believe that this is a Merit question because it is a multi-step problem requiring it to be equated to zero before it can be factorised and then solved. In fact, this is a Merit level problem because it requires the use of the quadratic formula, which is not immediately obvious in this example.

Conversations with teachers confirmed that they were familiar with the assessment criteria requirements for the various levels of the Standard, and that this helped them situate mathematical problems at given levels. Indeed the last comment by Mr Smith demonstrates understanding of the Excellence criteria for the Probability Standard. However, during the classroom observations, teachers were not generally observed making explicit links between mathematical question and the assessment criteria.

This extract also suggests that teachers expected some of the students to be strategic about their learning prior to the commencement of the mathematical unit. Mr Smith's

"just going for achieve" comment suggests that he believed that some students decide on the grade they are aiming for prior to the content material being taught, and that this dictates their approach to their learning in the unit. A possible implication of this practice is that students who only aim to secure an Achieve grade may concentrate on the Achieve material at the expense of a deeper understanding of the underlying content across all levels.

Data from the research student focus group interviews indicated that although students generally appreciated teachers identifying questions at a given level of the standard during the teaching of the unit, several students were concerned that this practice might limit the learning opportunities of some other students.

> I don't know that I like the fact that we get told that it is Achieve, Merit or Excellence. Some people could tend to be more lazyish and just learn the Achieve material and so they just pass. So they don't actually learn the Merit and Excellence stuff. I think we should just learn it all. (Sally, FG2)

Notably, none of the research students felt that they would adopt this approach. An examination of the students' approaches to learning occurs in section 4.6.4

Identifying levels on assessment tasks

Prior to the introduction of NCEA, the NZQA had proposed that questions in the assessment task be identified as Achieve, Merit or Excellence questions. This proposal caused much debate amongst teachers. Some teachers argued that it would help students be strategic about which questions to answer when sitting the assessment task. Others argued that it could potentially limit the achievement levels of students by encouraging them to only attempt part of the assessment task. This feature was subsequently removed prior to the introduction of Level 1 in 2002. However, in most current NCEA mathematical assessment tasks, although not specified as such, the first few questions are primarily designed to assess students against the Achieve criteria. While, later questions are primarily designed to assess Merit and/or Excellence criteria, they also gather evidence towards the Achieve criteria.

To investigate the extent to which teachers encourage students to be strategic about which questions to answer when sitting the assessment task, the issue of specifying question levels in assessment tasks was raised during the focus group interview with the teachers. While agreeing that the transparency of assessment criteria by level helped to focus their teaching during the unit of learning, they all endorsed the decision not to identify Achieve, Merit and Excellence questions in assessment tasks.

> I would never tell them in the test. You tell them in the unit so they should be aware anyway, and they all know by now that the first 4 or 5 questions are Achieve then Merit. I think all three of us will tell kids as many times as we can what Achieve Merit and Excellence were, but we wouldn't tell them in the test. (Ms Brown, FG)

The general consensus amongst the teachers was that if questions were identified by level in the assessment task then some of the students might only attempt the questioned labelled as Achieve, and this might limit their final grade. All three teachers reported that they never encourage students to only answer the easier questions. Rather they encourage students to try to answer all the questions in an assessment task. Ms Clarke indicated that students are often "surprised by how much they can do if they give it a go". Additionally, the teachers identified that a partial answer to a more difficult question may provide 'replacement evidence' for lower achievement grades.

Comments from the research students generally concurred with those made by the teachers. The students identified that labelling the questions by level in the assessment task may result in some students attempting a limited range of questions and hence potentially limiting their achievement levels.

> For people who are like "I just want to Achieve" then it wouldn't be because they'd just be like "I can do these ones so I won't even try for anything more", whereas they could probably get them if they tried. (Kate, FG2)

To explore the issue in more depth it was suggested to the students that, since the first questions tend to be the Achieve questions, if a student wanted to be strategic in sitting the assessment surely they could just attempt the first few questions. Rather than seeing this as being strategic, one of the students pointed out the fallacy of this

approach to sitting the assessment task, indicating their awareness of replacement evidence.

> And they probably wouldn't try them so if they got one of the Achieve ones wrong they couldn't use these other ones to fill in. (Tracy, FG2)

4.4.2 The 'I can do' sheets

Students are given a course outline book at the start of the year, referred to as their 'Blue Book'. This book has a timeline with the topics for the year listed, the relevant school assessment policies, and a set of 'I can do' sheets for the Achievement Standards used in the course (there are no 'I can do' sheets for the Unit Standards assessed in the course). 'I can do' sheets specify the assessment criteria in a language designed to be easily understood by students. The sheets were developed by the New Zealand Association of Mathematics Teacher (NZAMT), with schools being encouraged to adapt them to suit their own needs.

The following extract is from the students' Blue Book, and illustrates their intended purpose from the schools point of view.

> They provide a clear guide to the skills that students are required to have for each level of achievement. Use these sheets to set goals and guide your study for each topic.

The following figure is an example of the 'I can do' sheet for the algebra standard (see also Appendix M). Of note is the inclusion of an 'I have completed' column rather than 'I can do' column. This alignment to curriculum coverage, rather than curriculum mastery, possibly opens the way to misinterpretation by students or parents.

Figure 4.4 Example of part of an 'I can do' sheet.

Achievement Standard 2.1 - 90284 **Manipulate algebraic expressions and solve equations**	**I have Completed**
Achievement: I can • Expand brackets up to three factors	
• Factorise expressions where there is a common factor	
• Factorise quadratics with coefficient of $x^2 = 1$	
• Factorise quadratics with coefficient of x^2 an integer	
• Solve quadratic equations that can be factorised	
NB part of the Achievement criteria has been omitted for brevity	
Merit: I can • Use the quadratic formula to solve equations	
• Solve linear / non linear simultaneous equations	
• Solve exponential equations like $13^{4x-5} = 6$	
• Solve problems using algebra skills	
Excellence: I can	
• Understand the nature of roots of an equation	
• Use the roots of an equation to solve a problem in context	
• Choose and use algebra skills to solve problems	

Evidence from this project suggests that, despite clear recognition of the formative potential of the 'I can do' sheets by both the teachers and the students, the sheets were underutilised. Although there were occasions during the classroom observations when teachers referred to the Blue Book and the 'I can do' sheets, it appeared that this was not a normal occurrence. When direct references were made, very few of the students located their Blue Books and followed what the teacher was saying.

Students appeared to be uncomfortable when interviewed about their use of the 'I can do' sheets. Many admitted that they hadn't looked at them, with some commenting that they didn't know where their Blue Book was. Megan incorrectly thought that the 'I can do' sheets were handed out individually at the start of each unit rather than being in the Blue Book. "Well normally the teacher passes them out and I just like chuck them in the bin anyway" (Megan, FG2). Another student commented: "Sorry, I don't know what the 'I can do' sheets look like" (Chelsea, Q1).

Evidence indicated that although a number of students had initially used the 'I can do' sheets, their use of the sheets rapidly decreased during the year. For example, in the first focus group interview, James reported that: "I like, look up and see, oh yeah,

that's what I need to do to pass." By the second term, however, James revealed that he had not been using the sheets. He could offer no reason why he had stopped using them as part of his study strategy. Later in the year, James reported that he had tried to use the 'I can do' sheets to study for the mock examinations, but his comments revealed that he was not using them effectively. He used them to gauge his perceived achievement against the assessment criteria, but was not proactive in following up on his identified areas of weakness.

> For me it is just like "Can I do this?" And I'll just go tick. "Can I do this? No". So I just tick no and carry on to the next one. I don't try and go back and do it. (James, FG3)

Sally also stated in the first interview and questionnaire that she used the 'I can do' sheets. Despite recognition that the sheets detail the required assessment criteria, she incorrectly reported that the 'I can do' sheets specify exercises from the text-book to be completed.

> Yes, I commonly refer to the 'I can do' sheet when studying for tests. I find it gives good detail on what I need to know. I do not use it as a check list but I do make sure I have completed all of the exercises it states. (Sally, Q1)

During the second term Sally still appeared unsure about the content and structure of the 'I can do' sheets. She seemed unaware that the sheets specified assessment criteria for all three levels of the Achievement Standard.

> Sally: She hadn't really defined what we were going to do in that. I knew we had some expanding and factorising and stuff for Achieve level, but we didn't actually know what we were doing for the Merit and Excellence level, so I got there and thought "I don't know how to do these".
>
> Researcher: So she didn't get out the Blue Book and say "ok for Merit …"
>
> Sally: It's only for Achieve level. Oh no, it has Merit and Excellence levels, doesn't it? I think it does have Merit and Excellence?

Further comments in the fourth term indicated that Sally had looked at the 'I can do' sheets. Despite familiarising herself with their structure and content, she still found them of limited value.

> No I didn't refer to my 'I can do' sheets, I looked at them and they were quite hard to distinguish so I put them aside and just worked from the exercise list Ms Brown gave us.

> But they are not really clear... I don't know...it is just like, "I can do this" and I am just ... "well can I?" I need a little bit of help interpreting whether I can or can't. Where would I look? It needs to say, refer to chapter 23.3 for this can I do. (Sally, FG3)

The provision of references to text-book exercises was also mentioned by a number of other students as a potential improvement to the 'I can do' sheets.

> If there was an example question to go with each requirement, but that would be a bit difficult. (Andrew, Q1)

> I found with the Blue Book that it didn't really give me like exercises to do. It would just say "I can draw a linear graph" and I would be just like "Yeah, I suppose I can, but what exactly is a linear graph". It's not like "ok exercise 13.5", I've covered that so I can flip back to 13.5 and say "Oh yeah I can do that". (Raewyn, FG2)

On a number of occasions during the year, students stated that they used revision sheets provided by their teacher as part of their study strategy. Many of these revision sheets are copies of past assessment tasks, or exemplar assessment tasks off the internet. Although these revision sheets gave examples of the sort of mathematical problems that might be expected in the assessment tasks, they typically did not specify the achievement level of questions, nor were they linked with the assessment criteria specified in the 'I can do' sheets. Accordingly, although students found these revision sheets useful as part of their study strategy, they did little to help students develop an understanding of the assessment criteria they were working towards.

While admitting that they did not actively use the 'I can do' sheets, many students could recall the basic structure and content of the sheets, and could reflect on how they might be used to benefit learning. The comments below also highlight recognition of the importance of the self-monitoring and self-regulating components of metacognition.

> You can mark off what you definitely know and what you need to work on as you study. (Andrew, FG1)

> It shows the basic things you need to know and you can see what you do know and what you don't. (Raewyn, Q1)

The importance of metacognitive strategies was highlighted by one student. Her perception of her understanding of the content affected her preparation prior to the assessment task.

> If I don't understand the material I get very frustrated and find it really hard to then attempt to study because I say to myself "well I don't know it, so why bother". (Megan, Q1)

Despite Megan's ability to self-monitor, she has no metacognitive strategy to cope with the perceived difficulty.

Only one of the research students consistently claimed throughout the year to actively use the 'I can do' sheets as the following quotes from the first and fourth term indicate.

> I like the 'I can do' sheets a lot. I like knowing exactly what I must know and understand for various levels of achievement in order to be able to achieve my highest. (Kate, Q1)

> I find the 'I can do' sheets very helpful. I use it as a checklist; what I know and what I need to learn/brush up on. I use them to know what to study for the tests. (Kate, FG 4)

Despite Kate's proactive use she did express some frustration that there were no 'I can do' sheets for the Unit Standards.

4.5 Provision of feedback

The second key characteristic of Sadler's (1989) theory of formative assessment is that students must "compare the *actual* (or current) *level of performance* with the standard" (p. 121, emphasis in original). In addition to student's self-monitoring, this characteristic implies that students must receive appropriate feedback on the gap between their actual level of performance and that specified by the Standard. In most instances this feedback is provided by the teacher.

As noted in chapter 2, the provision of quality feedback has been widely identified in the literature as an important feature that facilitates students' formative use of assessment information (Black & Wiliam, 1998a; Crooks, 1988; Hattie & Timperley, 2007; Kluger & DeNisi, 1996). Students must first receive and understand feedback

before they will be in a position to engage in appropriate action to address deficiencies in their learning.

The following section examines students' perception of and preferences for the feedback they receive, provides examples of the actual written feedback given to the students, and discusses teachers' perceptions and preferences for giving feedback related to assessment events.

4.5.1 Characteristics of written feedback

During the focus group interviews, and in the written questionnaires, students were asked to discuss the nature of the feedback they received in both their assignments and assessment tasks. Their comments concurred with the data from the AEQ (section 4.3) indicating that students perceived that they got reasonable amounts of feedback that helped them to understand their errors, or incomplete understandings, and suggested strategies for addressing them.

Evidence from the interviews clarified that the quantity and type of feedback was dependent on the nature of the assessed work. Students identified that teachers typically provided reasonable amounts of feedback on assignments, with significantly lower amounts of feedback on both practice assessment for externally assessed Standards, and internally assessed Standards. In the following quote, Megan describes the feedback she receives from assessment tasks.

> Sometimes comments like you made a simple error. But normally there are just ticks on the test paper that say what you got right and wrong but not which part of the problem is wrong. (Megan, FG1)

An examination of the students' assessed work generally confirmed this observation although a number of good examples of written feedback were identified in assessment tasks.

Megan's comment highlights the notion that, at the very least, students want to know exactly where they went wrong in an incorrect answer. The teachers' perspectives of providing feedback on assessments tasks is examined in more detail in section 4.5.3.

In order to explore students' understanding of the term 'feedback', the definition adopted in this research was not discussed with the students. When asked to describe the feedback they received, students distinguished between written comments and other indications of achievement, such as ticks, crosses and grades. They typically talked about written comments as feedback. They found other written indication of achievement to be of limited value.

> Ticks and crosses are not very helpful at all, and neither is just an 'M' on the front. (Sally, Q1)[10].

When discussing the feedback they received on assignments, students could readily identify characteristics they found useful.

> If I complete an assignment my teacher usually writes good comments which are very helpful and also corrects mistakes so I can look back and see what to do for next time. (Andrew, Q1)
>
> He tells us exactly where we went wrong, how to change it, and what to do next time. My maths teacher shows us exactly where we went wrong. It's very helpful. I wish all teachers would. (Raewyn, Q1)
>
> Shows us where we went wrong and sometimes the correct answer. Also an overview of how we could improve our answers in general. I can normally figure out where I went wrong and how to fix it with their notes. (Hannah, Q1)

Of particular note in these extracts is the students' expectation that, in addition to identifying errors in their work, feedback should suggest a corrective strategy for future learning opportunities. The notion of scaffolded corrective strategies was a common theme in the students' comments throughout the research.

Responses to the questionnaire in the fourth term reinforced that students appreciated: "comments about mistakes and improvements that need to be made in my work" (Andrew, Q3), and also wanted teachers to "give some constructive criticism, not just plain old criticism" (James, Q3). One student wanted the teacher to go even further saying: "It would be really good if he could write out the steps I needed to do on all questions" (Tracy, Q3).

---

[10] The 'M' referred to in this quote represents a grade of Merit in the Achievement Standard.

Students were, however, also aware that providing quality feedback is time-consuming for teachers, and expressed appreciation of teacher efforts. “When Ms Brown has time she will write working … this is very helpful” (Sally, Q1).

It emerged during the course of the year that the characteristics of feedback students found helpful were independent of whether the assessment carried any credit value, or whether they were going to be assessed on that work again during the year. Most students appeared to have an inherent interest in knowing ‘what they did wrong’ and ‘what they should have done to get it right’, and this interest was not dependent on whether they were offered a subsequent resit opportunity worth credits. It suggests that the research students have an intrinsic interest in their own learning. This observation has implications for the provision of feedback and will be discussed in more depth in the discussion chapter. (For further evidence of students’ intrinsic interest in their learning see sections 4.6.4, & 4.6.5)

Scaffolded comments

The following extract from a focus group interview exemplifies the view that students prefer feedback that details errors and scaffolds remedial action. The discussion refers to a scenario where students had received their assessed scripts back after the mock examinations in the third term, and were told to go back and correct their mistakes, and then submit their work for remarking. The Achievement Standards in question are externally assessed at the end of the year so their current grade was indicative of their achievement level in this Standard at the time of testing.

Sally: She said, “OK next Monday I want you to hand them in done”, but she gave us just the answers for each of the questions. I found that hopeless. I could not figure out how they got to that answer and I was just sitting there and ...

James: You sit there for about an hour on your calculator and you do all sorts of things.

Sally: And you just cannot figure it out and I found that really useless.

James: There were a couple of the questions, the really crazy ones, where she did the first bit for us and then told us what to do next, like you need to factorise, expand solve and then ...

Sally: That was good. That was helpful but she didn't do that for all of them.

James: It's like, pointing you in the right direction, you are learning more, whereas, when you just get the answer you are just like, "oh so what".

Sally: It's like, I didn't know how to do that to start with so it is not going to help me just knowing the answer is 0.182.

James: I could have got the answer by myself if I knew how to work it out.

Megan: I like to know where I went wrong, and what I have to do to fix it, rather than just being given the answer.

James: You see where to go, and all of a sudden it kinda clicks into place, and you are like, "oh I see it is just that simple process".

This extract illustrates a number of important points. Firstly, students perceive that the provision of answers only does not constitute quality feedback. Secondly, students prefer scaffolded comments, encouraging them to engage with the feedback to develop corrective strategies. Thirdly, although it did not work quite as the teacher might have expected, this extract illustrates the teacher's attempt to get students to actively engage with feedback formatively.

Limitations of feedback

As noted earlier, students often received limited feedback on their assessment scripts. One implications of this practice is that students are often confused by the limited feedback they receive. This was highlighted by Sally who showed how frustrated she was when she could not work out why she got a question wrong in an assessment task.

Sally: That's why I questioned my graphs. She had ticked my table method and crossed my graph and I was like... "what's going on" because my table method was an exact replica of my graph. I was like "Miss, I don't understand. Why did I get this wrong" and she said "you didn't extend it".

Researcher: So perhaps if she had written "extend your graph" next to it?

Sally: Yeah, yeah. I would have been like "Oh I get it".

A number of other students also indicated that they had found feedback on assessment scripts difficult to understand at times.

> Sometimes I get lots of "????" and things like "what's this?" written on my sheet. I don't find this useful because I don't really know where I went wrong or why I got that particular question wrong. (James, FG1)

In one particular instance Kate talked about an example of feedback she had received on an assessment task that she felt was potentially useful. Her teacher had marked the assessment script and had mistakenly written a formative comment next to a question that she had answered incorrectly. The teacher, reviewing his marking, subsequently decided to offer the student a resubmission opportunity and so had to cover up his formative comment so she was not advantaged by it.

> Often there are not very many comments. However, in our algebra processes I confused an explanation of a question but I was able to change this through resubmission. However, the comment was covered up so I was unable to see it. When I re-examined the question and answered it correctly, Mr Smtih then uncovered the comment. It was very helpful and if I had gotten the question wrong, such comments would be very useful. It explained what I had done and what should have been done. (Kate, Q1)

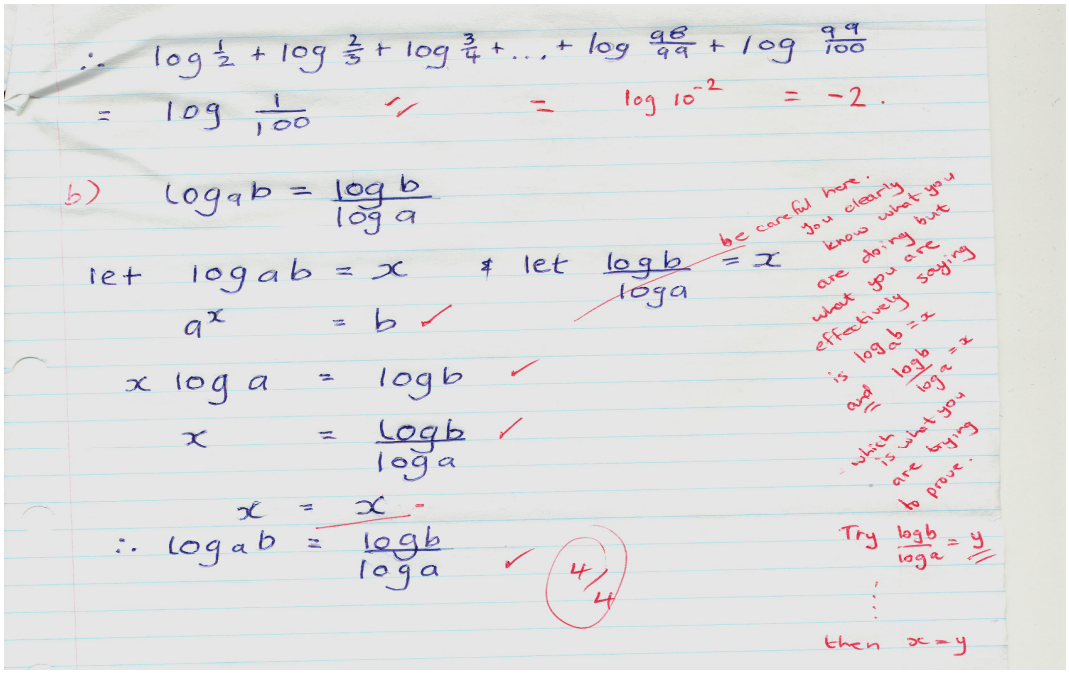
The potential for this technique to address some of the concerns regarding differential feedback practices on assignments and assessments will be discussed more in section 5.3.3.

4.5.2 Examples of feedback on student work

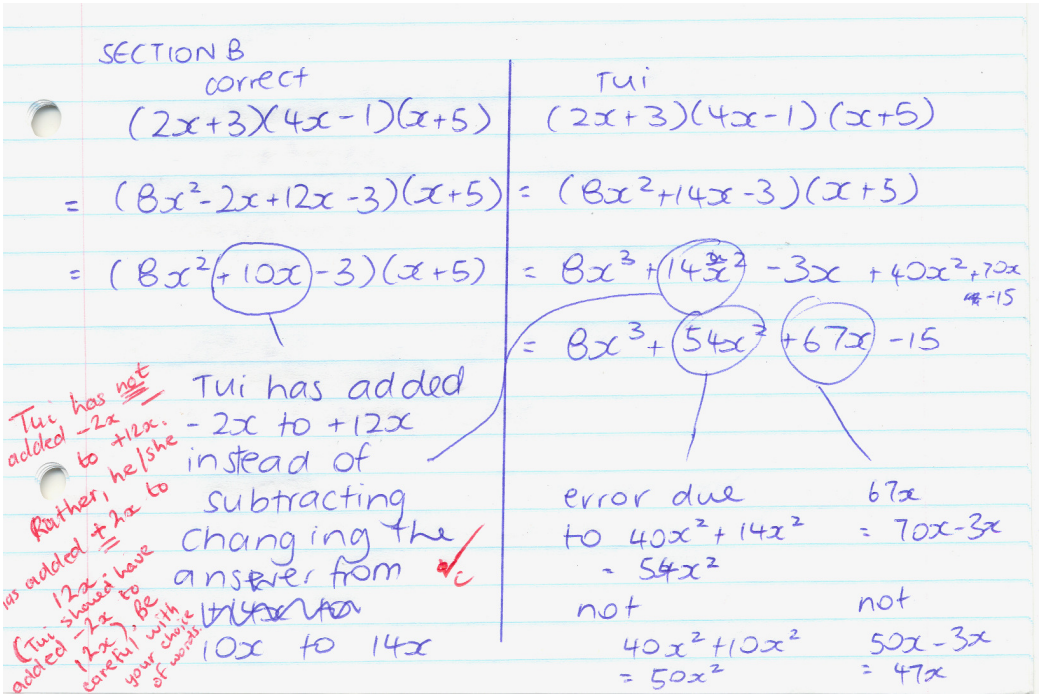
The following extracts from students' scripts provide examples of the types of feedback offered to students. An examination of students' assessed work confirms students' perceptions that they receive more written feedback on assignment work, although many examples of good feedback were found on assessment scripts.

The initial set of examples are related to the algebra Achievement Standard 2.1 'Manipulate Algebraic Expressions and Solve Equations'. This standard is an external standard that is summatively assessed in the external examinations at the end of the year. The examples for this standard come from three sources: from assignments completed during the unit of work, from formative assessments sat at the end of the unit in the first term, or as part of the 'mock' examinations sat in the third term.

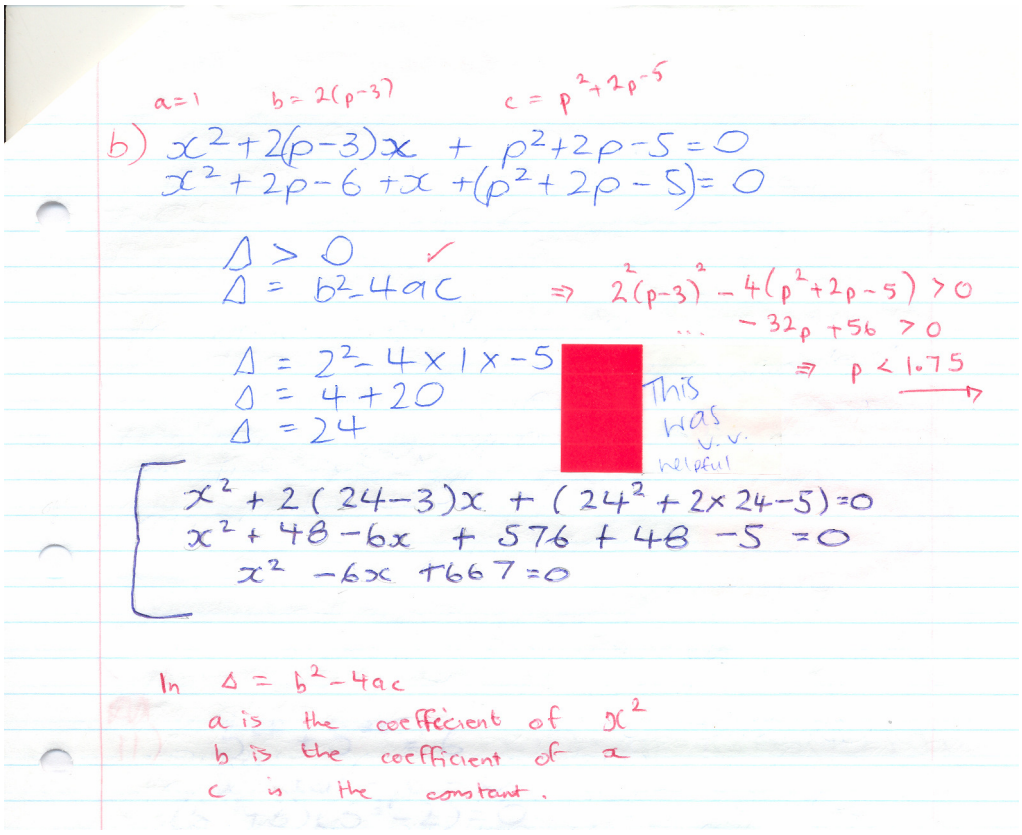
The first example is from a student's assignment and relates to a question that has been marked correct. While the feedback acknowledges Kate's understanding of the problem, the teacher identifies potential flaws in the presentation of her answer and suggests a more acceptable solution.



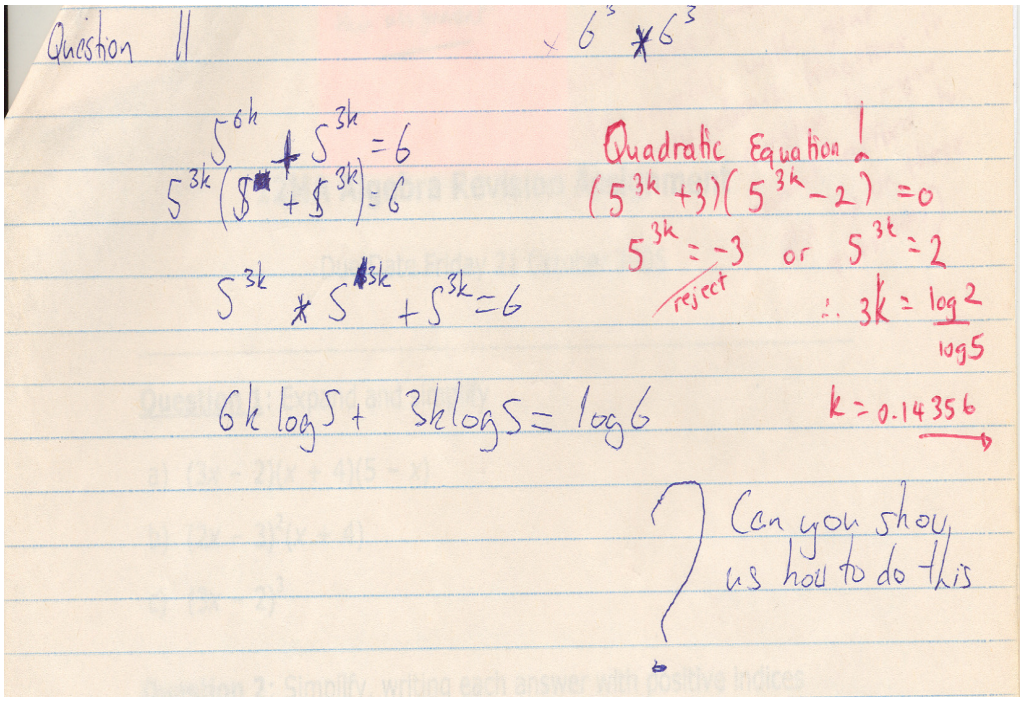
The next example is from a question in an assignment where the student was given a problem and an example of how a fictitious student, 'Tui' solved the problem. This question is essentially asking the student to do some peer marking of someone else's work. The student (Raewyn) was asked to find the mistake in Tui's working and identify what she should have written. Once again Raewyn displays a high level of mathematical understanding, but has not quite presented her answer in a coherent manner. The feedback attempts to clarify the small error in the Raewyn's answer and guides her in how to present the answer in an acceptable form.



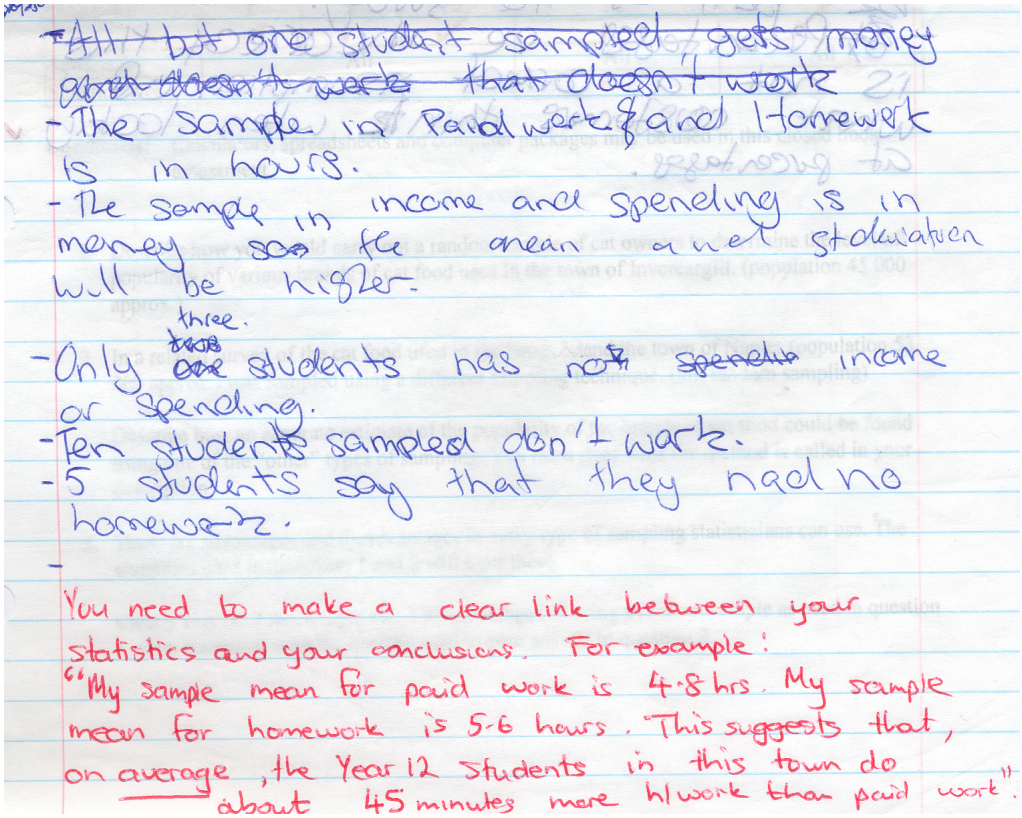
The following example is also from an assignment. Here the feedback identifies the error and also the next steps in the working leading to a solution. Raewyn knew that her assessed work would be looked at to select examples of good feedback practice. Accordingly she annotated her assignment to highlight the feedback that she found particularly helpful.



In the next example, Hannah perceived that she could not do this question, and had written a note on the script asking the teacher for some help. It is interesting to note that Hannah has written "can you show us how to do this?" rather than "can you tell us the answer?" suggesting a desire for a corrective strategy, rather than just the answer. I talked to Hannah individually about this immediately after one of the focus group interviews. She reported that, not only had the teacher provided the written feedback you can see on the script, but the teacher had also followed it up with her individually in class.



The last example is from an internally assessed Achievement Standard in statistics. This example is from the first opportunity students had to sit this Standard. The feedback offered to Andrew is about the shortcomings of his conclusion and how it could be strengthened: "You need make a clear link between your statistics and your conclusions." The teacher follows the advice with an example solution. These comments are particularly pertinent as students were offered a reassessment opportunity if they wanted to increase their grade.



4.5.3 Teachers' feedback practices

When describing their feedback practices, all three of the research teachers indicated that they prefer to give oral rather than written feedback on assessed work; they view this as an efficient way of communicating feedback to students. Such feedback is usually provided in a whole-class setting, concentrating on common mistakes made by the cohort.

Ms Clarke: I find I am writing the same comments.

Ms Brown: Same here.

Ms Clarke: So you may as well talk to them as a class.

Ms Brown: It is usually the same things coming up again so it is easier if I write it up on the board.

Two of the teachers expressed this view more strongly, with the third teacher, Mr Smith, describing his practice as a more mixed delivery of written and oral feedback practices.

All three teachers commented that providing written feedback was very time consuming. They questioned whether this was the best use of their time, especially since there was some doubt in their minds as to whether many students read, understood or acted on the written feedback provided.

> I just find that a lot of people just don't want to read written mathematical things if they got it wrong in the first place. Some do but it's just such a mission to write it down. (Ms Clarke)

Consistent with the classroom observation noted in section 4.2, the teachers also commented that many of the students lacked focus during these whole-class oral feedback sessions. Oral feedback sessions were associated with classroom management concerns.

> It is something that I do find difficult, how to give that feedback. How to go over a test that would make the class listen and want to do it, without having half of the class muck around. If you let them go over it themselves then you will always have about 15 kids who will just have a conversation, and all of those management issues. (Ms Brown)

Despite the teachers' personal preferences for giving oral feedback, the teachers were interested to learn about feedback practices that would increase effective learning, and decrease management issues. In a later conversation, Ms Brown indicated that she was uncomfortable with her existing feedback practices, and expressed a wish to find out more about the best way to provide assessment feedback to students. Once again, she expressed doubts about whether students read any written feedback given.

> I almost never give written feedback because I just can't be bothered writing it down. I will say it to the whole class. I suppose that is bad. All teachers should spend time writing it down. But I just find that often the kids don't even read it. It would be nice to know if they read it, that's true, and especially whether they find written, or personal, or group, or board work as the best delivery of the feedback. I'd be very interested in what would be most effective in terms of our delivery. (Ms Brown)

Ms Brown's expression of interest in what constitutes best practice is indicative of all the teachers' high level of reflective interest throughout the research project. All three teachers maintained their willing participation in the research process and were open to a critical examination of their existing practices.

In describing his feedback practices, Mr Smith explained that he tried to give reasonable amounts of written feedback when time permitted, but also noted that the amount of feedback was often determined by the grades achieved by the students. Students awarded Not Achieved grades received the most feedback, with Mr Smith arguing that: "They would want to find out what they did wrong so they can pass". Students who received an Achieve grade, but who were not close to receiving a Merit, received limited feedback. Students who were close to securing the next achievement level (i.e. close to the Merit or Excellence boundary) received moderate targeted written feedback designed to help them improve their performance in subsequent assessment tasks.

As identified in section 4.2, the school has a policy of allowing students resubmission opportunities. This policy has implications for the amount of written feedback teachers put on internally assessed NCEA tasks. The resubmission policy effectively means that teachers need to make a judgment as to the 'correctness' of a particular answer, and whether the students will be given the opportunity to resubmit their

answer, prior to the teacher indicating errors or writing feedback. By not identifying errors, the teacher can give the assessment script back to the student for them to review and correct any errors themselves without unduly advantaging the student.

The research teachers commented that the resubmission policy did impact on the amount of feedback they offered. It was time consuming to "go back and mark a question a second time", therefore adding to time it took to complete the marking of students' assessment scripts. Although marking workload, and the desire to return work quickly, were consistently identified by the teachers as factors affecting the amount of written feedback they offered, these factors appeared overshadowed by the teachers' belief that students don't read written feedback.

> Yeah I'd write more if I really thought the kids would read it. (Ms Clarke)

4.6 Engagement with feedback

The third component of Sadler's (1989) theory of formative assessment is that students must engage in appropriate action to address any gap between actual and desired achievement levels. From the literature we know that a number of complex interwoven factors, including teachers' practices, students' motivation, classroom culture, and conceptualisations of ability, all influence students' predisposition to engage with feedback (Black & Wiliam, 1998a; Dweck & Leggett, 1988; Harlen & Deakin Crick, 2003; Kluger & DeNisi, 1996; Weiner, 1979). These factors have a co-dependence, and it is recognised that it is the interplay between these factors that will ultimately determine whether students use the feedback to positively impact on their learning.

To examine students' engagement with feedback in the current study, the following section initially reports the data relating to the opportunities students have to engage with the feedback, beginning with the return of assessed scripts and assignments. It then examines students' motivations and goals for learning that influence their predisposition to use formative feedback. Lastly, it investigates a number of factors that help to constitute the assessment culture in the classroom: pressures on

pedagogical and learning strategies, learning orientations, and conceptualisations of assessment and ability.

4.6.1 Return of assessed scripts

With regard to NCEA, the first opportunity students had to engage with the teachers' written and oral feedback from assessment tasks was when their marked assessment scripts and assignments were returned to them during a normal mathematics period. Although class periods when assignments were returned were not observed, discussions with students and teachers confirmed that the structure of these periods were similar to the periods observed when assessment scripts were returned.

As noted earlier, responses from the AEQ reveals that the cohort students perceived that they read the feedback they are offered and tried to relate it to their assessed work, and did not simply focus on the grade. These perceptions are in accord with data from the focus group interviews. In the interviews, most students reported that despite being initially interested in their grades and those of their friends, they were more interested in looking at the questions they found difficult in the assessment task.

> Normally, I am interested to see if I got it right first. I may have been unsure and if I got it wrong I like to know what has to be done to correct it. (Kate, FG2)
>
> I am interested to see where I went wrong because I get frustrated when I can't do something in a test and spend ages thinking about it. (Hannah, FG2)
>
> I want to see if I'm right but if I'm wrong I want to see how to improve. (Raewyn, FG2)
>
> Firstly I look at the ones that I found hard to see if I got them correct then I will look over the rest of the test and see if I got questions I found easier, correct. (Andrew, FG2)
>
> I only double check with the questions that I got wrong. (Chelsea, Q1)

These accounts are consistent with the evidence from the classroom observations. During the observations students were generally seen to be discussing particular questions rather than comparing grades, suggesting a 'learning' rather than a 'performance' orientation (Dweck & Leggett, 1988).

However, there was some support for the teachers' perception that students do not read the written feedback offered. A small number of students indicated that the amount of time they spent looking over questions was dependent on either the amount

of effort they put into preparing for the test or the level of perceived difficulty for them.

> If I know I didn't try very hard I just kind of look over the test, realise that I probably should have studied harder, and then shrug it off. (Tracy, FG2)

Students expressed mixed perceptions when asked about the teachers' practice of providing whole-class oral feedback. Some clearly saw it as a valuable component of their learning subsequent to an assessment event.

> After an assessment our teacher is very helpful. If there is a question that most people got wrong then he often shows the whole-class how to do it correctly on the board. (Kate, FG2)

> I also like the whole-class oral as if there is an uncertainty I am able to ask about it and have it explained. (Hannah, FG3)

> Ms Brown is usually very good in allowing us time to read over our papers and ask any questions we have about them. I think that we need to understand what we did wrong so that we can correct it in the future, so yes I always use the opportunity to ask questions. (Sally, FG1)

Other students, however, regarded the practice of the teacher orally going over large portions of the test to be an inefficient use of time. In some instances this was because: "Going through the entire test is a bit boring if you get the majority of it right" (Hannah, Q3), while in other instances: "If you got too many questions wrong it can be a bit overwhelming" (Raewyn, Q3).

> It depends on how many questions I found hard. If there are many then no, but if it was a certain one question, then yes. (Tracy, FG3)

During the class observations when the assessed work was returned, it was noted that very few students asked questions of their teacher, either in the whole-class forum or subsequently on an individual level. When asked about this observation in the subsequent focus group interview, one student indicated quite strongly that she does not ask the teacher for help, although she was reluctant to expand on why not: "Nope, I never discuss any of the questions" (Chelsea, Q1). Two other students indicated that they only sought help in situations where they had major difficulties, while another student identified class size as a contributing factor in her decision not to ask for help from her teacher.

> There are heaps of kids in my maths class so it's hard to get a lot of individual attention so if I am really having a huge problem I will ask the teacher but if it's just something we will go over at the end of the year I don't bother. (Megan, FG2)

The majority of students indicated that they would not like the teachers to initiate individual oral feedback with them. They commented that they needed time to reflect on any written feedback offered, and try to correct the errors themselves

> Researcher: How would you feel if the teacher came round and sat down with you individually and talked with you about your assessed work?
>
> Raewyn: I think it would annoy me a bit because I like to try and work it out first myself and if I can't work it out then I will ask them. But if they just came round and sat watching me …
>
> Kate: Yeah you would get kinda freaked out.

Preference for working with peers

As noted in section 4.2, classroom observations revealed that many of the students did not appear to be focusing on the teacher led oral feedback following the return of the marked assessment scripts. Some of the students appeared to be ignoring their returned assessment scripts. Others were observed to be discussing their assessment scripts with their classmates rather than attending to the teacher led oral feedback.

The nature of these peer interactions was explored in the focus group interviews. Students reported that, after a brief initial discussion of the grades received by their peers, the conversations quickly turned to discussions of particular questions on the assessment task. Despite recognising their teacher's openness to answering individual queries, the majority of students indicated a clear preference for working with their peers to develop corrective strategies. Students would discuss individual questions with their peers in an attempt to understand how the written feedback related to their work. They would also ask around their peer group to see if anyone got the question right and could help them with a corrective strategy.

> Normally if we have a question wrong we compare and look at the other's working to find out where we went wrong. If we both got it wrong then we discuss how we view the question. If I cannot see where I went wrong from my friend's assessments then I ask the teacher. (Kate, FG2)

> We discuss how we arrived at our conclusion. For example in algebra we talked about how we interpreted the application question, and how we turned that information into a number problem. (Hannah, FG2)
>
> Often I will ask my peers, especially my great friend Jenny, about a question before I ask the teacher. It seems to work well because we can fill most gaps between ourselves. (Sally, FG2)

This description of the peer interactions is consistent with field notes made during the class observations and reported in section 4.2. The field notes record that the: "Students are briefly asking their friends what grade they got. They seem to quickly move past this and are asking each other about individual questions. They don't appear to be establishing a pecking order as we might expect in a Norm-Referenced class."

4.6.2 Review of marked scripts

Responses from a number of the focus group interviews and written questionnaires indicate that many of the students appreciated feedback consisting of written comments because you could "refer back to them later" (see section 4.5.1). Additionally, the value of using past assessed scripts, and the embedded feedback, as a study resource was highlighted in the AEQ (see section 4.3.2). The notion of referring back to past assessed work was subsequently followed up in the focus group interviews.

Despite recognising the formative potential of previously assessed work, there seemed to be little evidence that students used them for this purpose. It should be noted, however, that students did not retain their assessed work. Their assignments and assessment tasks were stored by their teachers in individual student folders. The school had a policy of keeping these as evidence for judging aegrotat passes. Students had access to their assessed work either during class or during Support Maths[11], but were not able to take assessed scripts home. Interviews with the teachers and

[11] Support Maths is a mathematics department initiative to provide students with the opportunity to receive additional mathematics help. It runs on selected lunchtimes and a number of afternoons after school.

interviews with the students revealed that very few students asked to see their assessment folders. While the majority of students reported that they were aware that they had access to their assessment folders, two students indicated that they were not aware they had access to previously marked assessments scripts. “I did not even know we were allowed access to our old marked scripts” (Sally, FG2). Another student commented that their previous work would be of limited use to them arguing that: “I did so badly I wouldn't know where to start” (Tracy, FG2).

4.6.3 Support Maths

The school ran a ‘Support Maths’ programme during a number of lunchtime and after school sessions. The programme was staffed by mathematics teachers, and students could attend sessions to receive individual help. This was one of the opportunities where students could access their past assessed work. Most students recognised that, in principle, Support Maths was a good idea.

> I think that is what Support Maths is all about. If you need that extra work then you go to Support Maths and you get one on one teaching there. (Tracy, FG1)

However, nearly all students reported that they were not regular attendees at the Support Maths sessions. A number reported that school based and outside commitments made it difficult to attend Support Maths regularly. Some students indicated that they would attend if they were experiencing severe difficulties, but many students stated a preference for attempting to resolve their learning difficulties by themselves, a finding consistent with previously stated data (see section 4.6.1).

> I find if I’m really stuck I will go to learning support but not usually just for day-to-day work, I would usually try to figure it out on my own. (Sally, FG1)

> I am sure that I will take up the learning support some time closer to the exams but for now I like to do things by myself. (James, FG1)

A number of students indicated that they had used Support Maths in preparation for the external examinations in their previous year (Y11), and planned to attend closer to the mock or end-of-year external examinations. However when this was followed up in the fourth term focus group interviews, none of the students had been to Support

Maths prior to the third term mock examinations, and none of them had started attending as they prepared for their Y12 external examinations.

During the time of the research the school was trialling a new system in Year 12 Traditional Mathematics. In this system students who had failed an assessment for an internal Standard were required to attend at least one Maths Support session before being able to resit the assessment. This was generally seen as a good policy with students arguing that it was important that individuals take responsibility for their learning, making an effort to correct misunderstandings and learn the required material before being able to resit the standard.

> I think you need to make sure that you understand it or like don't even waste the paper. If you are not going to even bother to learn then why do it anyway. (Megan, FG2)
>
> Why waste the teacher's time marking it if you are not going to learn. (Andrew, FG2)

Conversations with the teachers suggested that despite general support by both staff and students, there were practical considerations with the application of this policy. They reported that busy students often found it very difficult to attend Support Maths sessions, and that, in some instances, students were not prevented from resitting an assessment if they had not attended a Support Maths session.

4.6.4 Motivation to act on feedback

As detailed in chapter 2, research has consistently demonstrated that students' motivation for learning and learning orientation are central to their disposition to engage with feedback (Dweck & Leggett, 1988; Harlen & Deakin Crick, 2003; Weiner, 1979). AEQ responses illustrated that students generally found sitting examinations to be a constructive learning experience, suggesting that assessments positively impact on motivation for learning. Accordingly, during the focus group interviews, students were asked to describe their motivation and goals for learning.

An examination of the data from throughout the year reveals that most students consistently reported that a conceptual understanding of the underlying mathematics was a primary goal for their learning, and hence a primary motivation for them to read and act on feedback from assessed work. Students did not just want to 'rote learn'

techniques so they can 'pass the test'. They wanted to learn the mathematics, so that they could be flexible mathematical thinkers.

> I cannot use methods, formulas etc without understanding how and why they are used. (Sally, Q1)

> I get frustrated when I don't understand. It is important because if you come across something slightly different to what you are used to you can usually nut it out if you understand the concepts. (Hannah, Q3)

Students identified that a sound conceptual understanding is important for both short and long term goals. In the short term, students identified that understanding the material was going to help them when they had an assessment task. A number of students claimed that a sound understanding helped them: "to remember it" (Hannah Q1), and that that they will be: "prepared for a wide range of questions if you understand the material well" (Raewyn, FG3). These views expressed by the students are consistent with Bigg's (1987) 'deep-achieving' approach to learning.

Students also identified that a sound conceptual understanding was important for the medium and long term goals of preparing themselves for future learning opportunities.

> I feel it is very important to understand as much as I can to further my education. (Andrew, Q1)

> Yes it is important because to use it in later life you need to understand it. (Tracy, FG1)

Although parental expectations to do well were acknowledged, most students presented themselves as self-motivated learners with a strong internal locus of control; no students felt they were being unduly pressured to achieve.

> Very self directed. I want to do well and I know that will take work so I do it. (Hannah, Q3)

> I feel that I am like the only person actually pushing myself but I think I might need an extra shunt, like kick start type of thing to get me going. (James, FG3)

> My mum is like you need to do well this year but I think the motivation is more for what I want to do. (Megan, FG3)

4.6.5 Securing a grade that reflects their understanding

Students also expressed a link between a sound understanding of the mathematics and securing higher grades. There was awareness that Achieve questions test relatively low level thinking skills, but Merit and Excellence questions require the ability to integrate a range of methods and use higher order thinking skills.

> I think that in this course you are probably able to get Achieve without understanding the material fully; just do some questions here or there. In having a good understanding of the material you are able to apply it to harder level questions as Excellence often requires a combination of skills. (Kate, FG3)

The provision of a hierarchy of achievement grades in Achievement Standards was generally appreciated by all students, and acted as a source of motivation for learning. Contrary to the perceptions of the research teachers, the students did not appear to be 'credit chasing' or settling for a minimal Achieve pass.

> I think having a Merit, Excellence system is a good idea. It gives you a higher sense of achievement than just 'passing' a test. Personally I like having Merit/Excellence because it gives me something to aim for. (Sally, Q1)

> If I do not pass as highly as I would have liked then I feel disappointed with myself. I like to understand so that I am able to achieve highly … I think it is important because I like to do well and I know that I should be able to gain Merit/Excellence … I feel that, for me, getting Achieved is like failing as I like to aim for the best. (Kate, Q1)

Additionally, students saw securing a high grade as a measure of one's work ethic.

> I feel if I get an Achieved I have failed myself because to get an Achieved just means you did the bare minimum to pass. It's important to me because I want to become extremely successful in my life and I feel if I don't get good marks then when I try and go to apply for Med School they will just be like, "whatever, she only got Achieves. That means she didn't work hard". (Megan, Q1)

> It is sort of important to gain a Merit or Excellence because those grades are going to be my learning record. (Chelsea, Q1)

Students identified that securing a grade that reflects their mastery of the material was important for their own sense of self-efficacy. Students did not just want a high grade, they wanted the 'right' grade; a grade that reflected their ability: "Whether that is an Achieve or Excellence" (Andrew, Q1), because it "confirms that what I think I

know is what I do actually know" (Hannah Q1). These comments illustrate that students naturally form perceptions of their ability.

Despite recognising that achieving good grades provided a motivational impetus, there was also awareness that the grade secured on an assessment task is a reflection of their performance at a given time, and not necessarily a true indication of their ability.

> I do think that it's important for you to try and gain Excellence but I don't think that these marks really show how much you know. (James, Q1)

During one of the focus group interviews, students were given a scenario where they had received a lower grade than their self-perceived levels of understanding. Students reported that they would be frustrated about getting a lower grade than they felt they should have. Moreover, their comments indicated awareness that an initial grade can be altered either through a 'resit' opportunity for the internal Standards, or in the end of year examinations for the external Standards.

Andrew:	If you understand something more, then you want to try and do as well as you can.
Researcher:	So if you got an Achieve and in actual fact you thought you understood the material better?
Raewyn:	I'd resit.
Kate:	Yeah, I'd want a high mark.
Raewyn:	I don't like settling for Achieve if I know I could have done better.

Similar views were expressed by students in the other focus group.

Sally:	Well I would be frustrated. Because I got a Merit in this last one so I am resitting it on Friday because I want to get an Excellence, because I made a stupid mistake on the Excellence question and I know I know it. So I am going to resit it. So that does frustrate me.
Researcher:	So you think it is good that you get the opportunity to resit it?
Megan:	Yeah that's really good.

Students were also given a scenario where they received a grade that was higher than their perceived level of understanding. Whilst an attractive outcome for some, students expresses reservations about being awarded a grade that was not necessarily an indication of their true depth of knowledge in the subject.

> I would feel relieved, but it wouldn't be like 100% like "yes". I think I would go like "I got it" but it's kinda like an empty win in a way. (James. FG3)

> I do not like to do well in a test if I just guessed my way through it, because I know then that I only passed on luck. I like to understand what I'm doing. (Sally, FG3)

4.6.6 Classroom assessment culture.

In order to gain a sense of the assessment culture within the classroom, students' perceptions of the classes in which the assessed work was returned to them was examined. Students expressed mixed reactions to getting assessed work returned. A number of students voiced negative reaction to assessment in general: "I am always nervous. I don't like tests" (Sally, FG3). Others were more positive about having assessed material back: "I love getting tests back. It helps me heaps; but I like keeping them" (Raewyn, FG3). For some students, their reaction was linked to the perceived level of difficulty for the assessment: "I like it if I know that I have done really well, and other people found it hard" (James, FG2), suggesting that some students' reactions are influenced by a performance orientation (Dweck & Leggett, 1988).

A number of students reported that they were concerned about their teacher's reactions to their grades.

> I was just thinking what the teacher was going to say to me. I don't like it when she doesn't like me. (James, FG3)

> It is a bit disheartening when she says: "You are not going to get into maths next year". (Sally, FG3)

This last comment by Sally refers to an episode that occurred while I was observing Sally's classroom. It appeared to me that the teacher had made the comment to try and motivate Sally to work harder. Unfortunately this was not the effect.

All students appreciated the prompt return of assessed work.

> Last year you would get tests back after a week and a half and it was like...what test was that. (James, FG3)

Students received their assessed scripts back individually during a normal mathematics class. No grades were read aloud, however, during one of the class observations the teacher wrote on the board the number of students who received Not Achieved, Achieved, Merit and Excellence grades. When discussed in the focus group interviews, students expressed appreciation that individuals were not identified to the whole class. Some recognised that it provided a general profile of achievement throughout the class although its value as a practice was limited: "It lets you know where you are, but it is not really educational (James, FG2). Other students found that it had merit as a motivational practice.

> It is sort of like a wake up call though. So the people in the Not Achieved should probably do something about that. (Megan, FG2)

One student noted that her primary teacher used to read out everybody's results after a test. She clearly had bad memories about this practice and said that she wouldn't like it if her personal result was read out to the class. Another student commented that they felt their assessment information was personal, so preferred not to show their test score to anyone.

> I like to keep my marks to myself if possible because I don't feel that it's anyone's business who got what. (James, FG2)

During one of the class observations the teacher compared her current class with last year's class. Later in the class episode the teacher had remarked that she felt they were a more able group than their grades indicated. When this observation was raised with the teacher after the class, she felt that this would act as a motivational prompt. Recounting this episode in the focus group interview, the students reported that they did not appreciate her initial comment, and felt that the teachers were in competition to see who had the 'best' class. Her later comments were perceived as a much more positive and motivational.

James: She did say that she was disappointed because she thought that there were a lot of intelligent people in the class who she knew could have done better.

Sally: Yeah that was afterwards. She worded it differently and that sounded better.

James: That did sound better and I was like "oh maybe that was me". I'll try harder.

4.6.7 Pressures on pedagogy and learning strategies

A small number of students identified that deep-achieving learning strategies (Biggs, 1987), and associated pedagogies focusing on understanding, were sometimes compromised by time constraints and the number of assessments in the mathematics course. In the following extract from one of the focus group interviews, Megan expresses her frustration at what she characterises as 'teaching to the test' at the expense of a deeper understanding of the mathematics. This extract illustrates that pedagogical practices concentrating solely on preparing students to pass assessment may not be suited to all students' preferential learning strategies. Megan's comments are referring to a revision session immediately prior to an assessment event.

Megan: I reckon that since we have had so many assessments in maths class I guess… like I think…like I feel that I am learning something but I feel that I'm learning it because I'm doing the assessment. So that I'm learning how to take the assessment instead of actually learning the material that I'm going to use at the end of the year. Like it might be the same thing but I don't really know that.

Researcher: So what you are saying is that you are learning it because it is going to come up in the assessment that you are going to have.

Megan: Yeah like she will give us a worksheet and it will be exactly like the assessment that we have the next day, like the same questions but different numbers or something. And so I guess you are learning how to take the assessment rather than, like, thinking logically I guess.

Megan also perceived that Support Maths was structured to help students pass the assessment rather than gaining mastery of the underlying mathematical material.

> And then there is Support Maths that you have to go to after you take the assessment and it will be like "this is what you have to do to get it right" and that's all that really matters. So it is not really whether you understand it, it is what you need to do to pass. Like I know it is really hard because in any school system I'm sure that they are trying just to get people to pass but I don't really feel like I'm learning it and retaining it rather than regurgitating it. (Megan, FG2)

Data from the focus group interview with the teachers also identified a perceived pressure resulting from lack of class time, and that this affected teachers' pedagogical practice. The teachers commented that many periods were lost to other school activities and, under these conditions, teachers concentrated solely on the assessment criteria specified in the Achievement Standard. Although all teachers acknowledged that, under these circumstances, they lost opportunities to explore the mathematical content in more depth, focusing on the assessment criteria was not seen as a major concern as long as the curriculum and assessment criteria are closely aligned (Biggs, 1998).

> Mr Smith: At the end of the day, this term they have to know this, this and this to pass the test. That is what I am teaching. I would have liked to have done all sorts of stuff. Like in sequences and series there is all sorts of good stuff there.
>
> Ms Brown: I trust that the test assesses what they need to know. So it is not a test that is about something that is completely random. Somebody decided prior to me teaching it that this was important, so they put it in the test so I will teach what is in the test.
>
> Mr Smith: So you rely on the assessments to be subscribing to the curriculum.
>
> Ms Brown: Yeah and to reflect what kids should know.
>
> Mr Smith: So if you teach to the test then it should be the same as teaching to the curriculum
>
> Ms Brown: Yeah. Well you would hope that it was a process of continual change, that we were improving all of the time and thinking about what kids should know and learn. But as a teacher all I can do is trust that someone has thought about that reasonably deeply before.[12]

[12] The Mathematics in the New Zealand Curriculum dates from 1992 and is in the process of being updated through the Curriculum Stocktake.

4.6.8 Conceptualisations of assessment

Teachers' perceptions

As identified in chapter 2 teachers' conceptualisations of the role of assessment in learning influences how assessment information is portrayed and used in the classroom by both teachers and students. When describing their views of the role of assessment in learning, the teachers in this project expressed a variety of conceptualisations of formative assessment. The balance of comments indicated that they viewed formative assessment opportunities as primarily helping students to identify current achievement levels and the next steps in learning. All three teachers identified assignment work as a formative assessment opportunity, however there was some disagreement on whether classroom tests offered during the year were also formative assessment opportunities for the students.

Ms Brown: Their assignment shows what they know and what they need to do, so it should just be about them. Whereas their hard core real test stuff is for us. The test is for me, the assignment is for them.

Mr Smith: The first algebra test exposed their lack of understanding in a lot of cases and hopefully made them realise that, if they want to do well at the end of the year, then they have got a lot of work to do. It is helping them in the end realise that they have a long way to go.

Researcher: So the purpose of formative assessment is mainly to help them?

Mr Smith: I would say almost all of it. Certainly some of it is helping us know where they are and what areas to help them focus on but primarily at that stage it was a great way, from my perspective, to say to them "you guys have got a lot of work to do".

Additionally, Ms Clarke indicated that students achievement levels provided her with a mechanism to examine the effectiveness of her teaching, and can therefore be formative for her as well. This suggests a link between formative assessment and professional growth for teachers.

Ms Clarke: I look back at the end of algebra and I might see that they are all horrid at algebraic fractions so perhaps I need to rethink about how did I teach it or maybe I need to go over it again. It is something to investigate. If all of them didn't do well in a question then maybe I could have used a different approach.

Although the teachers could recognise assignments as formative assessment opportunity for the students, they noted that not all students saw them as a valuable learning tool. The teachers perceived that, in many instances, students are not motivated to complete assignments conscientiously because there are no 'credits' attached.

Ms Clarke: They don't seem to care about assignments. They do it five minutes before it is due.

Researcher: So they don't seem to see it as a valuable learning tool.

Mr Smith: Some of them do. There are a handful who will endeavour to do their best, regardless of what you give them. But there are others who will discriminate between those that count and those that don't.

Ms Clarke: They are after credits and that is it.

Mr Smith: Assignments don't count. I have got a couple of students in my class who are doing really well in assessments, always getting Merit or better, but hand in hopeless assignments because they have learned to discriminate between what counts and what doesn't.

Ms Brown: I mean they should really do the assignments but I think what happens, especially for Unit Standards, is that the test becomes the formative and the resit becomes the summative.

Students' perceptions

An examination of the research students' data reveals that their understanding of the terms 'formative' and 'summative' were underdeveloped. They all identified the end of year examination of the external Achievement Standards as summative, but when asked to describe the nature of the assessments that they had currently sat so far this year there were some difference of opinion. Some students identified the internal assessed Achievement Standards and Unit Standards as being formative as they generally had a number of resit opportunities for these Standards. However, there was little recognition that the 'practice' external Achievement Standards held any formative potential. Additionally, no student identified the unit assignment, usually based on previous assessment tasks, as a formative assessment opportunity.

As identified in the AEQ, although students recognise the assignments as valuable and challenging, they generally do not expend more effort leading up to an assignment

being handed in. Evidence from the focus group interviews concurs with these findings. Despite recognising that doing well in the assignment would benefit their preparation for the assessment task, most students cited conflicting pressures on their time as the main reason why they do not spend a lot of time completing assignments to a high standard. Nearly all students commented on their general assignment and class-work load across their subject, and more specifically about the number of assessments they perceive they have.

> It seems that we have an assessment almost every week in one of our subjects. (Hannah, FG2)

Additionally, there was no student evidence to support the teachers' view that students are not interested in assignments because they 'don't count'. Indeed the fact that assignments do not have 'credits' awarded was not mentioned by any of the research students as a reason why they did not expend a lot of effort doing the assignments.

4.6.9 Students' perceptions of ability

Literature suggests that for students to be disposed to use feedback in a formative manner, they should believe that their ability is 'unstable' rather than 'stable' (Weiner, 1979), and that effort, for example in the form of increased study, will improve their mathematical ability. Data from this project reveals that most students perceive mathematical ability as unstable. Students argued that although people have a natural mathematical ability, increasing the amount of time and effort spent on studying mathematics will improve that ability, and hence achievement. However, their comments suggest that they perceive that most people have a natural upper limit to their mathematical ability

> I think that to a certain degree some people will always be better than others., but I think it is more complex than that, and if the motivation and dedication are there then anyone would be able to learn and do well in mathematics this year. (Raewyn, FG3)

> Like I know for me, like I know if I had have studied on my mock exams I would have done way better, but I know I am never going to be like 'miss math' or anything. I think there is a line. Like you can practice and practice but you can only get to a certain level. (Megan, FG3)

> I think people have a natural ability to a certain level, but everyone has to work a bit if they want to achieve higher than the level they are previously at. Maths is not something that is

> hardwired into your brain. It's a skill that has to be learned and developed over a long period of time. I think you progress with increased motivation and work effort, but I don't think it is possible to learn all maths. There is only so much you can take in at any certain time, and if you spent too much time on maths I think your enthusiasm would drop. (Sally, FG3)

Kate, a very able student with high expectations, identified a clear link between the amount of effort she puts in and her overall achievement levels

> For me, if I do well, it is because I have studied and kept up in class and done the homework and assignments. If I don't do very well then I wouldn't have studied very much, or done the work. (Kate, FG3)

4.7 Conclusion.

This chapter has presented the qualitative and quantitative data for the current study. The data were organised into five broad sections. This chapter initially provided a synopsis of the observations conducted in the three mathematics classrooms involved in the study. The purpose of this was to help situate the reader into the current research. It then presented the results from the AEQ. The purpose of this was to establish the representativeness of the research sample and help to establish the trustworthiness of the qualitative data that followed. The final three sections of the chapter were organised around the three aspects of Sadler's (1989) theory of formative assessment that form a theoretical and organisational lens to view the data. The data have been presented with limited discussion. A fuller discussion of the data in light of the research discussed in chapter 2 occurs in the next chapter (chapter 5).

5 DISCUSSION AND CONCLUSIONS.

5.1 Introduction

This chapter brings together key findings in relation to the research questions and discusses these with reference to the literature. Three factors that provided the impetus for this study have been identified. All three of these factors have potential implications for existing assessment practice. Firstly, contemporary theories of learning have reconceptualised the role of the social environment in students' learning. Secondly, despite mathematics education being informed by sociocultural theories of learning, teachers' assessment practices are slow to align with changing perspectives of how students learn mathematics. Lastly, the introduction of an SBA high-stakes summative assessment system in New Zealand has provided an opportunity for teachers to reflect on their existing assessment practice.

Research evidence suggests that effective use of assessment promotes effective learning (e.g., Black & Wiliam, 1998a; Carr et al., 2000; Crooks, 1988; Natriello, 1987; Sadler, 1989). However, a tension exists in practice between the summative and formative purposes of assessment. For assessment systems to provide both formative and summative information the system needs to be designed primarily around the formative purpose (Black & Wiliam, 1998a; Brookhart, 2001; Sadler, 1989, 1998). However, institutional demands for accountability and high stakes can result in summative assessment dominating the assessment process, resulting in formative practices being compromised (Biggs, 1998; Black, 2000; Black & Wiliam, 1998a, 2006; Crooks, 1988; Harlen, 2006).

To explore the potential for NCEA to provide effective formative assessment in mathematics classrooms, a theoretical framework developed by Sadler (1989) was used to examine key aspects of the use of NCEA assessment systems within the mathematics classroom. Sadler's framework consisted of three key characteristics necessary for assessment systems to be considered formative. He stated these characteristics from the students' perspective, reflecting the notion that for assessment

information to be considered formative requires particular knowledge and actions from the student. Sadler's three key components require that the learner has to:

> (a) possess a concept of the *standard* (or goal, or reference level) being aimed for, (b) compare the *actual* (or current) *level of performance* with the standard, and (c) engage in appropriate *action* which leads to some closure of the gap. (p.121, emphasis in original)

Consistent with the premise that assessment is only formative if the student purposefully engages with the formative feedback, a case study was conducted examining students' perceptions of the formative potential of NCEA within mathematics. It is important to note that the findings discussed in this chapter are drawn from one school. As such, it is acknowledged that the findings of this project are a subset of the possible range of perceptions of secondary school students and teachers.

From Sadler's (1989) theoretical framework a series of three research questions were established for the current study:

1. How do students derive understandings of the assessment criteria they are working towards and how does this impact on their learning?

2. What feedback do students receive from in-class assessment events and in what ways do students interpret this feedback?

3. In what ways do current assessment practices impact on students' engagement with feedback and subsequent learning behaviours?

This chapter draws on the results of the current study to address these research questions. Although findings from this study are discussed under separate headings, it is noted that in reality, students' formative use of feedback is a complex process relying on the interplay between all three areas. Finally, the potential limitations from the current research, and some suggestions for future research studies are discussed.

5.2 Understanding assessment criteria

The first characteristic identified by Sadler (1989) necessary for the formative use of assessment is that students and teachers must have a shared understanding of the assessment criteria they are working towards. Using a sociocultural perspective of learning, it can be argued that these understandings are developed within the social structures of the classroom environment, and indirectly through the school environment and wider community, through interaction between students and their teacher, and amongst students themselves. Knowledge of the assessment criteria initially rests with the teacher, but ideally the students will internalise the criteria and monitor their own progress as they work towards them. Accordingly, interactions that support the development of students' understanding of assessment criteria should be integrated into the normal classroom teaching of the unit of work.

A key finding from this research is that the research students in this study had a limited and underdeveloped knowledge of the assessment criteria for the assessment Standards they were working towards. This limited understanding restricted students' independent use of self-assessment strategies to monitor and regulate their learning. Furthermore, an underdeveloped knowledge of the assessment criteria potentially reduced students' ability to understand and engage with feedback provided to them.

Students in this study developed their knowledge of the assessment criteria predominantly from two sources: The structure of the unit of mathematics work and the provision of the 'I can do' sheets. These two areas are examined separately.

5.2.1 Structure of the unit of work.

Within the classes involved in this study it was apparent that the teaching and learning of the units of mathematical content were structured around the assessment criteria for the assessment Standards. Despite the fact that Unit Standards are included in the Y12 Trad Maths course, it was noticeable, in reviewing the data, that nearly all of the conversations, with either the research students or teachers, were about Achievement Standards.

Teachers reported that in structuring their lesson sequences, they started with content consistent with the assessment criteria at the Achieve level. Once the teachers perceived that the students were ready to progress they advanced to content based on the Merit level assessment criteria (see section 4.4.1). The assessment of students' readiness by the teacher was informally conducted across the whole class rather than on an individual level. Teachers gained a sense of the whole class's understandings based on the social interactions that occurred within the classroom, for example, the number of questions asked by students, either in whole-class teaching, or when students were doing set work out of the text book or worksheet.

Following this model of progression through the levels, a unit of content should conclude with material based on the assessment criteria at the Excellence level. It was recognised by the teachers, however, that, in many instances, Excellence material was not actively taught; the responsibility to master Excellence material often being left up to the individual student.

> Sometimes I will just do the Achieve and Merit and I will say "look exercise such and such is covering the Excellence. Sorry we don't have time so you are going to have to do it at home". So if they really want to get Excellence then they are going to have to get off their bum and do it at home. (Ms Clark)

Excellence level material generally requires higher level thinking skills, and, arguably, higher levels of formative interaction between the teacher and students. It is acknowledged that this level of interaction is often difficult given the number of students and range of ability in a typical Y12 mathematics class. From a sociocultural perspective, greater use of structured peer interaction between students of similar ability could potentially improve the learning of Excellence level material within a community of practice. Additionally, if students are working more collaboratively, it potentially can free up some of the teachers' time to work with groups of students as well as individuals.

The majority of the research students recognised the intent of the structure of the units of mathematical content, and identified that the implicit structure supported them in gauging the level of difficulty required at specific levels of the assessment Standards. One notable observation from the classroom teaching is that although teachers

generally tended to talk about a particular mathematical problem being at a specific level of the Standard, they were not proactive at explicitly linking characteristics of the question that relate to specific assessment criteria at that level (see section 4.2.2 & 4.4.1). The way in which insufficient links impact on students' understanding of assessment criteria is a concern. Students may come to believe that 'harder' questions are at the higher achievement levels of the Standard, but may not develop links between specific aspects of the questions and the relative assessment criteria from the Standard.

While acknowledging that the assessment criteria should not drive the enacted curriculum, a suggested pedagogical strategy to strengthen the links could be to start from a specific assessment criteria and then give an example of a mathematical problem that illustrates that criteria. For instance, in section 4.2.2 an example was given where the teacher provided a difficult quadratic to solve without specific reference to the relevant assessment criteria. It would be preferable for the teacher to start from the assessment criteria (i.e. the use of the quadratic formula to solve quadratic expression) and then give a mathematical problem making explicit links back to the criteria. It is not intended to suggest that this pedagogical strategy should be used in all lessons. There will be many lessons when the content will be focused on developing the mathematics skills to solve problems. However this level of detail, on occasions, would be beneficial to the students, since a lack of specific knowledge of the assessment criteria effectively limits their ability to monitor and regulate their progress towards the assessment Standard.

One of the teachers expressed some concern that structuring the unit around the levels of the assessment Standard might be considered an improper practice: "You are allowed to do that aren't you? That's good? Or is that bad?" (Ms Brown). Ms Brown defended the practice arguing that:

> I trust that the test assesses what they need to know. So it is not a test that is about something that is completely random. Somebody decided prior to me teaching it that this was important, so they put it in the test so I will teach what is in the test. (Ms Brown)

The notion of 'alignment' between curriculum and assessment implicit in Ms Brown comment concurs with research suggesting that such alignment is feasible and

beneficial to both teaching and learning (Biggs, 1998; Clune, 2001; Linn & Herman, 1997; Porter & Smithson, 2001). Birmingham (2001) concluded that a strong alignment between assessment and teaching allowed for more student-focused teaching and learning, resulting in higher success rates. Similarly, research by Preece and Skinner (1999) and Wilson and Floden (2001) has shown that the clear objectives inherent in SBA help to focus teaching and raise expectations for teachers and students.

The structuring of the units of work around the levels of the Standard may potentially encourage some students to restrict their learning to material based on the Achieve level of the standard. Under NCEA, students are required to demonstrate that they have sufficiently mastered the assessment criteria for given achievement levels of the Standard. In a mathematics assessment task, the first few questions are primarily designed to assess students against the Achieve criteria. However, in most instances, later questions primarily designed to assess Merit and/or Excellence criteria, also gather 'replacement' evidence towards the Achieve criteria. For example, in level 2 algebra, one of the Merit skills requires students to solve quadratic equations that are given in an expanded, rather than a factorised form. In solving this form of question students must use one or both of two Achieve skills, namely manipulate algebraic expressions and solve quadratics given in a factorised form. A student may incorrectly factorise the quadratic therefore preventing them from getting the correct answer, but may provide answers consistent with their incorrectly factorised expression, therefore demonstrating the achieve skills of solving quadratics given in a factorised form.

From the current study, both the teachers and the students were critical of restricting learning to just the Achieve material.

> Some people could tend to be more lazyish and just learn the Achieve material and so they just pass. So they don't actually learn the Merit and Excellence stuff. I think we should just learn it all. (Sally, FG2)

Students were also aware of the potential for replacement evidence from later questions. As Tracy notes: "And they probably wouldn't try them so if they got one of the Achieve ones wrong they couldn't use these other ones to fill in." These

findings are consistent with Meyer et al.'s (2006) research which identified that students whose motivation could be characterised as 'doing my best' were more likely to achieve better grades than those students whose motivation could be characterised as 'doing just enough'.

5.2.2 'I can do' sheets

Students in this study can potentially develop knowledge of the assessment criteria through the use of 'I can do' sheets. These sheets list the assessment criteria for Achievement Standards[13] in student language and are incorporated into the course booklet (the 'Blue Book') given to the students at the start of the year. The 'I can do' sheets were developed nationally by the New Zealand Association of Mathematics Teachers, with schools being encouraged to adapt them to best suit their students.

Data generated across this project provided evidence that teachers referred to the 'I can do' sheets during the normal teaching of the mathematics unit, but the incidence of this was relatively low. Although teachers asked students to get out their personal copies of the 'I can do' sheets, many students either did not have the sheets with them, or did not bother getting them out. During the focus group interview a number of students appeared embarrassed when asked about why they did not use the 'I can do' sheets. Although the majority of students recognised the potential of the sheets, most tended not to use them. In some instances they did not know where the 'I can do' sheets were, whilst in other instances they did not find them easy to use (refer to section 4.4.2).

Significantly, even though the 'I can do' sheets were written in a language designed to be easily understood by students, the majority of the research students found it difficult to judge their competence against the assessment criteria detailed in them.

> It would just say "I can draw a linear graph" and I would be just like "yeah, I suppose I can, but what exactly is a linear graph". It's not like "ok exercise 13.5", I've covered that so I can flip back to 13.5 and say "Oh yeah I can do that". (Raewyn, FG2)

[13] The school does not use 'I can do' sheets for Unit Standards.

These findings suggests that the research students are not developing sufficient understanding of the assessment criteria from their normal classroom work to be able to recognise the criteria from the 'I can do' sheets, effectively limiting the independent use of the sheets for self-assessment.

A small number of students indicated that they did sometimes use the 'I can do' sheets as a tool for assessing competence levels but that they did not follow up on any required remedial action.

> For me it is just like: "Can I do this?" … and I'll just go 'tick'. "Can I do this?... no". So I just tick 'no,' and carry on to the next one. I don't try and go back and do it. (James, FG4)

Only one student consistently claimed to use the sheets independently as a self-assessment tool to direct them towards material they need to study.

Despite self and peer-assessment being identified in the literature as an integral component of formative assessment, neither the research students nor the teachers could cite instances where the 'I can do' sheets were used as a structured in-class self- or peer-assessment tool. This finding concurs with research that identifies that a focus on student self- and peer-assessment is not common practice at secondary school level, even amongst those teachers who take assessment seriously (Black, 2001; Black & Wiliam, 1998a; Crooks, 1988).

With respect to the current project, small modifications to the 'I can do' sheets' could increase their effectiveness in developing students' understanding of assessment criteria, and supporting self- and peer- assessment opportunities. For example, the addition of a column where a peer could sign off that they have seen work that satisfies particular assessment criteria. Additionally, there are currently no 'I can do' sheets for Unit Standards. If these were developed it would help to embed the use of 'I can do' sheets as a normal practice in all units of work, rather than just those assessed by Achievement Standards. Teachers could also plan for structured classroom sessions where students actively use the 'I can do' sheets as self- and peer-assessment tools.

The provision of reference exercise numbers on the 'I can do' sheets alongside each of the assessment criteria was cited by a number of students as a potential improvement. Although this would allow students to judge their competence against specific mathematical problems, the provision of reference exercise numbers alone will not necessarily improve student understanding of the assessment criteria. To be effective, explicit links between the mathematical questions in the reference exercises and the respective assessment criteria must be made. Developing the links between specific assessment criteria and exercises linked to those criteria could be done as a structured class activity. Activities could also be designed where students are given specific assessment criteria and are required to write mathematical problems consistent with the given criteria. These mathematical problems could then be given to their peers to see if they can identify the assessment criteria they are designed to exemplify.

Modifications in the design and use of the 'I can do' sheets will result in students gradually being exposed to the full set of assessment criteria and the rules for assessing them. By requiring students to think about and apply criteria to their own work and the work of others, students build up a body of evaluative knowledge, bringing them into what Sadler (1989) refers to as the 'guild' of people who are able to determine quality using multiple criteria. This transfer of some of the responsibility for making decisions from the teacher to the learner encourages autonomy as students act in the role of the assessor and judge of quality (Black & Wiliam, 1998a.; Black et al., 2003; Crooks, 2001; Gipps, 1994; Sadler, 1989, 1998). This increased autonomy helps contribute to a shift in the conceptualisation of the roles of the teacher and student in assessment and learning (Black et al., 2006) as students become "insiders rather than consumers" (Sadler, 1989, p. 135) of the assessment process; a view consistent with the sociocultural perspective of knowledge being socially situated.

5.3 The Nature of Feedback

The second characteristic identified by Sadler (1989) as important in formative assessment systems is a mechanism for students to compare their current level of performance with the Standard required. Apart from self- and peer-assessment and informal teacher feedback, the main activities designed to provide this comparison are assignments and assessment tasks marked by the teacher. To be considered formative, feedback identifying discrepancies between the students' actual and desired achievement levels will need to be provided. The provision of feedback alone, however, will not necessarily lead to its formative use by students (Sadler). Students must, in the first instance, read and understand the feedback, and, in the second instance, they must purposefully engage with the feedback to positively impact on their learning. This latter point will be discussed in more detail in section 5.4. The current section examines the extent to which students read and understand the feedback they receive.

The main finding from this research project is that the research students read and valued written feedback provided on assessed work. Notably, these findings conflict with the views expressed by the research teachers who felt that students did not generally read the written feedback offered, and therefore did not value, or use it, in a formative way. This perception of the teachers, coupled with a school assessment policy on resubmission, resulted in teachers providing differential levels of feedback between assignments, internal and mock external assessment tasks. The practice of providing limited written feedback on internally assessed Standards is a barrier to the formative use of assessment information becoming a normal part of the assessment process in the mathematics classroom.

5.3.1 Accessing feedback

Students in the current study read and valued written feedback offered prior to, and subsequent to, an assessment event. A number of sources of data converge to make this conclusion. Data from the Assessment Experience Questionnaire for the entire Y12 mathematics cohort identified both individual questions and factor scores relating to feedback that were statistically significant. Students strongly agreed with the

statement: 'I read the feedback carefully and try to understand what the feedback is saying' (I25, 4.0), and rejected the statement that they tend to only read the marks (I30, 3.7).

Students also indicated that they found the feedback valuable. They identified that feedback helped them to understand why they received the mark they did (I22, 3.7), helped them to understand the content covered in more depth (I20, 3.8), and also how to improve in subsequent assessments (I21, 3.9). Overall, the factor score for the 'Quality of Feedback' was 3.4, with the 95% confidence interval showing that the score was statistically higher than the neutral value of 3. This indicates that students perceived they received plenty of timely feedback that was understandable and useful; explaining grades, misunderstandings and improvements.

Students' ability to readily recount episodes of particular feedback during the focus group interviews provided further evidence that they read the written feedback they had been given. Students frequently described the sorts of feedback they received, illustrating the discussion with recollections of particular pieces of feedback, and their impressions of its usefulness (see section 4.5.1). These results are encouraging, and challenge the research teachers' expressed doubts about whether students read the feedback.

5.3.2 Students' feedback preferences

One of the key characteristics of effective feedback identified in the research literature is the need to focus on the specific details of what needs to be done to improve (Black & Wiliam, 1998a; Crooks, 1988; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Tunstall & Gipps, 1996; Wiliam, 1999). Results from the current project concur with this observation. All students reported that the most useful feedback identified errors and provided a scaffolded (Wood et al., 1976) solution method for obtaining the correct answer (refer to section 4.5.1). The following quote exemplifies the views expressed by the students.

> He tells us exactly where we went wrong, how to change it, and what to do next time. My maths teacher shows us exactly where we went wrong. It's very helpful. I wish all teachers would. (Raewyn, Q1)

Despite the generally positive view of feedback reported by the research students, there were some instances where concerns were expressed about feedback perceived to be of limited value. For example, a number of students highlighted the shortcomings of what could arguably be considered 'traditional' marking. "Ticks and crosses are not very helpful at all, and neither is just an 'M'[14] on the front" (Sally, Q1). These views concur with Sadler's (1989) argument that such "dangling data" (p. 121) cannot be considered effective feedback. Although a summary grade provides students with knowledge of their result, it is too deeply coded to describe the appropriate action needed to improve learning. Moreover, 'crosses' do nothing to indicate the errors in reasoning that led to the wrong or incomplete answer, and do not suggest a strategy to improve their learning and performance in the future.

Feedback does not need to be extensive in volume before it is considered useful by students. In the following extract, we see how a small addition to the feedback provided would have made it more effective for Sally.

Sally: That's why I questioned my graphs. She had ticked my table method and crossed my graph and I was like ... "what's going on" because my table method was an exact replica of my graph. I was like "Miss, I don't understand. Why did I get this wrong" and she said "you didn't extend it".

Researcher: So perhaps if she had written "extend your graph" next to it?

Sally: Yeah, yeah. I would have been like "oh I get it".

The provision of 'answer only[15]' feedback was not commonly observed in the current study, but one particular episode was reported by the students. The students' negative reaction to this episode reinforces the value that they place on scaffolded feedback that encourages them to actively engage with the feedback to work out a solution method.

Megan: I like to know where I went wrong, and what I have to do to fix it, rather than just being given the answer.

---

[14] The 'M' referred to in this quote represents a grade of Merit in the Achievement Standard.
[15] 'Answer only' feedback refers to feedback where students are supplied with only the correct answer to a question rather than the steps leading to a correct answer.

James: You see where to go, and all of a sudden it kinda clicks into place, and you are like, "oh I see it is just that simple process".

Due to the nature of most mathematics questions asked in a secondary school environment, a number of specific task-related feedback practices can be suggested. Identifying the student's error, or undeveloped understanding, and providing the next correct line of working is effective for both conceptual errors—where the student does not know what to do next—or procedural errors—for example incorrectly factorising a quadratic. Additionally, the provision of feedback designed to cause 'cognitive conflict' is effective in highlighting to students misconceptions in their reasoning.

The concept of scaffolded feedback is consistent with contemporary sociocultural learning theories that hold that students' knowledge and understandings are not wrong, so much as 'partial' (Sfard, 2003). Learning, therefore, is seen to help students build on these partial understandings, and feedback has a significant role to play in this. Learning is enhanced when these partial understandings are recognised and commented on. Such task-oriented rather than ego-oriented feedback concentrates students' attention on the specifics of the mathematical problem they are trying to solve (Tunstall & Gipps, 1996).

5.3.3 Teacher feedback practices

One of the notable findings of the current research is the differing opinions between the teachers and the students regarding optimum feedback practices. When interviewed, the teachers expressed a clear preference for providing whole-class oral feedback coincident with the return of the assessed scripts to students. The teachers reasoned that, as much of the feedback is common to a large number of students, feedback to the whole class is an efficient and effective pedagogic practice. "It is usually the same things coming up again so it is easier if I write it up on the board" (Ms Brown). Despite her preference for providing oral feedback, Ms Brown did express some concern that her existing practice might be seen as 'poor teaching': "I will say it to the whole class. I suppose that is bad. All teachers should spend time writing it down" (Ms Brown).

Although teachers identified marking workloads, the desire to return assessed work quickly, and the schools' resubmission policy as factors affecting the amount of written feedback they offered, doubts about whether students read written feedback was the most significant factor associated with their feedback practices. "I just find that a lot of people just don't want to read written mathematical things if they got it wrong in the first place" (Ms Clark).

The research teachers identified that the whole-class oral feedback sessions were often difficult. Trying to provide the oral feedback while simultaneously trying to manage student behaviour was challenging.

> It is something that I do find difficult, how to give that feedback. How to go over a test that would make the class listen and want to do it, without having half of the class muck around. If you let them go over it themselves then you will always have about 15 kids who will just have a conversation, and all of those management issues. (Ms Brown)

Classroom observations confirmed that the whole-class oral feedback sessions were strained, with many of the students appearing not to listen to the teacher during these sessions (see 4.2.3). Although some students were clearly off-task, others were conferring with their peers about their assessed work. It cannot be assumed, therefore, that students are not interested in feedback per se simply because they do not appear to listen to oral feedback or ask questions of the teacher in whole-class settings.

The school currently operates a resubmission assessment policy that effectively limits the written feedback that can be provided on internal Achievement and Unit Standards. The rationale behind this policy is that it allows students who have made minor procedural errors, preventing them from securing the next grade, to identify and correct the errors themselves. This practice is designed to prevent students having to resit the entire assessment unnecessarily (see section 4.2.3). Although the resubmission policy does not explicitly state that teachers should limit the feedback they provide on all students' scripts, the research teachers reported that this policy did affect the amount of written feedback they offered on assessed scripts. It is certainly time consuming for teachers to read through a question and decide whether the

student will be offered a resubmission opportunity before deciding to provide any written feedback.

Given that reassessment opportunities are only offered to a small number of students, the policy of providing limited feedback on internal Standards could usefully be reviewed by the school. Although the policy is well intentioned, it does limit the formative information offered to the majority of students who are not participating in resubmissions. To avoid a blanket restriction of feedback on all scripts, one possible solution is to mask feedback on those questions requiring resubmission. Using 'post-it' notes to cover feedback comments and photocopying could be feasible strategy for the relatively low number of occurrences of resubmission. This practice would allow teachers to adhere to the schools resubmission policy without restricting the amount of feedback offered to other students.

The research teachers in this project could be characterised as reflective practitioners and were clearly interested in finding the best way to provide feedback.

> It would be nice to know if they read it, that's true, and especially whether they find written, or personal, or group, or board work as the best delivery of the feedback. I'd be very interested in what would be most effective in terms of our delivery. (Ms Brown).

Providing written feedback is time consuming for teachers—a factor that is acknowledged by both students and teachers. Teachers tend to prioritise the time and effort spent on their classroom teaching over the time on providing written feedback. Despite the teachers' clear preference for providing oral rather than written feedback it cannot be assumed that students will engage with generic feedback offered in an oral format. Given that teachers have limited time to teach and assess their students, and given the value students place on written feedback, it could be suggested that teachers might want to re-evaluate their allocation of time to better support formative assessment practices.

5.4 Student Engagement with Feedback

The third characteristic identified by Sadler (1989) as necessary for an assessment system to be considered formative is that students must engage in appropriate action to develop corrective strategies in their learning. The knowledge of assessment criteria, and the provision of feedback identifying the gap between actual and desired performance, are necessary, but not sufficient, to ensure effective formative assessment practices. Students must also engage with the feedback to address any perceived gaps in understanding. Despite research evidence that high-achieving students engage with feedback in a variety of integrated ways (Brookhart, 2001; Higgins et al., 2002; Hyland, 2000), there is little evidence in the research literature that the majority of students independently engage with feedback in a formative way. Furthermore, little is known about effective pedagogical strategies that encourage students to engage with feedback in a sociocultural learning environment.

The following sections examine the evidence associated with student engagement and the assessment culture in the research classroom.

5.4.1 Working with peers.

Despite students' claims in the AEQ that they use past assessments for revision purposes (I29, 3.7), evidence from the focus group interviews revealed that the majority of students did not access their assessment folders containing their past assessed work[16] to help them prepare for subsequent assessment events. Students primarily engaged with the feedback during the class period when the assessed work was returned to them (see section 4.6.2). Students reported in the AEQ that they used the feedback to go over material in the assignments (I26, 3.8), with the 95% confidence interval for the factor score for 'Use of Feedback' showing that the mean value of 3.7 was statistically higher than the neutral value of 3 on the Likert scale. Although the class periods when assignments were returned were not observed during this research, both the teachers and students reported that the classroom periods

[16] Assessed work consisted of class assignments as well as work from formal assessment tasks.

followed a similar format to when the assessment scripts from more formal assessment events[17] were returned.

The majority of students in the current research expressed a preference to go through their assessment scripts and engage with written feedback either individually, or more commonly, with their peers first rather than with the teacher. As noted earlier, classroom observations (see section 4.2.3) and teachers' concerns (see section 4.6.1) revealed that many students did not appear to be focused on following and understanding the oral feedback. Evidence from the student focus group interviews confirmed and expanded on the observation that some students appeared to be discussing their scripts with their peers rather than listening to what the teacher was saying. Although students acknowledged, and appreciated, that their teacher provided oral feedback and answered any specific questions, the majority of students preferred to work with their peers in the first instance. The following extract from Kate exemplifies a commonly held view.

> Normally if we have a question wrong we compare and look at the other's working to find out where we went wrong. If we both got it wrong then we discuss how we view the question. If I cannot see where I went wrong from my friend's assessments then I ask the teacher. (Kate, FG2)

There was a reluctance to ask for specific help from the teacher and indeed only a few said that that they ever asked their teacher for help during these sessions.

> Often I will ask my peers, especially my great friend Jenny, about a question before I ask the teacher. It seems to work well because we can fill most gaps between ourselves. (Sally, FG2)

Additionally, for a variety of reasons, there was a reluctance to use the 'Support Maths' system to address learning needs. These findings are consistent with classroom research that identifies that help-seeking from teachers is not a preferred learning strategy common to senior secondary students, especially in whole-class settings (Anthony, 1994; Bell & Cowie, 2001).

---

[17] More formal assessment events includes assessments for internal Achievement and Unit Standards and 'practice' assessments for External Achievement Standards.

5.4.2 Classroom assessment environment

The nature of the dominant classroom assessment environment affects students' engagement with feedback (Black & Wiliam, 1998a; Harlen & Deakin Crick, 2003). In order to gain a sense of the assessment environment within the research classes, and to contextualise students' comments, a number of indicative factors were explored, namely: Intrinsic/extrinsic motivation levels, deep/surface approaches to learning, learning/performance orientation of participants, and conceptualisations of assessment systems held by participants. These factors have a strong inter-dependence. Accordingly any discussion of these factors in isolation is not meant to imply that they operate separately, with no apparent relationship between them.

A number of observations can be made from the findings of the current research project. Firstly, the majority of the research students exhibited a deep-achieving approach to learning (Biggs, 1987). This approach is characterised by students who are motivated by an intrinsic interest in the content, with the aim of maximising their grades. Within the current study, the research students optimised their understanding, and therefore their grade, by strategically engaging in tasks, including assignments, class activities, and feedback.

The following findings from the current project support this conclusion. The majority of students claimed that a significant motivation to learn the mathematics was a need to understand the underlying content (see section 4.6.4). They saw this as necessary to become flexible mathematical thinkers. This deep approach to learning is reflected in the comment:

> I get frustrated when I don't understand. It is important because if you come across something slightly different to what you are used to you can usually nut it out if you understand the concepts. (Hannah, Q4)

Further evidence for a deep approach to learning can be seen in Sally's claim that she couldn't: "use methods, formula's etc. without understanding how and why they are used".

The 'achieving' subscale of Biggs' (1987) work is evident when students recognise that good understanding will help them with the short term goal of performing well in

the assessment tasks. “I like to understand so that I am able to achieve highly” (Kate). Many students also perceive that you can secure an ‘Achieve’ grade with a surface approach to learning, but to secure the higher grades you need to have a deeper understanding of the content.

> I think that in this course you are probably able to get Achieve without understanding the material fully; just do some questions here or there. In having a good understanding of the material you are able to apply it to harder level questions as Excellence often requires a combination of skills. (Kate, FG4)

Further evidence of a deep-achieving approach can be seen in the data from the AEQ. The factor score for ‘Learning from the Exams’ (3.54) was statistically above the neutral value of 3. Rather than simply adopting a surface approach to preparing for the examination, students indicated that they “learnt new things while preparing for the exam” (I33, 3.6), that they “understand things better as a result of the exam” (I34, 3.7), and that “doing the exam brought things together for me” (I32, 3.3).

Many students indicated that a deeper understanding of the content would also help with the medium and long term goals of preparing them for future learning opportunities. For example: “[T]o use it in later life you need to understand it” (Tracy), and, **“**I feel it is very important to understand as much as I can to further my education” (Andrew). It could be conjectured that these findings are a direct result of the research teachers encouraging all students to develop strong content knowledge for its own sake, rather than simply trying to do enough to ‘pass the test’. These findings are consistent with Meyer et al.’s (2006) research which identified that the most successful students could be characterised as ‘doing my best’ rather than ‘doing just enough’.

Closely related to students’ expressed need to understand the underlying content, the way in which students preferred to use feedback to reach this understanding was also highlighted in the current research. Rather than simply wanting to ‘understand’ the feedback, it appeared that students wanted to be able to *use* feedback to ‘work it out’ themselves thus regarding ‘working it out’ as integral to their learning. In other words, students want to construct their own deep understandings of the mathematics, and effective feedback has a central role to play in how these understandings are

constructed. This observation has implications for the nature and level of detail required in the written feedback and will be discussed more in section 5.6.

Although participant students recognised that engaging with feedback would improve their knowledge for future assessment events, the findings also suggest that students' expressed need to understand the feedback appeared to be independent of whether that knowledge will be assessed in future assessment tasks during the year. Their understanding and engagement with feedback appeared to be more related to improving their knowledge of the mathematics, rather than improving their chances in a reassessment opportunity. Due to the nature of secondary school mathematics assessment formats, students often had a strong sense of whether they had answered a question correctly. This was evident in the students' observed and reported actions when they got their assessed scripts back. They quickly went to the questions they thought they might have got wrong (see section 4.6.1). An implication of this observation is that students should be provided with feedback irrespective of whether they receive an opportunity to resit that assessment Standard or not. The provision of formative written feedback completes the 'learning loop' which appears to be a primary goal for these research students (Sadler, 1989).

5.4.3 Intrinsic motivation

The findings of this study reveal that the research students were predominantly intrinsically motivated to do well in their studies: "I want to do well and I know that will take work so I do it." (Hannah). Although most students recognised and appreciated encouragement from their parents, they did not perceive that they were under undue pressured to excel. Evidence of intrinsic motivation can also be seen in students' comments about getting a grade that reflects their understanding, rather than just a high grade based solely on luck (see section 4.6.5).

> I do not like to do well in a test if I just guessed my way through it, because I know then that I only passed on luck. I like to understand what I'm doing. (Sally)

For some students, securing a good grade was seen as a measure of their work ethic and self efficacy. This is exemplified in the following quote.

> If I do not pass as highly as I would have liked then I feel disappointed with myself. ….I think it is important because I like to do well and I know that I should be able to gain Merit/Excellence….I feel that, for me, getting Achieved is like failing as I like to aim for the best. (Kate)

5.4.4 Learning orientation.

High levels of intrinsic motivation amongst the research students, coupled with an emphasis on deep approaches to learning are consistent with a learning orientation (Dweck & Leggett, 1988). This conclusion is supported by data from the classroom sessions when students received their marked scripts back (see section 4.6.1). After asking their peers what grade they received, students' conversations quickly turned towards discussions about how to do individual questions. Although some students preferred to keep their grades to themselves, a majority of students were happy to share their grades with their close friends. Students did not appear to be trying to establish an informal ranked list of students' achievement levels as might be more commonly observed in a performance oriented classroom (Dweck & Leggett). Under SBA, students are no longer operating in an assessment environment where they are competing with their peers for a limited pool of marks or grades. As such, students can collaborate rather than compete, and indeed collaborative interaction with their peers was clearly expressed by the students as their preferred method for engaging with written feedback.

All of the research students appreciated the absence of a socially comparative atmosphere. Notably, the students displayed a negative reaction when the teacher compared them with last year's classes, or with the other Y12 Trad Maths classes (see section 4.6.6). The findings suggest that the students are concentrating on their individual performance rather than their relative performance with their peers, further indication of a learning orientation (Ames, 1992; Dweck & Leggett, 1988). This does not, however, preclude them from being pleased if they have done well on an assessment task that many found difficult. "I like getting the test back if everybody else found it really hard." (Andrew). This is, perhaps, a natural reaction to succeeding in a difficult task.

5.4.5 High-stakes impacts on learning and teaching.

Within the current research, only a minority of students identified that time pressures and the number of assessments negatively impacted on deep learning strategies: "I feel that I am learning something but I feel that I am learning it because I am doing the assessment" (Megan). However, time pressure was consistently identified by teachers as a factor affecting pedagogy (see section 4.6.7). Teachers commented that a number of class periods were lost to other school activities, resulting in a perceived restriction in the depth of mathematical content they were able to cover. The reduction in the time to teach the curriculum resulted in teachers focusing on the assessment criteria as specified in the assessment Standards. The teachers, however, were not overly concerned with this practice, often characterised as 'teaching to the test', since they perceived that the curriculum and assessment were closely aligned. This notion of alignment is consistent with research claiming that the clear assessment criteria inherent in SBA provides a focus to teaching, and enhanced learning for students (e.g., Barrington, 2004; Kannapel et al., 2001; Preece & Skinner, 1999; Wilson & Floden, 2001). If assessment and curriculum are closely aligned then 'teaching for assessment' will result in 'teaching the curriculum'.

Although both participant students and teachers recognised that the pressure of time exerted some influence on teaching and learning strategies, the students and class atmosphere could be characterised as calm. This is in contrast with some studies that have found that the impact of high-stakes assessment has resulted in high test anxiety amongst students (e.g., Leonard & Davy, 2001), with observed classes described as being at 'fever pitch' (Reay &Wiliam, 1999). Within the current study, neither the classroom observations immediately prior to students sitting the assessment tasks, nor the interviews with the students, revealed any abnormally heightened anxiety levels relating to assessment events.

Furthermore, the students in this research project had a developing knowledge of the difference between summative and formative assessment and were aware of feedback designed to be formative (see section 4.6.8). These observations are in contrast with research by Pollard et al. (2000) who found that students in high-stakes environments

were only aware of assessment as a summative activity, and found no evidence that teachers were communicating any formative information to students.

Deep-achieving approaches to learning, high levels of intrinsic motivation, and a predominant learning orientation suggest that, on balance, the impact of NCEA on the classroom environments in the current research is not negatively influencing either teachers' pedagogy or students' learning. While cautious that these findings are drawn from a subject-specific case study, they appear to contradict observation by Hipkins et al. (2007) that over-assessment negatively impacts on learning and teaching strategies. The findings from the current study suggest that a range of approaches to implementing NCEA in schools might result in varied impacts on teaching and learning. Indeed, the findings concur with evidence from the research literature (e.g. Gipps, 1994; Black & Wiliam, 1998a; Clune, 2001; Hipkins et al., 2004; Kannapel et al., 2001; Wilson & Floden, 2001) suggesting that Standards-Based reform, and its implied pedagogical changes, can have a positive impact on teaching, learning and achievement.

5.5 The formative potential of NCEA

This research project sought to examine the extent to which New Zealand's SBA exit qualification system can serve both the summative and formative purposes of assessment. Despite recognition that a single assessment system can serve a duality of purpose, the assessment system should ideally exhibit a design consistent with formative assessment principles (Black & Wiliam, 1998a; Brookhart, 2001; Sadler, 1989, 1998).

Using Sadler's (1989) theory of formative assessment as a framework, an examination of the philosophical and structural design of NCEA reveals a strong potential for it to serve a duality of purpose. The provision of explicit assessment criteria in the assessment Standards is consistent with Sadler's (1989) first criteria for effective formative assessment. Despite the current projects' finding that students' knowledge of assessment criteria is underdeveloped, the potential for this characteristic to significantly impact on students' learning and sense of autonomy is significant. The

knowledge of assessment criteria is a necessary first step in effective self- and peer-assessment practices (Black & Wiliam, 1998a, Sadler, 1989). These effective assessment practices support the development of students' metacognitive knowledge and learning strategies that allow students to take a more autonomous role in their own learning (Black et al., 2006). Increasing student autonomy in learning, and the resulting decrease in the reliance on the teacher to address learning needs, also has the potential to free up teachers' time to provide targeted assistance when required (OECD, 2005). Teachers therefore become one of many learning resources within a community of learners. This reduction in the hierarchical relationship between students and teachers is consistent with a sociocultural perspective of learning (Lave & Wenger, 1991).

The explicit nature of assessment criteria inherent in SBA focuses students' and teachers' attentions on important learning objectives (Birmingham, 2001; Clune, 2001; Wilson & Floden, 2001). This focus potentially enhances the effectiveness of task-oriented feedback, identifying gaps between students' current achievement levels and levels of achievement as specified by the Standard (Sadler, 1989). Task-oriented feedback that scaffolds students' development of learning strategies to address incomplete or underdeveloped mathematical knowledge was consistently identified by the students in the current research project as the preferred style of feedback.

Scaffolded feedback recognises that students' knowledge is partial and developing and encourages students to make connections between their existing knowledge and the feedback provided. Making connections has been identified as a key component of the work that mathematicians do (Boaler, 2002; Burton, 1999). Accordingly, by engaging in the work of practising mathematicians, students "progress to more complete participation in the practices, beliefs, conventions and values of communities of practitioners" (Goos, Galbraith, & Renshaw, 2004, p. 91). Such active engagement with task-oriented feedback is consistent with a sociocultural perspective of learning (Lave, 1988, Lave & Wenger, 1991).

Sadler's (1989) third condition for effective formative assessment is that students must engage in appropriate action to address the identified discrepancies between current and desired achievement levels. The provision of reassessment opportunities

within NCEA potentially encourages students' engagement with formative feedback. However, given the majority of the research students' preference for a deep-achieving (Biggs, 1987) approach to learning, the provision of formative feedback, irrespective of whether students will be reassessed on that knowledge, would appear to be beneficial to students' long-term learning outcomes.

It is argued in this research that the abandonment of a Norm-Referenced assessment system in favour of an SBA system, and the inherent reduction in the socially competitive assessment environment, may have encouraged greater peer-interaction to address learning needs, and a predominant learning orientation. This research project has consistently identified students' preferences for working with their peers in the first instance, rather than their teacher. This collaborative rather than competitive approach to learning, combined with higher levels of intrinsic motivation, are seen as preferable to a performance orientation because they encourage the metacognitive strategies that enhance students' abilities to engage in future learning opportunities (Ames, 1992; Dweck & Leggett, 1988). Therefore it can be hypothesised that NCEA supports and strengthens preferable learning orientations within a sociocultural learning environment.

It was argued in chapter 1 that the ultimate goal of formative assessment should be to teach students to regulate their own learning (Black & Wiliam, 1998; Gipps, 1994; Sadler, 1989, 1998). Increasing students' autonomy in learning has been identified as a desirable outcome of students' school experiences, as it promotes and supports life long learning skills, necessary for adult life in the 21st century (MoE, 2006a). Furthermore, increasing peer interaction in learning within a sociocultural learning environment encourages social discourse and levels of cooperation consistent with the 'Key Competencies' identified in the New Zealand Curriculum: Draft for consultation (MoE, 2006a).

While acknowledging the formative potential for NCEA, the provision of an assessment system consistent with formative principles will not guarantee improvement in formative assessment practices. Research has demonstrated that in situations where high-stakes dominate formal assessment systems there is a risk that

summative practices will take precedence over formative use of assessment information (Broadfoot et al., 1998; Pollard et al., 2000; Reay & Wiliam, 1999).

Much of the research into the effect of high-stakes assessment on learning and teaching, however, has been conducted in countries where secondary school exit qualifications are considered very high-stakes for both students and schools. Schools are often held accountable for the achievement levels of their students, with possible financial consequences. Similarly, students' entry into their preferred tertiary institution is often dictated by their relative position on a rank list based on marks from secondary school exit qualifications. In contrast, New Zealand's approach to improving the achievement of secondary school students differs from approaches adopted in other countries. Rather than imposing high-stakes pressure on schools through accountability processes, New Zealand's approach has been through raising the capacity of schools and teachers, and through offering assessment frameworks that give more choice to students and schools (Hipkins et al., 2007). Similarly, entry to the large majority of tertiary courses is controlled by a general entry criteria (e.g., University Entrance).

To help raise the capacity of teachers and schools, a professional development programme that encouraged teachers to reflect on their beliefs about assessment and its link to students' learning was an integral component of the implementation of NCEA. Implicit in this professional development programme was an increased recognition of the potential for formative assessment to improve students' learning. Research into high-stakes SBA systems signals an increased awareness and use of formative assessment practices in the classroom (Clune, 2001; Harlen & Deakin Crick, 2003; James, 2000; OECD, 2005; Wilson & Floden 2001). Indeed, research within New Zealand suggests that many schools are beginning to acknowledge the value of SBA to learning (Barrington, 2004; Hipkins & Vaughan, 2005).

However the potential for undesirable pedagogical and learning outcomes resulting from the introduction of NCEA should also be acknowledged. The identification and division of a set of assessment criteria split into achievement levels has the potential to negatively impact on learning and teaching in a number of ways. Firstly, critics of SBA have highlighted the potential for atomisation of the curriculum (Zepke et al.,

2006). The current mathematics Unit and Achievement Standards are constructed along traditional content lines consistent with the strands identified in Mathematics in the New Zealand Curriculum (MoE, 1992). This atomisation potentially restricts teachers' and students' opportunities to make connections between related mathematical content and strategies. For example, there are currently separate Standards assessing algebra techniques and graphical techniques. However, many mathematical problems may equally well be solved using either an algebraic approach or a graphical approach (e.g., simultaneous equations). Under the existing set of Standards, students who answer a question using an appropriate graphical approach within an algebra Standard will not receive credit for their work.

To address this shortcoming, NZQA is in the process of redesigning some of the mathematics Standards to assess students' ability to solve 'types' of mathematical problems by a range of techniques, rather than assessing particular content based areas. It is hoped that this revision will encourage students to make connections between different areas of mathematics, a notion that was highlighted earlier as an important component of the work that practising mathematicians do (Boaler, 2002, Burton, 1999).

Secondly, the identification of achievement levels within Achievement Standards has the potential to limit students' learning by encouraging a 'doing just enough' (Meyer et al., 2006) approach to learning, where students limit themselves to learning only the Achieve material. While both teachers and students in the current research were critical of this approach to learning, they recognised that this approach might be attractive to some students. Within the discipline of mathematics, the potential for replacement evidence in more difficult questions provides a suitable incentive for students to maximise their achievement levels by 'doing their best' (Meyer et al., 2006). Accordingly teachers will need to encourage students to approach all learning opportunities with an intrinsic rather than extrinsic motivational orientation.

Lastly, the flexibility offered under NCEA to design courses with any number or combination of assessment Standards can be considered both a strength and a potential weakness of the assessment system. Its strength lies in the ability of schools to offer "enhanced flexibility … to offer broader and deeper learning for all students"

(Mallard, 2001, p.5). However, many schools have chosen to provide for able students' learning needs by increasing the number of Standards that students sit in order to maximise the number of credits students received rather than maximising students' achievement levels in a smaller number of Standards. (Hipkins et al., 2007) Accordingly, it could be argued that schools are offering 'broader', but not 'deeper' learning opportunities for students. Offering more assessment Standards has the effect that it reduces the time to teach any given standard. This over-assessment has been identified as a major factor negatively influencing students' approach to learning (Hipkins et al., 2007).

Furthermore, the reduction in time to teach the mathematical content for a given Standard, combined with New Zealand's historically egalitarian approach to the provision for gifted and talented students (McAlpine & Reid, 1987, Rawlins, 2000), potentially results in students being left to try and master the Excellence level material with limited help from their teachers. Indeed the research teachers in the current study recognised that students were often required to undertake the learning of the Excellence material themselves. It could be argued that the higher achieving students are more likely to go on to use higher level mathematical skills in future careers. As such, these students are entitled to receive a level of education that will assist them to achieve to their highest potential, as detailed in the Ministry of Education's National Education Goals (Rawlins, 2000).

While a number of potentially undesirable outcomes for learning and teaching are acknowledged, it should be noted that these are not inevitable consequences of the introduction of NCEA. NCEA is an assessment system designed to support rather than direct teaching and learning. The decisions of *what* and *how* to teach are school decisions—with guidance from the New Zealand Curriculum Framework (MoE, 1993). Schools and teachers have the potential to use the structure of NCEA to realise more fully the goal of students obtaining "[t]he highest standards of achievement, through programmes which enable all students to realise their full potential as individuals, and to develop the values needed to become full members of New Zealand's society" (National Education Goal 1).

5.6 Assessment in a sociocultural learning environment

The conceptualisation of a more complete theory of learning in a social environment will need to consider the interplay between learning, teaching and assessment. In particular, theory will need to more overtly address the role that assessment, including high-stakes assessment, plays in shaping the actions and understandings of individuals, and groups of individuals, within that learning environment.

Existing conceptions of assessment within a sociocultural learning environment need to explain more fully the role that the social environment plays in determining students' individual and collective engagement with formative feedback. Similarly, existing conceptions of assessments need to more fully consider *how* students develop metacognitive knowledge of the potential impact of formative assessment practices on their learning. The following section seeks to provide a perspective of these two issues.

5.6.1 A sociocultural perspective on student engagement with formative assessment

Pedagogies based on contemporary sociocultural perspectives of learning need to consider assessment systems and practices that are consistent with this perspective of learning. Implicit in much of the research on effective formative assessment is the assumption that it stimulates the *individual* to construct the next step in learning. Theory should also be cognisant of the potential for formative practices to stimulate the *community of learners* to engage collectively with formative assessment.

Although there is a considerable body of evidence describing characteristics of effective feedback (e.g., Black & Wiliam, 1998a; Crooks, 1988; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Tunstall & Gipps, 1996), much of this research is centred on the individual use of feedback to stimulate learning. To date, research and theory has placed too little attention on the potential for collective use of feedback to stimulate learning in a sociocultural environment. The decrease in a competitive element between students for grade allocation inherent in SBA provides greater opportunities for teachers to encourage collective, as well as individual, use of formative feedback.

This current research has demonstrated that students have a natural propensity to work with their peers when assessment feedback is returned to them. Students are reluctant to ask their teacher for help, and little attention is paid to teachers' oral feedback delivered immediately after the return of assessment scripts. This natural propensity for students to work with their peers can be enhanced by teachers providing opportunities for peer-collaboration as a planned activity therefore establishing it as a valid formative practice. Collaborative use of feedback can also be considered a valid learning activity in its own right, extending beyond its use to correct errors in individual student's work. Students can use the peer discussion of feedback to learn new material or alternative approaches to mathematical questions. In this way, formative assessment practices will become accepted as an integral part of learning and a cultural norm in the classroom.

Collaborative activities could include the use of feedback generated from 'peer-enhanced self-assessment'. Embedded in the concept of 'self-assessment' is the notion that it is conducted by the individual in relative isolation from social stimulus. While peer-assessment recognises the involvement of social processes, it involves one student assessing the work of another. Peer-enhanced self-assessment would combine these two elements. Within existing theory, little emphasis has been placed on the potential for group discussion of an individual student's work to enhance that student's ability for self-assessment. Support for this perspective comes from research which identified that students' use of peer-assessment enhanced their ability to self-assess (Black et al., 2003; Ross et al., 1993; Sadler, 1989; 1998).

While collaborative use of formative feedback and 'peer-enhanced self-assessment' have the potential to enhance learning, it cannot be assumed that students are proficient at understanding feedback, self- and peer-assessing, or working with their peers (Sadler, 1998; Tunstall & Gipps, 1996). They will need to be provided with opportunities to learn and practise these skills in their regular classroom setting. This will help to establish these formative practices as a normal part of the mathematics classroom, which is consistent with a sociocultural perspective on how students develop understandings of the impact of assessment on learning. As the collective use of feedback is legitimised as a formative strategy within a community of practice, it has the potential to improve the independent use of feedback by students.

Implications

Two major implications stem from the perspective that feedback can stimulate learning in a sociocultural environment. Firstly, this notion has implications for the structure of any planned feedback sessions subsequent to marked scripts being returned. Teachers should plan suitable activities where students interact with their peers and engage with the written feedback to enhance learning and deepen understanding of their work. An example of a potentially suitable activity was illustrated in section 4.5.1. One of the teachers had returned the students' assessment scripts after the third term mock examinations and told the students to correct their errors and hand the scripts in for re-marking. Unfortunately, the full potential of this activity could not be assessed as the teacher did not explicitly encourage students to work with their peers, and no class time was allocated for this activity.

Secondly, the notion of collaborative use of feedback has implications for the nature of the written feedback given to students. Feedback must include enough specific detail about errors or next steps, and how to address them, to allow peer-supported development of strategies to deepen students' learning. It is not sufficient to provide them with 'just the answers' as evidenced from the extract cited in 4.5.1. Students require scaffolded feedback that allows them to think through and discuss questions with other students, therefore placing them in the ZPD (Vygotsky, 1978). Collective engagement with feedback to 'work it out' is consistent with contemporary perspectives of learning mathematics that argue that making connections is an important part of being an effective mathematics student (Boaler, 2002; Burton, 1999). Furthermore, since within an SBA environment students will only be able to improve their grade by improving their understanding of the mathematical content, teachers' and students' attentions is focussed more keenly on important task-oriented characteristics of effective formative feedback.

The suggestion that students should initially work with their peers to engage with written feedback does not preclude the use of oral feedback. It does, however, require an examination of the timing of the delivery of such feedback. Students in the current research preferred to use teacher help only when they could not identify errors and how to correct them from the written feedback (see section 4.6.1). It could be

conjectured that students' reluctance to ask their teacher for help is, in part, because they have had insufficient time to examine the feedback, discuss it with their peers, and identify whether they need to ask any questions of the teacher. In an analogous way to the value of 'wait time' following teachers' questions, a delay in teachers asking for questions could encourage students to critically analyse the written feedback given, and help them to be in a better position to ask more focussed questions of the teacher. This might, in turn, help teachers develop awareness of particular characteristics of their written feedback that are clearly understood or, perhaps more importantly, not understood by students.

5.6.2 Development of formative assessment strategies

Sociocultural perspectives of learning are underpinned by the belief that students should be actively involved in developing knowledge within communities of learners (Even & Tirosh, 2002; Lave, 1988; Lave & Wenger, 1991; Sfard, 1998). In addition to content knowledge and skills, students should also develop knowledge about, and dispositions towards, active involvement in their own learning, both individually and collectively. Active involvement in learning must also include an active involvement with assessment practices.

Research has identified that high achieving students independently develop strategies to formatively engage with assessment information (Brookhart, 2001; Higgins et al., 2002; Hyland, 2000). However, there is little support in the literature for widespread independent development of effective formative assessment strategies by middle and low achieving students. Because it is hypothesised in this research project that students' development of knowledge of the formative potential of assessment events is socially and culturally situated in the classroom assessment environment, it is imperative that teachers' instructional practices include explicit support for these students.

The current research has presented evidence that the provision of well intentioned tools designed to help students understand assessment criteria, and feedback practices designed to be formative, may not be sufficient to embed the formative use of assessment into the classroom norms of effective learning. Structuring teaching

around levels of the assessment Standard may not, in itself, lead to strong student understanding of the assessment criteria. Similarly, the provision of the 'I can do' sheets designed to help students' self- and peer-assess may not, by itself, lead to frequent independent use of these assessment strategies by students.

Additionally, there is a lack of agreement between the participant students and teachers regarding optimum feedback practices. The research teachers perceived that students did not read written feedback offered, and this perception partially influenced the amount of written feedback offered. In contrast, students have provided clear evidence that they read, valued and used written feedback. Teachers perceived that oral feedback is the most efficient and effective feedback practice, and interpreted a lack of student focus during these session as a lack of interest in feedback per se. In contrast, students reported that they preferred to engage with feedback in collaboration with their peers in the first instance. It could be argued that this lack of agreement is a result of students and teachers independently developing knowledge of effective formative strategies, rather than collectively developing them.

The way that teachers use and talk about assessment practices sends important messages to students; many of these messages are implicit and unintentional (Black & Wiliam, 1998a; Harlen & Deakin Crick, 2003; Stiggins, 2001). As such, greater engagement in formative assessment practices should be seen as a shared challenge. Although it is reasonable to expect students to take responsibility for their learning and engagement in assessment practices, given the importance of the learning environment, teachers have a pivotal role in creating appropriate conditions for greater student engagement in formative assessment practices. This extends beyond teachers 'informing' students how they should use assessment information formatively. From a sociocultural perspective the teacher must create an environment in which students collectively develop this knowledge. Students must come to believe that understanding assessment criteria, the use of self- and peer-assessment, and active engagement in feedback are normal practices in learning mathematics. In this way students recognise and accept the benefits of active participation in assessment processes.

As Hattie and Timperley (2007) argue:

> Teachers and parents often assume that students share a commitment to academic goals, whereas the reality is that developing this shared commitment needs to be nurtured and built. (p. 89)

The development of formative practices and tools needs to become part of the cultural norms of the classroom through social discourse. This can best be achieved if "teachers lead and encourage students to come with them" (Hipkins et al. 2007, p. 25). An implication for teachers of this position is that, rather than assuming they understand students' propensity and preferences for engaging with feedback, teachers should work with their students to co-develop a shared understanding of what constitutes effective feedback practices. Teachers need to expose all students to a range of formative practices and make explicit links between these practices and benefits to students' learning. This knowledge then resides in the minds of students and teachers alike, and is a common language that they can share.

5.7 Reflections on the research process

5.7.1 Students' voice

As noted earlier, teachers' and students' perceptions of effective formative practices are not always in agreement. This observation highlights the value of the students' voice in helping to enhance the teaching and learning that occurs in classrooms. From a sociocultural perspective the students' voice is a central component of the community of learners in the classroom. However, too often students are not consulted about educational matters that directly affect them (Vaughan, 2003).

Both the research teachers and the students in this project actively engaged with the research process. The teachers could be characterised as reflective practitioners with a clear interest in improving their current practice. They entered into the research process enthusiastically and maintained a keen interest throughout the year. Similarly, the students in the research project appeared motivated to participate by a genuine desire to enhance the collective knowledge about effective formative use of

NCEA systems. It could be argued that, given an understanding of the value of consultation by participants, systems could be set up within the school to allow a collective development of best practice in assessment. Such systems will need to be transparent and respect the rights and responsibilities of both teachers and students. Given the relative success of the use of text messaging in opening up contemporary communication channels, similar technologically such as 'blogs' could be used as a consultative device, provided that suitable guidelines and protections were established.

5.7.2 Limitations

It is important to acknowledge the limitations in any research study. Limitations are recognition of the fact that any educational research study is a partial picture of a complex set of interdependent variables. The limitations inherent in using case study methodology have been discussed in chapter 3, however, others are worth mentioning here.

Firstly it should be recognised that this research investigated students' perceptions of their experiences. The interpretivist stance adopted in this research acknowledges that multiple views of reality exist. As such, the students' perceptions are their interpretation and recollections of their assessment experiences, and these may differ from interpretations made by their teachers or the researcher. However, as identified above, research must be cognisant of the lived experiences of all participants in the teaching and learning process. Knowledge of students' perspectives provides an opportunity to examine the impact of teachers' assessment practices on students, and to determine whether the actual impact was what was intended. To help contextualise the students' perspectives, classroom observations and an interview with the teacher were also conducted.

Secondly, the self selected research students were not representative of all ability levels (see Appendix L). How closely such a sample represents the wider cohort is often debated. Although the self selecting students were moderate to high achievers, the AEQ confirmed the existence of a close relationship between the sample students and the wider Y12 Trad Maths cohort. It is recognised, however, that the perceptions

of the research students reported in this project may not be representative of all students in this cohort.

Lastly, recognition is also made of the relatively small size, and subject specificity, of the research sample. Investigating nine students within one subject area at one school limits the generalisable nature of the findings of this particular study to a wide range of educational settings. However, as argued by Stake (1995) the nature of generalisation is that it refers to the learning process through which we individually acquire concepts and information and steadily generalise them to other situations. It is through careful replication of case studies over time, that 'petite' generalisations are expanded into 'grandes' generalisations (Stake).

Accordingly, it is not meant to imply that the perceptions reported here are necessarily representative of all Y12 mathematics students throughout New Zealand. Rather the perceptions reported here are indicative of the range of possible perceptions and further investigation will be required in order to more fully appreciate the complex nature of students' involvement in formative assessment practices. Additionally the fluid nature of education systems in general, and New Zealand's senior secondary school assessment system in particular, means that any research is a snapshot within a dynamic system. Multiple research projects over time may serve to illustrate and map the changes, tensions and directions of a necessarily complex, multi-faceted system.

Any extended period of research is also subjected to the development of perceptions amongst the research students. To a certain extent many of these perceptions will develop naturally over time as students' experience base expands. However, involvement in the research project may also have caused an increased reflection on the inter-relationship between aspects of students' learning. Since this project examined students' perceptions of formative assessment practices they are exposed to, it is reasonable to expect that students will increasingly examine how they currently use assessment information, and how these practices may be improved. Accordingly, it is recognised that students' perceptions are socially, culturally and historically situated.

5.7.3 Implications for further study.

This research project has investigated students' perceptions of the formative potential of NCEA. It has provided some insights in to how students and teachers use a high-stakes assessment system, designed primarily for summative purposes, for their own formative ends. In recognising that the current research has only investigated a small part of a larger concern, a number of suggestions for future research can be made.

This research proposed perspectives on two aspects of conceptions of assessment within sociocultural theories of learning that are underdeveloped. Firstly, the role that feedback plays in generating collaborative learning opportunities, and secondly, collective development of knowledge of effective formative assessment practices is socially and contextually situated. Futher research into pedagogies that help develop and support effective collaborative formative assessment practices should be conducted.

Given the historically under-represented nature of the students' voice, further research should be cognisant of the important role that students play in developing inclusive and effective practices within communities of learners. Given the importance placed on the situated nature of cognition within this current project, it is suggested that case study is an appropriate methodology for future research. A valuable addition to the data gathering methods could be to include student diaries as a way of gathering students' reflections of their assessment experiences.

A number of other suggested areas for future research can be made. Firstly, research on the effective use of 'I can do' sheets as a way of helping students learn about the assessment criteria that they are working towards, and how to use 'I can do' as effective self- and peer-assessment tools is warranted. Secondly, research on the structured use of activities designed to encourage collaborative use of formative feedback would enhance the sociocultural perspective on assessment in communities of practice. Lastly, research could also investigate the most effective methods of combining oral and written feedback on assessment tasks.

The current research also suggested that differences in the way schools implement NCEA will result in a range of potential impacts on teachers' pedagogy and students' learning. Accordingly, ongoing research into the integration and impact of formative assessment practices in senior classrooms will need to be undertakne. In particular, concerns were raised over the potential impact of over-assessment on students' learning and achievement, especially higher achieving students who are often left to try and master the more difficult material with limited teacher assistance. Research should be undertaken into the relative merits of valuing higher number of credits over higher grades.

5.8 Final thoughts

Assessment practices in secondary schools in New Zealand are currently in a state of flux. Recent changes to the senior secondary school exit qualification have caused many to question the "long-held tradition of viewing assessment-for-qualification as *the* main purpose for learning in the senior secondary school" (Hipkins et al., 2007, p. 18). The potential for assessment systems to support and enhance students' learning has received increasing recognition recently.

This research project investigated students' perception of the formative potential of NCEA in the mathematics classroom. Using a theoretical framework developed by Sadler (1989), it identified that the structure inherent in SBA has the potential for NCEA to satisfy both summative and formative purposes of assessment. It can be concluded that this potential is still to be fully realised.

Despite teachers' pedagogical strategies, students' knowledge of assessment criteria is underdeveloped. This underdevelopment reduces the potential for students' independent use of self-assessment strategies. This project has also identified that teachers and students held differing views on preferred feedback practices. The teachers perceived that students did not read written feedback and this perception significantly influenced the amount of written feedback that they offered to students. In contrast, students clearly displayed that they read, valued and used scaffolded written feedback to improve their learning. While the teachers preferred to offer oral

feedback, students preferred to engage with their peers to use the feedback to develop corrective strategies and deepen learning.

The project has made a number of practical and theoretical suggestions to improve students' understandings of the assessment criteria they are working towards, and to more effectively integrate the collaborative use of formative feedback into students' learning experiences. In particular, it has suggested two additional perspectives on the development and use of formative assessment in a sociocultural learning environment. Firstly, that students' knowledge of the role of formative assessment is socially and contextually situated, and develops through the social interactions that occur in the classroom. Secondly, the potential exists for formative assessment practices to stimulate collaborative learning opportunities within communities of practice.

It should be noted that a move towards using models of assessment that promote learning will not in itself bring about changes in teaching and learning, It is not enough that we have an assessment tool that potentially allows for a formative model of assessment to become the norm in New Zealand classrooms. Teachers must be supported with appropriate professional development and research. It will take time before the effective use of formative assessment as an integral component of learning becomes the norm in senior mathematics classes. As Harlen (2005) notes, "it takes a good deal of support—and courage—for teachers to turn round their practices from being test-oriented to being learning-oriented (p. 210). However, if we can achieve this then the perception of the role of the teacher can change from being the 'judger of student competence' to the 'guider of students' learning' (Crooks, 2006).

References

Albrecht, T. L., Johnson, G. M., Walther, J. B., & Krueger, R. A. (1993). Understanding communication processes in focus groups. In D. L. Morgan (Ed.), *Successful focus groups: Advancing the state of the art* (pp. 51-64). Newbury Park, CA: Sage Publications.

Alton-Lee, A., & Nuthall, G. (1990). Pupil experiences and pupil learning in the elementary classroom: An illustration of a generative methodology. *Teaching and Teacher Education, 6*(1), 27-45.

Alton-Lee, A., Nuthall, G., & Patrick, J. (1993). Reframing classroom research: A lesson from the private world of children. *Harvard Educational Review, 63*(1), 50-84.

Alton-Lee, A., & Nuthall, G. (1998). *Inclusive instructional design: Theoretical principles emerging from the understanding learning and teaching project*. Wellington: Ministry of Education.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261-271.

Anderson, G. L. (1990). *Fundamentals of educational research*. London: Falmer Press.

Anderson, A. (1993). Assessment: A means to empower children? In N. L. Webb & A. F. Coxford (Eds.), *Assessment in the mathematics classroom* (pp. 103-110). Reston, VA: National Council of Teachers of Mathematics.

Anthony, G. J. (1994). *Learning strategies in mathematics education.* Unpublished doctoral thesis, Massey University, Palmerston North, New Zealand.

Anthony, G. (2000). Factors influencing first-year students' success in mathematics. *International Journal of Mathematical Education in Science & Technology, 31*(1), 3-14.

Anthony, G., & Walshaw, M. (2003). *Fractions assessment using a real context: NEMP probe report*. Palmerston North, New Zealand: Massey University.

Anthony, G., & Walshaw, M. (2007). *Effective pedagogy in mathematics/pangarau: Best evidence synthesis iteration [BES]*. Wellington: Ministry of Education.

Askew, M., Brown, M. L., Rhodes, V., Johnson, D., & Wiliam, D. (1997). *Effective teachers of numeracy: Final report*. London: King's College School of Education.

Australian Education Council. (1994). *Mathematics: A curriculum profile for Australian schools / a joint project of the States, Territories and the Commonwealth of Australia initiated by the Australian Education Council*. Carlton, Victoria: Curriculum Corporation.

Barrington, E. (2004, March). *Teaching snippets.* Paper presented at the Vice-Chancellor's Symposium on National Certificate of Education (NCEA), Auckland University.

Bassey, M. (1999). *Case study research in educational settings*. Philadelphia: Open University Press.

Begg, A. (1998). Assessment and school mathematics. *SAMEpapers*, 4-25.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education, 85*(5), 536-553.

Benmansour, N. (1999). Motivational orientations, self-efficacy, anxiety and strategy use in learning high school mathematics in Morocco. *Mediterranean Journal of Educational Studies, 4*, 1-15.

Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 361-392). Hillsdale, NJ: Lawrence Erlbaum Associates.

Berg, B. L. (2004). *Qualitative research methods* (5th ed.). Boston: Pearson.

Biggs, J. B. (1987). *Student approaches to learning and studying*. Hawthorn, Vic: Australian Council for Educational Research.

Biggs, J. B. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 347-364.

Biggs, J. (1998). Assessment and classroom learning: A role for summative. *Assessment in Education: Principles, Policy & Practice, 5*(1), 103-110.

Birmingham, K. (2001). *Who will pass? Validity of standards-based assessment measure to predict high school mathematics success.* Unpublished doctoral dissertation, Alfred University, NY.

Bishop, R. (2003). Changing power relations in education: Kaupapa Maori messages for 'mainstream' education in Aotearoa/New Zealand. *Comparative Education, 39*(2), 221–238.

Bishop, R., Berryman, M., Tiakiwai, S., & Richardson, C. (2003). *Te Kotahitanga: The experiences of year 9 and 10 Maori students in mainstream classrooms*: New Zealand, Ministry of Education.

Black, P. (2000). Research and the development of educational assessment. *Oxford Review of Education, 26*(3&4), 407-419.

Black, P. (2001). Dreams, strategies and systems: Portraits of assessment past, present and future. *Assessment in Education: Principles, Policy & Practice, 8*(1), 65-85.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice.* Maidenhead: Open University Press.

Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: A theoretical inquiry. *Research Papers in Education, 21*(2), 119-132.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.

Black, P., & Wiliam, D. (1998b). *Inside the black box: raising standards through classroom assessment*. London: Kings College.

Black, P., & Wiliam, D. (2003). 'In praise of educational research': formative assessment. *British Educational Research Journal, 29*(5), 623-637.

Black, P., & Wiliam, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning* (pp. 9-26). London: Sage Pubications.

Bloom, B. S., Hastings, T., & Madaus., G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

Boaler, J. (1998). Open and closed mathematics: Students experiences and understandings. *Journal for Research in Mathematics Education, 29*(1), 41-62.

Boaler, J. (1999). Participation, knowledge and beliefs: A community perspective on mathematics learning. *Educational Studies in Mathematics, 40*(3), 259-281.

Boaler, J. (2002). Exploring the nature of mathematical activity: Using theory, research and 'working hypotheses' to broaden conceptions of mathematics knowing. *Educational Studies in Mathematics, 51*(1-2), 3-21.

Bogdan, R., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theory and methods* (5th ed.). Boston, MA: Pearson A & B.

Boss, T., Endorf, D., & Duckendahl, C. (2001). *Informing state assessment from the local level: A district's reflections.* Paper presented at the Annual Meeting of the Mid-Western Education Research Association, Chicago, IL

Boud, D. J. (1995). *Enhancing learning through self assessment*. London; Philadelphia: Kogan Page.

Bourke, R. (2000). *Students' conceptions of learning and self-assessment in context.* Unpublished doctoral thesis, Massey University, Palmerston North, New Zealand.

Broadfoot, P., Pollard, A., Osborn, M., McNess, E., & Triggs, P. (1998, April). *Categories, standards and instrumentalism: Theorising the changing discourse of assessment policy in English primary education.* Paper presented at the Annual Meting of the AERA, San Diego, CA.

Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education: Principles, Policy & Practice, 8*(2), 153-169.

Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record, 106*(3), 429-458.

Brookhart, S. M., & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education, 12*(4), 409-425.

Brooking, K. (2006, Dec). *The NCEA: Competing viewpoints since implementation and the debate about its future.* Paper presented at the NZARE Conference, Rotorua.

Brown, G. (2002). Student beliefs about learning: New Zealand students in year 11. *Academic Exchange Quarterly, 6*(1), 110-114.

Brown, G. T. L. (2006). *Secondary school students' conceptions of assessment: A survey of four schools (Conceptions of Assessment and Feedback Project Report #5).* Auckland, New Zealand: University of Auckland.

Brush, G. (1997). *A year with criterion-referenced assessment: Setting my students up for success*. Retrieved 22 July, 2004, from http://www.bctf.ca/inquirer/support/teacherassessemnt.html.

Burgess, T. (2007). *An investigation of the dynamic nature of teacher knowledge within the context of teaching statistics.* Unpublished doctoral thesis, Massey University, Plamerston North, New Zealand.

Burton, L. (1999). The practices of mathematicians: What do they tell us about coming to know mathematics? *Educational Studies in Mathematics, 37*(2), 121-143.

Bushnell, P. (1992). *The development of grade-related criteria in Sixth Form Certificate drama (Research Report No. 92-1).* Christchurch, New Zealand: University of Canterbury, Education Department.

Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology, 58*, 1-14.

Cardno, C. E. M. (2003). *Action research: A developmental approach*. Wellington: New Zealand Council for Educational Research.

Carr, M., McGee, C., Jones, A., McKinley, E., Bell, B., Barr, H., et al. (2000). *Literature review: the effects of curricula and assessment on pedagogical approaches and on educational outcomes*. Wellington, New Zealand: Research Division Ministry of Education.

Carroll, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology, 83*, 360-367.

Cavanagh, M. (2006). Mathematics teachers and working mathematically: Responses to curriculum change. In P. Grootenboer, R. Zevenbergen, & M. Chinnappan (Eds.), *Identities, cultures and learning spaces.* (Proceedings of the 29th annual conference of the Mathematics Education Research Group of Australasia, Vol 1. pp. 115-122). Sydney: MERGA.

Clune, W. H. (2001). Towards a theory of standards-based reform: The case of nine NSF statewide systematic initiatives. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 13-38). Chicago: University of Chicago Press.

Codd, J., McAlpine, D., & Poskitt, J. (1995). Assessment policies in New Zealand: Educational reform or political agenda? In R. Peddie & B. Tuck (Eds.), *Setting the standards: Issues in assessment for national qualifications* (pp. 32-58). Palmerston North: Dunmore Press.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London; New York: RoutledgeFalmer.

Confrey, J., & Kazak, S. (2006). A thirty-year reflection on constructivism in mathematics education in PME. In A. Guttiérrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present and future* (pp. 305-346). Rotterdam: Sense Publishers.

Cotton, T. (2001). Mathematics teaching in the real world. In P. Gates (Ed.), *Issues in mathematics teaching* (pp. 23-37). London: RoutledgeFalmer.

Covington, M. V., & Müeller, K. J. (2001). Intrinsic versus extrinsic motivation: An approach/avoidance reformulation. *Educational Psychology Review, 13*(2), 157-176.

Cowie, B. (2005). Pupil commentary on assessment for learning. *Curriculum Journal, 16*(2), 137-151.

Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 4*, 438-481.

Crooks, T. (2001, September). *The validity of formative assessments.* Paper presented at the British Educational Research Association Annual Conference, University of Leeds.

Crooks, T. J. (2002). Educational assessment in New Zealand schools. *Assessment in Education: Principles, Policy & Practice, 9*(2), 237-253.

Crooks, T. (2006). *Assessment for learning in New Zealand: What's happening and what's next?* Paper presented at the What's next with assessment for learning conference, Massey University, Palmerston North, New Zealand.

Davies, J., & Brember, I. (1998). National Curriculum testing and self-esteem in Year 2. The first 5 years: A cross-sectional study. *Educational Psychology, 18*(4), 365-375.

Daws, N., & Singh, B. (1996). Formative assessment: To what extent is its potential to enhance pupils' science being realized? *School Science Review, 77*(28 1), 93-100.

DES/WO. (1988). *Task Group on Assessment and Testing: A report*. London: Department of Education and Science and the Welsh Office.

Dobric, K. (2006). Drawing on discourses: Policy actors in the debate over the National Certificates of Educational Achievement 1996-2000. *Annual Review of Education in New Zealand, 15*, 85-109.

Dweck, C. S. (1992). The study of goals in psychology. *Psychological Science, 3*(3), 165-167.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273.

Education Commission of the States. (2002). *A guide to standards-based assessment: No child left behind*: National Center for Research on Evaluation, Standards and Student Testing.

Education Review Office. (2004). *Progress of NCEA Levels 1 and 2, and readiness of schools to implement Level 3*. Wellington, New Zealand: Author.

Egan, K. (1997). *The educated mind: How cognitive tools shape our understanding*. Chicago: University of Chicago Press.

Eng, G. (1992). *The use of grade-related criteria in Sixth Form Certificate English: A comparative study (Research Report No. 92-2)*. Christchurch, New Zealand: University of Canterbury, Education Department.

Ernest, P. (2004). Images of mathematics, values and gender: a philosophical perspective. In B. Allen & S. Johnston-Wilder (Eds.), *Mathematics education: exploring the culture of learning* (pp. 11-25). London; New York: RoutledgeFalmer.

Even, R., & Tirosh, D. (2002). Teacher knowledge and understanding of students' mathematical learning. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 219-240). Mahwah, NJ: Lawrence Erlbaum.

Fielding, M. (2006). Listening hard to the student voice. *Education Review, 11*(6), 6-7.

Fontana, D., & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology, 64*, 407-417.

Forman, E. (2003). A sociocultural approach to mathematics reform: speaking, inscribing, and doing mathematics in a community of practice. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 333-352). Reston, VA: National Council of Teachers of Mathematics.

Frederiksen, J. R., & White, B. J. (1997). *Reflective assessment of students' research within an inquiry-based middle school science curriculum.* Paper presented at the Annual Meeting of the AERA, Chicago.

Gates, P. (Ed.). (2001). *Issues in mathematics teaching*. London: RoutledgeFalmer.

Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education: Choosing and using diverse approaches* (pp. 41-53). Buckingham: Open University Press.

Gibbs, G., & Simpson, C. (2003, September). *Measuring the response of students to assessment: The Assessment Experience Questionnaire.* Paper presented at the 11th Improving Student Learning Symposium, Hinckley, Leicestershire.

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports student learning. *Learning and Teaching in Higher Education,* (1), 3-31.

Gillham, B. (2000). *Case studies research methods*. London: Continuum.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. London: The Falmer Press.

Gipps, C., & Tunstall, P. (1998). Effort, ability and the teacher: Young children's explanations success and failure. *Oxford Review of Education, 24*(2), 149-165.

Goos, M. (2004). Learning mathematics in a classroom community of practice. *Journal for Research in Mathematics Education, 35*(4), 258-291.

Goos, M., Galbraith, P., & Renshaw, P. (2004). Establishing a community of practice in a secondary mathematics classroom. In B. Allen & S. Johnston-Wilder (Eds.), *Mathematics education: Exploring the culture of learning* (pp. 91–116). London: RoutledgeFalmer.

Harlen, W. (2005). Teachers' summative practices and assessment for learning—tensions and synergies. *Curriculum Journal, 16*(2), 207-223.

Harlen, W. (2006). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 61-80). London: Sage Publication.

Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice, 10*(2), 169-208.

Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships. *Assessment in Education: Principles, Policy & Practice, 4*(3), 365-380.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70*(2), 151-179.

Hiebert, J., Carpenter, T., Fennema, E., Fuson, K. C., Wearne, D., Murray, H., et al. (1997). *Making sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.

Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education, 27*(1), 53-64.

Hipkins, R. (2004). Changing school subjects for changing school times. *SPANZ Journal, August*, 13-21.

Hipkins, R., Vaughan, K., Beals, F., & Ferral, H. (2004). *Learning curves: Meeting student learning needs in an evolving qualifications regime: Shared pathways and multiple tracks: A second report* (No. 1877293342). Wellington, New Zealand: New Zealand Council for Educational Research.

Hipkins, R., Vaughan, K., with Beals, F., Ferral, H., & Gardner, B. (2005). *Shaping our futures: Meeting secondary students' learning needs in a time of evolving qualifications: Final report of the learning curves project* (No. 1877293490). Wellington, New Zealand: New Zealand Council for Educational Research.

Hipkins, R., Wylie, C., & Hodgen, E. (2007, April). *Can standards-based qualifications improve low-performing students' engagement in learning, and their achievement?* Paper presented at the Annual Meeting of the AERA, Chicago.

Hitchcock, G., & Hughes, D. (1995). *Research and the teacher: A qualitative introduction to school-based research* (2nd ed.). London; New York: Routledge.

Hosking, P., & Shield, M. (2001). Feedback practices and the classroom culture. In J. Bobis, B. Perry, & M. Mitchelmore (Eds.), *Numeracy and beyond.* (Proceedings of the 24th annual conference of the Mathematics Education Research Group of Australasia, Vol. 1, pp. 289-297). Sydney: MERGA.

Hughes, I. E., & Large, B. J. (1993). Staff and peer-group assessment of oral communication skills. *Studies in Higher Education, 18*(3), 379-385.

Hyland, P. (2000). Learning from feedback on assessment. In A. Booth & P. Hyland (Eds.), *The Practice of University History Teaching* (pp. 233-247). Manchester: Manchester University Press.

James, M. (2000). Measured lives: the rise of assessment as the engine of change in English schools. *The Curriculum Journal, 11*(3), 343-364.

James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 47-60). London: Sage Publication..

James, M., & Gipps, C. (1998). Broadening the basis of assessment to prevent the narrowing of learning. *The Curriculum Journal, 9*(3), 285-297.

Johnston, J., & McClune, W. (2000). *Selection Project Sel 5.1: Pupils motivation and attitudes—self-esteem, locus of control, learning disposition and the impact of selection on teaching and learning*. Belfast: Department of Education for Northern Ireland.

Kane, R., & Maw, N. (2005). Making sense of learning at secondary school: Involving students to improve teaching practice. *Cambridge Journal of Education, 35*(3), 311-322.

Kannapel, J. P., Aagard, L., Coe, P., & Reeves, C. (2001). The impact of standards and accountability on teaching and learning in Kentucky. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the States* (pp. 217-241). Chicago: The University of Chicago Press.

Kellaghan, T., Madaus, G., & Raczek, A. (1996). *The use of external examinations to improve student motivation*. Washington, DC: AERA.

Khisty, L. L., & Chval, K. B. (2002). Pedagogic discourse and equity in mathematics: When teachers talk matters. *Mathematics Education Research Journal, 14*(3), 154-168.

Klenowski, V. (1995). Student self-evaluation processes in student-centred teaching and learning contexts of Australia. *Assessment in Education: Principles, Policy & Practice, 2*(2), 145-163.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.

Koch, A., & Shulamith, G. E. (1991). Improvement of reading comprehension of physics texts by students' question formulation. *International Journal of Science Education, 13*, 473-485.

Krueger, R. A., & Casey, M. A. (2000). *Focus groups: A practical guide for applied research* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Leder, G. C., & Forgasz, H. J. (1992). Perspectives on learning, teaching and assessment. In G. C. Leder (Ed.), *Assessment and learning of mathematics*. Hawthorn, Victoria: Australian Council for Educational Research.

Lee, H., & Lee, G. (2000). The National Certificate of Educational Achievement (NCEA): "Fragile—handle with care". *New Zealand Annual Review of Education, 10*, 5-38.

Lennox, B. (2001). Where did NCEA come from? *QA News, 38*, 10-11.

Leonard, M., & Davey, C. (2001). *Thoughts on the 11 plus*. Belfast: Save the Children Fund.

Lerman, S. (2000). The social turn in mathematics education research. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 19-44). Westport, CT: Ablex Pub.

Lewis, R. K. (2005). *School based assessment within a standards based assessment regime: A mathematics community's perspective.* Unpublished doctoral thesis, University of Otago, Dunedin, New Zealand.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The handbook of qualitative research* (2nd ed., pp. 163-188). Thousand Oaks, CA: Sage Publications.

Linn, R., & Herman, J. (1997). *A policymaker's guide to standards-led assessment*. Denver: ECS Distribution Centre.

Linn, R. (2000). Assessment and accountability. *Educational Researcher, 29*(2), 4-16.

Loretz, M. R. (2002). *Changes in mathematics assessment tasks and practice as the NCEA is introduced.* Unpublished master's thesis, University of Auckland, New Zealand.

Mallard, T. (2001). Minister launches campaign *QA News, 38*, 5.

McAlpine, D., & Reid, N. (1987). The Gifted and Talented. In D. R. Mitchel & N. N. Singh (Eds.), *Exceptional Children in New Zealand* (pp. 318-322). Palmerston North: Dunmore Press.

McCurdy, B. L., & Shapiro, E. S. (1992). A comparison of-teacher-monitoring, peer-monitoring, and self-monitoring with curriculum-based measurement in reading among students with learning disabilities. *Journal of Special Education, 26*, 162-180.

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy & Practice, 10*(2), 209-220.

McNeil, L., & Valenzuela, A. (2000). *The harmful impact of the TAAS system of testing in Texas: Beneath the accountability rhetoric.* Houston, Austin: Rice University, University of Texas.

Meek, N. (2001). Information campaign lifts off. *QA News, 38*, 4.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* San Francisco: Jossy-Bass Publishers.

Meyer, L., McClure, J., Wakley, F., McKenzie, L., & Weir, K. (2006). *The impact of the NCEA on student motivation: Final report* (No. 0478135076 (pbk.)). Wellington, New Zealand: Victoria University of Wellington.

Ministry of Education. (1989). *Assessment for better learning: A public discussion document*. Wellington: Dept. of Education.

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. (1989). *Assessment for better learning: A public discussion document*. Wellington: Dept. of Education.

Ministry of Education. (1998). *Achievement 2001: Qualifications for 16 to 19 year olds*. Wellington: Ministry of Education; New Zealand Qualifications Authority.

Ministry of Education. (2006a). *The New Zealand curriculum: Draft for consultation 2006*. Wellington: Learning Media.

Ministry of Education. (2006b). *Numeracy Development in New Zealand: The story continues*. Wellington, N.Z.: Author.

Mishler, E. G. (1986). *Research interviewing: Context and narrative*. Cambridge, MA: Harvard University Press.

Morgan, D. L. (1988). *Focus groups as qualitative research*. Newbury Park, CA: Sage Publications.

Morgan, D. L., & Krueger, R. A. (1993). When to use focus groups and why. In D. L. Morgan (Ed.), *Successful focus groups: Advancing the state of the art* (pp. 3-19). Newbury Park, CA: Sage Publications.

Morse, J. L. (1994). Design funded qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The handbook of qualitative research* (2nd ed., pp. 220-235). Thousand Oaks, CA: Sage Publications.

Nachmias, D., & Nachmias, C. (1992). *Research methods in the social sciences*. New York: St. Martin's.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Retrieved 18th Feb, 2007, from http://standards.nctm.org/document/prepost/project.htm.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Retrieved 18th Feb, 2007, from http://standards.nctm.org/document/prepost/project.htm.

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologists, 22*, 155-175.

Neyland, J. (1996, 18th October). *Teacher's knowledge: The starting point for a critical analysis of mathematics teaching*. Retrieved 26th March, 2007, from http://www.people.ex.ac.uk/PErnest/pome/pompart4.htm

Nunes, T., & Bryant, P. (1996). *Children doing mathematics*. Oxford; Cambridge, MA: Blackwell.

Nuthall, G. A. (1999a). The way students learn: acquiring knowledge from an integrated science and social studies unit. *Elementary School Journal, 99*, 303–341.

Nuthall, G. A. (1999b). Learning how to learn: the evolution of students' minds through the social processes and culture of the classroom. *International Journal of Educational Research, 31*(3), 139–256.

Nuthall, G. A. (2002). Knowing what we know and need to know about effective teaching. In B. Webber (Ed.), *Teachers make a difference: What is the research evidence? Conference Proceedings* (pp. 41-63). Wellington: New Zealand Council for Educational Research.

Nuthall, G., & Alton-Lee, A. (1993). Predicting learning from student experience of teaching: A theory of student knowledge construction in classrooms. *American Educational Research Journal, 30*(4), 799-840.

O'Donnovan, B., Price, M., & Rust, C. (2000). The student experience of criterion-referenced assessment (through the introduction of a common criteria assessment grid). *Innovations in Education and Teaching International, 38*(1), 4-85.

OECD. (2005). *Formative assessment: Improving learning in secondary classrooms.* Paris: OECD.

Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education, 27*(4), 309-323.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology, 49*(1), 345.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). London: Sage Publications.

Pfannkuch, M. (2001). Assessment of school mathematics: Teachers perceptions and practices. *Mathematics Education Research Journal, 13*(3), 185-203.

Phillips, D. (2003). Lessons from New Zealand's National Qualifications Framework. *Journal of Education & Work, 16*(3), 289-303.

Pollard, A., Triggs, P., Broadfoot, P., McNess, E., & Osborn, M. (2000). *What pupils say: Changing policy and practice in primary education*. London; New York: Continuum.

Porter, A. C., & Smithson, J. L. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the States* (pp. 60-80). Chicago: The University of Chicago Press.

Preece, P., & Skinner, N. (1999). The National Assessment in science at Key Stage 3 in England and Wales and its impact on teaching and learning. *Assessment in Education, 6*(1), 11-25.

Rawlins, P. (2000). *Acceleration in mathematics: Students' perspectives.* Unpublished master's thesis, Massey University, New Zealand.

Reay, D., & Wiliam, D. (1999). 'I'll be a nothing': Structure, agency and the construction of identity through assessment. *British Educational Research Journal, 25*(3), 343-54.

Reynolds, M., & Trehan, K. (2000). Assessment: A critical perspective. *Studies in Higher Education, 25*(3), 267-278.

Ritter, L. (2000). The quest for an effective form of assessment: The evolution and evaluation of a controlled assessment procedure (CAP). *Assessment and Evaluation in Higher Education, 25*(4), 307-320.

Roderick, M., & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low achieving pupils to high-stakes testing. *Educational Evaluation and Policy Analysis, 23*, 197-228.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Ross, M., Radnor, H., Mitchell, S., & Bierton, C. (1993). *Assessing achievement in the arts*. Buckingham: Open University Press.

Ross, J. A., Rolheister, C., & Hogaboam-Gray, A. (2002). Influences on student cognitions about evaluation. *Assessment in Education, 9*(1), 81-95.

Ruthven, K. A. (2002). Assessment in mathematics education. In L. Haggerty (Ed.), *Teaching mathematics in secondary schools* (pp. 176-191). London: Routledge Falmer.

Ryan, G., & Bernard, R. (2003). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (2nd ed., pp. 259-309). London: Sage Publications.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Sciences, 18*, 119-144.

Sadler, R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education, 5*(1), 77-84.

Schraw, G., Horn, C., Thorndike-Christ, T., & Bruning, R. (1995). Academic goal orientations and student classroom achievement. *Contemporary Educational Psychology, 20*(3), 359-368.

Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33*, 359-382.

Sebatane, E. M. (1998). Assessment and classroom learning: a response to Black and Wiliam. *Assessment in Education: Principals Policy & Practice, 5*(1), 123-130.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4-13.

Sfard, A. (2003). Balancing the unbalanceable: The NCTM standards in light of theories of learning mathematics. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 353-392). Reston, VA: National Council of Teachers of Mathematics.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 1-14.

Silverman, D. (2003). Analysing talk and text. In N. K. Denzin & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (2nd ed., pp. 340-362). London: Sage Publications.

Skinner, B. F. (1976). *About behaviorism*. New York: Vintage Books.

Smith, E., & Gorard, S. (2005). 'They don't give us our marks': The role of formative feedback in student progress. *Assessment in Education: Principles, Policy & Practice, 12*(1), 21-38.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.

Stefani, L. A. J. (1994). Peer, self and tutor assessment: Relative reliabilities. *Studies in Higher Education, 19*, 69-75.

Stenhouse, L. (1988). Case study methods. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (pp. 48-54). Oxford: Pergammon Press.

Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Strachan, J. (2001). Assessment in change: Some reflections on the local and international background to the National Certificate of Educational Achievement (NCEA). *New Zealand Annual Review of Education, 11*, 245-274.

Sturman, A. (1999). Case study methods. In J. P. Keeves & G. Lakomski (Eds.), *Issues in Educational Research* (1st ed., pp. 103-112). Oxford; New York: Pergamon.

Tanner, H., & Jones, S. (2003). Self-efficacy in mathematics and students' use of self-regulated learning strategies during assessment events. In N. A. Pateman, B. J. Dougherty, & J. T. Zilliox (Eds.), *Proceedings of the 27th conference of the International Group for the Psychology of Mathematics Education held jointly with the 25th conference of PME-NA* (pp. 275-282): University of Hawai'i.

Thorndike, E. L. (1924). Mental discipline in high school studies. *Journal of Educational Psychology, 5*(1), 1-22.

Tiknaz, Y., & Sutton, A. (2006). Exploring the role of assessment tasks to promote formative assessment in Key Stage 3 Geography: Evidence from twelve teachers. *Assessment in Education: Principles, Policy & Practice, 13*(3), 327-343.

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom.* Buckingham: Open University Press.

Tripp, D. (1985). Case study generalisations: An agenda for action. *British Educational Research Journal, 11*(1), 395-408.

Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal, 22*(4), 389-404.

Vaughan, K. (2003). Changing lanes: Young people making sense of pathways. In B. Webber (Ed.), *Educating for the 21st century: Rethinking the educational*

outcomes we want for young New Zealanders. Conference Proceedings (pp. 3-17). Wellington: New Zealand Council for Educational Research.

Vaughn, S., Shay Schumm, J., & Sinagub, J. (1996). *Focus group interviews in education and psychology*. Thousand Oaks, CA: Sage Publications.

Vlaardingerbroek, B. (2006). Transition to tertiary study in New Zealand under the National Qualifications Framework and 'the ghost of 1888'. *Journal of Further and Higher Education, 30*(1), 75-85.

von Glaserfeld, E. (1995). A constructivist approach to teaching. In L. P. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 3-16). Hillsdale, N.J.: Lawrence Erlbaum.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* MA: Harvard University Press.

Watson, A. (1998). What makes a mathematical performance noteworthy in informal teacher assessment? In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd annual conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 169-176). Stellenbosch, South Africa: PME.

Watson, A. (2002). Instances of mathematical thinking among low attaining students in an ordinary secondary classroom. *Journal of Mathematical Behavior, 20*, 461-475.

Watson, A. (2004). Paradigmatic conflicts in informal mathematics assessment as sources of social inequity. In B. Allen & S. Johnstone-Wilder (Eds.), *Mathematics education: Exploring the culture of learning* (pp. 57-68). London; New York: RoutledgeFalmer.

Watson, A. (2006). Some difficulties in informal assessment in mathematics. *Assessment in Education: Principles, Policy & Practice, 13*(3), 289-303.

Watt, H. (2005). Attitudes to the use of alternative assessment methods in mathematics: A study with secondary mathematics teachers in Sydney, Australia. *Educational Studies in Mathematics, 58*(1), 21-44.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71*(1), 3-25.

West, A. (2001). A vital element in the knowledge society. *QA News, 38*, 3.

Wiliam, D. (1999). Formative assessment in mathematics. Part 2: Feedback. *Ennals, 5*(3), 8-11.

Wiliam, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment. *British Educational Research Journal, 22*(5), 537-548.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice, 11*(1), 49-65.

Wilson, S. M., & Floden, R. E. (2001). Hedging bets: Standards-based reform in classrooms. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the States* (pp. 193-216). Chicago: The University of Chicago Press.

Wood, D. (1998). *How children think and learn: The social context of cognitive development.* (2nd ed.). Oxford: Blackwell.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17*(2), 89-100.

Yin, R. K. (1993). *Applications of case study research*. Newbury Park, CA: Sage Publications.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). London: Sage Publication.

Zepke, N., Leach, L., Brandon, J., Chapman, J., Neutze, G., Rawlins, P., et al. (2006). *Standards-based assessment in the senior secondary school: A research synthesis.* Wellington, New Zealand: New Zealand Qualifications Authority.

Appendix A: Students' information sheet

(N.B. The original documents were supplied on Massey Letterhead)

Students' perceptions of the formative nature of NCEA. Information sheet for students and their parents/guardians.

Dear Student

My name is Peter Rawlins and I am currently doing Doctoral research into student's perceptions of the formative nature of mathematics assessments used as part of the new National Certificate of Educational Assessment (NCEA).

Overview of the research
With the introduction of the new NCEA qualification, there is a need for more research on the nature of assessment feedback and how students use it to influence future learning. Your voice is very important as you are in a unique position to help us understand the impact of assessment on you, the student. I'm interested in hearing what experiences you have had from your mathematics assessments. What types of feedback do you get and how do you use this feedback?

Involvement in this project will provide you with an opportunity to reflect on how you use assessment feedback to help you learn. Hopefully this will make your senior secondary school years more rewarding for you. Your participation in this research will contribute to our understandings of the complex role of feedback in student learning.

Why me?
You have been invited to take part in this research because you are currently in the Year 12 mathematics class taught by the teacher who has agreed to take part in this pilot project. Your participation is voluntary and is independent of any assessment procedures associated with your course of study. Although this research will focus on your perspectives it will gather some general data from the classroom so that a deeper understanding of your views can be gained.

What will I be asked to do?
There are two levels that will involve you in this project.

1. I will be conducting a series of classroom observations of your year 12 mathematics class. I will sit in the back of your mathematics classroom and make some background notes on the interactions involving assessment issues between students and your teacher. This will involve no time commitment from you other than your normal class time.

2. You may wish to be involved in the case study part of the project. This will involve you in two data gathering processes. The first of these is a focus group interview each term involving five students from your year 12 class. I'll ask you to talk about some of your perceptions of the mathematics assessments you have recently completed. Each focus group interview will last about 30 minutes at I'll arrange a mutually acceptable time. With your permission, I'll record the interview on audio tape and make a transcription of the recording. In the second process I will send you a series of questions for you to consider and answer in your own time. This will allow you to expand on anything mentioned in the focus group or to talk about things that did not arise during the interview. I would like to do this via email so that I can follow up on any of your experiences if I need to, but I am also happy to post the questions to you if you would prefer. The things you tell me will be confidential, but I will be sending copies of the questions to you to your parents/ guardians, so they can be kept up to date on the project.

All the data will be held securely in my University office and not made available to anyone without participants' written permission.

The data will be analysed and a summary of the findings made available to those who participated in the interview. This will give you an opportunity to correct or clarify any interpretation of the interview data. At the completion of the project a summary of the research report will be sent to your school. Your mathematics teacher will inform you when this is available. If you or your parents/guardians wish to read this summary please talk to the Head of the Mathematics Department.

What rights do I have?
In accordance with the requirement of the Massey University Ethics Committee you have the right:

- to decline to participate;
- to refuse to answer any particular questions;
- to withdraw from the study at any time;
- to ask any questions about the study at any time during participation;
- to provide information on the understanding that your name will not be used unless you give permission to the researcher;
- to be given access to a summary of the findings of the study when it is concluded;
- to ask for the audio tape to be turned off at any time during the focus group interview.

When any research is conducted it must be recognised that there is always a risk of a breach of confidentiality and that I can only give an assurance of confidentiality and anonymity to the extent allowed by law. It should be noted, however, that there is a clear expectation that all participants, including the researcher, will respect any information shared through the research process and will treat it with confidentiality. Neither the school nor any individuals will be identified either directly or indirectly in verbal or written form. Where direct quotes from the interview tapes or written correspondence are used in subsequent publications pseudonyms will be assigned to maintain anonymity.

What am I asking you to do?
The school Board of Trustees and your mathematics teacher have given me permission to observe your mathematics classroom. If, however, you would prefer that no observational data is recorded for you, please tick the appropriate box on the consent form and return it in a sealed envelope to your mathematics teacher. A nil reply implies consent.

If after reading this information sheet you are willing to be involved in the case study part of the research project, can you please complete the consent form and return it in a sealed envelope to your mathematics teacher. If more than five students from your class consent to be part of the case study I will select a representative sample and you will be informed in writing whether you have been selected to take part or not.

I would like to thank you for your careful consideration of this opportunity. Should you require further clarification please feel free to contact me or either of my supervisors here at Massey University.

Researcher: Peter Rawlins, Ph (06) 350 5799 ext. 8855 P.Rawlins@massey.ac.nz
Supervisors: Associate Professor Glenda Anthony, ext.8600 G.J.Anthony@massey.ac.nz
Dr. Jenny Poskitt, ext 8293 J.M.Poskitt@massey.ac.nz

This project has been reviewed and approved by the Massey University Human Ethics Committee, PN Application 04/168. If you have any concerns about the conduct of this research, please contact Professor Sylvia V Rumball, Chair, Massey University Human Ethics Committee: Palmerston North, telephone 06 350 5249, email humanethicspn@massey.ac.nz.

Appendix B: Students' consent form

(N.B. The original documents were supplied on Massey Letterhead)

Students' perceptions of the formative nature of NCEA.

STUDENT CONSENT FORM

THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF FIVE (5) YEARS

I have read the Information Sheet and have had the details of the study explained to me. My questions have been answered to my satisfaction, and I understand that I may ask further questions at any time. I agree to keep confidential all information concerning the project 'Students' perceptions of the formative nature of NCEA.'

I **do/ do not** (Please circle one) want the researcher to record any observational data on me during the class observations.

I **would/would not** (Please circle one) like to participate in the case study part of the research under the conditions set out in the Information Sheet.

I **agree/do not agree** (Please circle one) to the interview being audio taped.

Signature of student: ..... **Date:** .....

Full Name - printed .....

Signature of parent/guardian: ..... **Date:** .....

Full Name - printed .....

Appendix C: Board of Trustees' information sheet

(N.B. The original documents were supplied on Massey Letterhead)

To the Chairperson
Board of Trustees
Euclid College
Mathsville

Dear Sir.

Request for permission to conduct research into:
Students' perceptions of the formative nature of NCEA.

My name is Peter Rawlins and I am currently doing Doctoral research into students' perceptions of the formative nature of mathematics assessments used as part of the new National Certificate of Educational Assessment (NCEA). I am writing to request permission to conduct a research project in your school during 2005.

Overview of the research
In 2002 the new NCEA qualification began its implementation in New Zealand Schools. With the move to increased internal assessment and opportunities for reassessment, there is a need for research on the nature of assessment feedback and how students use it to influence future learning. Involvement in this project will provide the teachers, students and the school with an opportunity to reflect on current assessment practices in light of the implementation of NCEA and the findings will contribute to our understandings of the complex role of feedback in student learning.

This research will concentrate on the mathematics classroom and will investigate the following questions:

1. What feedback do students receive prior to, and subsequent to, an assessment event and in what ways do students interpret this feedback?
2. In what ways do student use any feedback given to influence future learning?

Participants
A case study methodology will be used to explore the research questions with two groups of five students from two separate year 12 mathematics classes. Although this research will focus on students' perspectives it will gather some data from the classroom as well as the classroom teacher so that a deeper understanding of the students' views can be gained. In consultation with the HOD mathematics I would like to ask two teachers from your school and their year 12 mathematics classes to be involved in this project as outlined below. Students from within these two classes will be invited to take part in the case study part of the research.

Phases of the research.
There are five phases to this project. The first two of these form the main part of the case study methodology, the last three stages are used for triangulation purposes.

3. Four focus group interviews will be conducted, one per term, with each group of 5 participant students. These students will be asked to respond to questions posed by the researcher regarding their experiences with assessment feedback as set out in the research questions (an overview of the interview schedule is attached). Each focus group interview will last about 30 minutes at a mutually acceptable time arranged with the participating students. It is intended that this research will not interfere with students' in-class time.

4. An email or postal conversation with the students will be established for the duration of the research project. This will allow students to respond to focus questions and explore their personal experiences with assessment allowing for individual and reflective responses that the focus group interview process may not bring out. This will be time effective for the students as they can respond at their convenience with no loss of class time. Students' responses will remain confidential but to ensure their personal safety, copies of all correspondence sent to the students will also be sent to the parents/caregivers.

5. A series of two class observations per term, for four terms, with each of the two year 12 mathematics class will be conducted. The purpose of this is to gain general background data on the interactions involving assessment issues between the students and their teacher. My role will be non-participatory and will involve written observational notes only. There will be no time commitment from the students other than their normal class time. Students who would prefer that no observational data is recorded for them have the opportunity to indicate this on the consent form.

6. An individual interview with each class teacher will be conducted so that general background information about classroom and departmental assessment practices can be gathered (an overview of the interview schedule is attached).

7. A document search of the school and the departmental policies on NCEA assessment. This search will be conducted by the researcher and can be conducted at the school if you would prefer. It may be necessary to spend a small amount of time with the Senior Manager in charge of NCEA assessment or Head of Department if any clarification is needed by the researcher.

Interviews will be recorded on audio tape, with the permission of the participants, and I will make transcriptions of the recordings. All data will be held securely in my University office and not made available to anyone without participants' written permission.

The data will be analysed and summaries of the findings sent to the relevant participants to clarify any interpretation of the data. At the completion of the project a summary of the research report will be prepared and sent to the school early in 2006. Students will be advised that this summary is available for them to read should they wish to.

Rights of the participants
In accordance with the requirement of the Massey University Ethics Committee participants have the right:

- to decline to participate;
- to refuse to answer any particular questions;
- to withdraw from the study at any time;
- to ask any questions about the study at any time during participation;
- to provide information on the understanding that their name will not be used unless they give permission to the researcher;
- to be given access to a summary of the findings of the study when it is concluded;
- to ask for the audio tape to be turned off at any time during the focus group interview.

When any research is conducted it must be recognised that there is always a risk of a breach of confidentiality and that I can only give an assurance of confidentiality and anonymity to the extent allowed by law. It should be noted, however, that there is a clear expectation that all participants, including the researcher, will respect any information shared through the research process and will treat it with confidentiality. Neither the school nor any individuals will be identified either directly or indirectly in verbal or written form. Where direct quotes from the interview tapes or written correspondence are used in subsequent publications pseudonyms will be assigned to maintain anonymity.

Permission from the students
I have written separate information and consent forms for the students and teacher (attached). To facilitate the consent process for the class observations (phase 2 above), I have asked students who do not wish to have observational data recorded for them to indicate this in the appropriate place on the consent form and return it, in a sealed envelope, to their mathematics teacher. A nil return implies consent. Students who wish to be involved in the case study part of the project should also indicate this on the consent form and return it in a sealed envelope to their mathematics teacher. If more than five students consent to be part of the case study I will select a representative sample and participants will be informed in writing whether they have been selected to take part or not..

This project has been reviewed and approved by the Massey University Human Ethics Committee, PN Application 04/168. If you have any concerns about the conduct of this research, please contact Professor Sylvia V Rumball, Chair, Massey University Human Ethics Committee: Palmerston North, telephone 06 350 5249, email humanethicspn@massey.ac.nz.

Should you require further clarification please feel free to contact me or either of my supervisors here at Massey University.

Researcher: Peter Rawlins, Ph (06) 350 5799 extn. 8855, P.Rawlins@massey.ac.nz

Supervisors:
Associate Professor Glenda Anthony
Ph: (06) 350 5799 extn. 8600
Email: G.J.Anthony@massey.ac.nz

Dr. Jenny Poskitt
(06)350 5799 ext 8293
J.M.Poskitt@massey.ac.nz

I would like to thank you in advance for your careful consideration of this opportunity. Once you have made your decision could you please complete the attached consent form and return it to me in the prepaid envelope supplied.

Appendix D: Teachers' information sheet

(N.B. The original documents were supplied on Massey Letterhead)

Students' perceptions of the formative nature of NCEA.
Information sheet for teachers.

My name is Peter Rawlins and I am currently doing Doctoral research into students' perceptions of the formative nature of mathematics assessments used as part of the new National Certificate of Educational Assessment (NCEA).

Overview of the research
In 2002 the new NCEA qualification began its implementation in New Zealand Schools. With the move to increased internal assessment and opportunities for reassessment, there is a need for research on the nature of assessment feedback and how students use it to influence future learning. Involvement in this project will provide the teachers, students and the school with an opportunity to reflect on current assessment practices in light of the implementation of NCEA and the findings will contribute to our understandings of the complex role of feedback in student learning.

This research will concentrate on the mathematics classroom and will investigate the following questions:

3. What feedback do students receive prior to, and subsequent to, an assessment event and in what ways do students interpret this feedback?
4. In what ways do student use any feedback given to influence future learning?

Participants
A case study methodology will be used to explore the research questions with two groups of five students from two separate year 12 mathematics classes. Although this research will focus on students' perspectives it will gather some data from the classroom as well as the classroom teacher so that a deeper understanding of the students' views can be gained. In consultation with the HOD mathematics I would like to ask two teachers from your school and their year 12 mathematics classes to be involved in this project as outlined below. Students from within these two classes will be invited to take part in the case study part of the research.

Phases of the research.
There are five phases to this project however, as a teacher you will only be involved in two phases.

8. A series of three class observations per term, for four terms, with each of the two year 12 mathematics class will be conducted. The purpose of this is to gain general background data on the interactions involving assessment issues between the students and their teacher. My role will be non-participatory and will involve written observational notes only. There will be no time commitment from the students other than their normal class time. Students who would prefer that no observational data is recorded for them have the opportunity to indicate this on the consent form.

9. An individual interview with each class teacher will be conducted so that general background information about classroom and departmental assessment practices can be gathered (an overview of the interview schedule is attached).

Interviews will be recorded on audio tape, with the permission of the participants, and I will make transcriptions of the recordings. All data will be held securely in my University office and not made available to anyone without participants' written permission.

The data will be analysed and summaries of the findings sent to the relevant participants to clarify any interpretation of the data. At the completion of the project a summary of the research report will be prepared and sent to the school early in 2006. You will be advised that this summary is available for you to read should you wish to.

Rights of the participants
In accordance with the requirement of the Massey University Ethics Committee participants have the right:

- to decline to participate;
- to refuse to answer any particular questions;
- to withdraw from the study at any time;
- to ask any questions about the study at any time during participation;
- to provide information on the understanding that their name will not be used unless they give permission to the researcher;
- to be given access to a summary of the findings of the study when it is concluded;
- to ask for the audio tape to be turned off at any time during the focus group interview.

When any research is conducted it must be recognised that there is always a risk of a breach of confidentiality and that I can only give an assurance of confidentiality and anonymity to the extent allowed by law. It should be noted, however, that there is a clear expectation that all participants, including the researcher, will respect any information shared through the research process and will treat it with confidentiality. Neither the school nor any individuals will be identified either directly or indirectly in verbal or written form. Where direct quotes from the interview tapes or written correspondence are used in subsequent publications pseudonyms will be assigned to maintain anonymity.

Consent
In order to provide a cross section of students' perceptions it will be important to look at a class with a range of student abilities. If you feel that your class would make interesting subjects for this research and you are interested in being involved, could you please indicate this to your Head of Department. If may be necessary for me to select from a number of people who have indicated a willingness to be involved and this will be done in consultation with the Head of Department and any teachers who have indicated their interest. If you are selected to take part in this research you will be asked to complete the attached consent form.

This project has been reviewed and approved by the Massey University Human Ethics Committee, PN Application 04/168. If you have any concerns about the conduct of this research, please contact Professor Sylvia V Rumball, Chair, Massey University Human Ethics Committee: Palmerston North, telephone 06 350 5249, email humanethicspn@massey.ac.nz.

Should you require further clarification please feel free to contact me or either of my supervisors here at Massey University.

Researcher: Peter Rawlins, Ph (06) 350 5799 extn. 8855, P.Rawlins@massey.ac.nz

Supervisors:
Associate Professor Glenda Anthony
Ph: (06) 350 5799 extn. 8600
Email: G.J.Anthony@massey.ac.nz

Dr. Jenny Poskitt
(06)350 5799 ext 8293
J.M.Poskitt@massey.ac.nz

Appendix E: Board of Trustees' information sheet for AEQ

(N.B. The original documents were supplied on Massey Letterhead)

To the Principal
Euclid College

Dear Mr. McKay

Request for permission to conduct additional research into:
Students' perceptions of the formative nature of NCEA.

My name is Peter Rawlins and I am currently conducting Doctoral research with a group of Euclid college year 12 mathematics students. I am writing to request permission to conduct some additional background research. The nature of this research is an Assessment Experience Questionnaire which all students from the three year 12 Traditional Mathematics classes will be invited to complete during their normal mathematics class. I anticipate that this will take no more than 10 to 15 minutes. I have received verbal approval from the teachers concerned.

The purpose of this questionnaire is to provide a fuller picture of student's experiences of assessment and to examine the representative nature of the research cohort. I have attached a copy of the questionnaire for your information.

I have written a separate information sheet for the students explaining the rationale for the questionnaire and also that there is no compulsion for them to complete it if they do not wish to.

This main research project has been reviewed and approved by the Massey University Human Ethics Committee, PN Application 04/168. If you have any concerns about the conduct of this research, please contact Professor Sylvia V Rumball, Chair, Massey University Human Ethics Committee: Palmerston North, telephone 06 350 5249, email humanethicspn@massey.ac.nz.

Should you require further clarification please feel free to contact me or either of my supervisors here at Massey University.

Researcher: Peter Rawlins, Ph (06) 350 5799 extn. 8855, P.Rawlins@massey.ac.nz

Supervisors:
Associate Professor Glenda Anthony
Ph: (06) 350 5799 extn. 8600
Email: G.J.Anthony@massey.ac.nz

Dr. Jenny Poskitt
(06)350 5799 ext 8293
J.M.Poskitt@massey.ac.nz

I would like to thank you in advance for your careful consideration of this opportunity. Once you have made your decision could you please sign the attached consent form and return it in the prepaid envelope.

Yours sincerely

Appendix F: Students' information for AEQ

(N.B. The original documents were supplied on Massey Letterhead)

Students' perceptions of the formative nature of NCEA. Information sheet for Assessment Experience Questionnaire

Dear Student

My name is Peter Rawlins and you may remember that I am currently conducting some research into student's perceptions of the formative nature of mathematics assessments used as part of the new National Certificate of Educational Assessment (NCEA). I have been working with a smaller group of your classmates and would like to examine whether their views are representative of all Y12 mathematics students. This is also an opportunity for you add your voice and the data will allow me to gain a fuller picture of all Y12 students' experiences of assessment.

I would like to ask you to take 10 to 15 minutes to complete the attached questionnaire and return it in the envelope supplied to the school office. I would appreciate it if you could complete the questionnaire without talking to your classmates.

As you can see the questionnaires are anonymous so you will not be identified in any way.

In accordance with the requirement of the Massey University Ethics Committee you have the right:

- to decline to participate;
- to refuse to answer any particular questions;
- to ask any questions about the study at any time during participation;

If you agree to participate then please read each statement and circle the number on the left that best expresses how you feel about that statement. If you would prefer not to participate then please leave your questionnaire blank.

Researcher: Peter Rawlins, Ph (06) 350 5799 ext. 8855 P.Rawlins@massey.ac.nz
Supervisors: Associate Professor Glenda Anthony, ext.8600 G.J.Anthony@massey.ac.nz
Dr. Jenny Poskitt, ext 8293 J.M.Poskitt@massey.ac.nz

This research project has been reviewed and approved by the Massey University Human Ethics Committee, PN Application 04/168. If you have any concerns about the conduct of this research, please contact Dr John O'Neill, Chair, Massey University Human Ethics Committee: Palmerston North, telephone 06 350 5799 x 8635, email humanethicspn@massey.ac.nz.".

Appendix G: Focus group overview

Overview of interview schedule for focus group interviews and questions for individual email/postal communication.

This research will utilise a semi-structured interview format. The following is a list of the general themes I want to explore in the **first focus group interview** and also in the **email/postal communication** with the individual students.

1. What is the nature of the verbal feedback given to students with reference to achievement levels on the AS:
   - Class, individually or in small groups?
   - During the teaching of the unit?
   - When the marked assessment is handed back?
   - How frequently does the teacher refer to the AS?

2. What is the nature of the written feedback students receive from assessment tasks and in what ways do students interpret this feedback?
   - Grades and/or comments
   - Just what is wrong or what needs to be done next?

3. What does the teacher talk about when they give marked assessments back?
   - Whole assessment question by question, or just concentrate on areas of common concern?
   - Correct answers only or common mistakes that were made?

4. In what ways do student use any feedback given to influence future learning?

In subsequent **focus group interviews** and **email/postal communications** I will be examining similar issues but I will be particularly interested in what things have changed since the last focus group interview or email/postal communication.

Appendix H: Example of follow up questionnaire to students

(N.B., the spaces for the students to write their responses after the questions have been removed for brevity.)

Hi everyone.

Thanks for your involvement in the project so far. I got lots of really good material from the first focus group interview. I was impressed by how articulate and reflective you all were.

This series of questions is designed to compliment the data gathered from the focus group interviews. I have designed these questions based on your responses, but please feel free to add anything that you feel I might have missed out on or misinterpreted.

If you could answer the questions and send the email back to me in the next week that would be great. Just enter your answers below each question.

1. What do you do to prepare for an assessment task?

2. Do you look at the 'l can do sheet' at the back of the blue book?

3. How important is it that that you understand the mathematical material?

4. How important is it that you get a grade, such as Merit or Excellence, that really reflects what you know? Why is this important to you?

5. What sorts of things would you like the teacher to do differently leading up to an assessment?

6. What sorts of comments are written on your assessment scripts, and do you find these useful?

7. What do you do when you get your assignments and assessment scripts back?

8. Do you discuss any of the questions or answers with other students when you get your marked script back and if so what sorts of things do you discuss?

9. Do you ever get the opportunity to ask the teacher why you got a particular question wrong and what you needed to do to get it right? Do you ever take this opportunity?

10. What would you like your teachers to write on your marked scripts that they don't currently write?

11. I heard your teacher stress that it was important for you to be sure that there were no questions about the material that were left unanswered and that you

could go to learning support to get some help. Did anybody go or are you intending to go, either now or closer to the exams?

12. When you are preparing for the next time you have to sit the same Standard do you use your previously marked scripts when you are studying?

13. How could the 'I can do' sheet be helpful?

14. Is there anything else that you didn't get the chance to say during the focus group interview and would like to say now?

Thanks again for taking the time to do this.
Hope you have a great holiday.

Cheers
Peter

Appendix I: Assessment Experience Questionnaire

Assessment Experience Questionnaire (AEQ)

Please answer every item quickly by giving your immediate response.
Circle the appropriate code number to show your response to assessment.

	strongly disagree	disagree	?	agree	strongly agree
1 Amount and distribution of study effort					
I do the same amount of study each week, regardless of whether an assignment is due or not.	1	2	3	4	5
I can be quite selective about what I study and learn and still do well.	5	4	3	2	1
I only study things that are going to be covered in the assignments.	5	4	3	2	1
I have to study regularly if I want to do well on the course.	1	2	3	4	5
On this course, it is possible to do quite well without studying much.	5	4	3	2	1
In weeks when the assignments are due I put in many more hours.	5	4	3	2	1
2 Assignments and learning					
Tackling the assignments really makes me think.	1	2	3	4	5
I learn more from doing the assignments than from studying the course material.	1	2	3	4	5
In completing the assignments you can get away with not understanding and still get high marks.	5	4	3	2	1
The assignments give very clear instructions about what you are expected to do.	1	2	3	4	5
When I tackle an assignment it is not at all clear what would count as a successful answer.	5	4	3	2	1
The assignments are not very challenging.	5	4	3	2	1
3 Quantity and timing of feedback					
On this course I get plenty of feedback on how I am doing.	1	2	3	4	5
The feedback comes back very quickly.	1	2	3	4	5
There is hardly any feedback on my assignments when I get them back.	5	4	3	2	1
When I get things wrong or misunderstand them I don't receive much guidance in what to do about it.	5	4	3	2	1
I would learn more if I received more feedback.	5	4	3	2	1
Whatever feedback I get comes too late to be useful.	5	4	3	2	1

	strongly disagree	disagree	?	agree	strongly agree
4 Quality of feedback					
The feedback mainly tells me how well I am doing in relation to others.	5	4	3	2	1
The feedback helps me to understand things better.	1	2	3	4	5
The feedback shows me how to do better next time.	1	2	3	4	5
Once I have read the feedback I understand why I got the mark I did.	1	2	3	4	5
I don't understand some of the feedback.	5	4	3	2	1
I can seldom see from the feedback what I need to do to improve.	5	4	3	2	1
5 What you do with the feedback					
I read the feedback carefully and try to understand what the feedback is saying.	1	2	3	4	5
I use the feedback to go back over what I have done in the assignment.	1	2	3	4	5
The feedback does not help me with any subsequent assignments.	5	4	3	2	1
The feedback prompts me to go back over material covered earlier in the course.	1	2	3	4	5
I do not use the feedback for revising.	5	4	3	2	1
I tend to only read the marks.	5	4	3	2	1
6 The examination and learning (only to be completed if there is an exam)					
Preparing for the exam was mainly a matter of memorising.	5	4	3	2	1
Doing the exam brought things together for me.	1	2	3	4	5
I learnt new things while preparing for the exam.	1	2	3	4	5
I understand things better as a result of the exam.	1	2	3	4	5
I'll probably forget most of it after the exam.	5	4	3	2	1
In the exam you can get away with not understanding and still get good marks.	5	4	3	2	1

Comments you would like to make about the way the assessment affected your learning on the course

Mean and percentage frequency distribution for Assessment Experience Questionnaire	Research Students n=9						All students n=59					
	Mean	% freq 1	2	3	4	5	Mean	% freq 1	2	3	4	5
1. Amount and distribution of study effort												
I do the same amount of study each week, regardless of whether an assignment is due or not.	2.2	22%	56%	0%	22%	0%	2.4	17%	53%	8%	20%	2%
I can be quite selective about what I study and learn and still do well.	2.6	0%	56%	33%	11%	0%	2.7	5%	45%	28%	21%	2%
I only study things that are going to be covered in the assignments.	2.6	11%	56%	0%	33%	0%	2.5	12%	51%	14%	20%	3%
I have to study regularly if I want to do well on the course.	4.0	0%	0%	22%	56%	22%	3.6	4%	16%	16%	42%	23%
On this course, it is possible to do quite well without studying much.	3.6	0%	22%	11%	56%	11%	3.5	0%	25%	15%	45%	15%
In weeks when the assignments are due I put in many more hours.	1.9	33%	56%	0%	11%	0%	2.2	27%	46%	8%	15%	3%
2. Assignments and Learning												
Tackling the assignments really makes me think.	4.2	0%	0%	0%	78%	22%	3.9	0%	3%	17%	66%	14%
I learn more from doing the assignments than from studying the course material.	3.6	0%	13%	25%	50%	13%	2.9	7%	25%	40%	23%	5%
In completing the assignments you can get away with not understanding and still get high marks.	3.8	11%	0%	11%	56%	22%	3.4	5%	19%	16%	47%	14%
The assignments give very clear instructions about what you are expected to do.	3.1	0%	33%	22%	44%	0%	3.4	5%	12%	24%	56%	3%
When I tackle an assignment it is not at all clear what would count as a successful answer.	3.0	0%	33%	33%	33%	0%	3.1	5%	22%	38%	31%	3%
The assignments are not very challenging.	3.6	0%	11%	44%	22%	22%	3.6	2%	10%	22%	54%	12%
3. Quantity and timing of feedback												
On this course I get plenty of feedback on how I am doing.	3.7	0%	22%	22%	22%	33%	3.6	3%	14%	22%	46%	15%
The feedback comes back very quickly.	4.0	0%	11%	0%	67%	22%	3.2	7%	24%	27%	32%	10%
There is hardly any feedback on my assignments when I get them back.	3.6	11%	11%	11%	44%	22%	3.6	3%	14%	20%	47%	15%
When I get things wrong or misunderstand them I don't receive much guidance in what to do about it.	3.8	0%	33%	0%	22%	44%	3.5	5%	19%	16%	40%	19%
I would learn more if I received more feedback.	2.7	22%	33%	11%	22%	11%	2.4	17%	42%	27%	12%	2%
Whatever feedback I get comes too late to be useful.	3.9	11%	0%	11%	44%	33%	3.6	5%	12%	19%	49%	15%

4. Quality of feedback												
The feedback mainly tells me how well I am doing in relation to others.	3.2	11%	11%	33%	33%	11%	3.2	4%	23%	32%	39%	4%
The feedback helps me to understand things better.	4.0	0%	11%	0%	67%	22%	3.8	0%	11%	7%	72%	11%
The feedback shows me how to do better next time.	4.1	0%	0%	22%	44%	33%	3.9	0%	7%	12%	67%	14%
Once I have read the feedback I understand why I got the mark I did.	4.0	0%	11%	0%	67%	22%	3.7	2%	11%	14%	63%	11%
I don't understand some of the feedback.	3.2	11%	11%	22%	56%	0%	3.0	2%	34%	27%	34%	4%
I can seldom see from the feedback what I need to do to improve.	3.6	11%	0%	11%	78%	0%	3.3	5%	16%	25%	49%	5%
5. What you do with the feedback												
I read the feedback carefully and try to understand what the feedback is saying.	4.2	0%	0%	0%	78%	22%	4.0	0%	5%	11%	67%	18%
I use the feedback to go back over what I have done in the assignment	4.0	0%	0%	22%	56%	22%	3.8	0%	7%	23%	54%	16%
The feedback does not help me with any subsequent assignments.	3.9	0%	0%	33%	44%	22%	3.7	2%	7%	21%	54%	16%
The feedback prompts me to go back over material covered earlier in the course.	3.4	0%	11%	33%	56%	0%	3.3	2%	19%	32%	39%	9%
I do not use the feedback for revising.	4.1	0%	0%	0%	89%	11%	3.7	2%	14%	12%	58%	14%
I tend to only read the marks.	3.9	0%	0%	22%	67%	11%	3.7	4%	11%	14%	56%	16%
6. The examination and learning												
Preparing for the exam was mainly a matter of memorising.	2.9	14%	29%	29%	14%	14%	3.1	8%	28%	25%	28%	11%
Doing the exam brought things together for me.	3.4	0%	14%	29%	57%	0%	3.3	4%	21%	23%	45%	8%
I learnt new things while preparing for the exam.	3.9	0%	14%	0%	71%	14%	3.6	2%	17%	10%	60%	12%
I understand things better as a result of the exam.	3.9	0%	0%	29%	57%	14%	3.7	2%	6%	25%	57%	11%
I'll probably forget most of it after the exam.	2.9	14%	14%	43%	29%	0%	2.4	25%	33%	23%	15%	4%
In the exam you can get away with not understanding and still get good marks	4.4	0%	0%	14%	29%	57%	4.0	4%	8%	8%	49%	32%

Appendix K: Research students' individual responses to AEQ

Item number	Brief code	Andrew	Chelsea	James	Kate	Sally	Megan	Raewyn	Tracy	Hannah	Average
1	amount	4	4	2	2	2	2	1	2	1	2.2
2	selective	2	2	3	4	2	2	3	3	2	2.6
3	covered	1	4	2	4	2	2	4	2	2	2.6
4	regularly	4	4	4	5	5	3	4	3	4	4.0
5	without	2	5	3	4	4	2	4	4	4	3.6
6	weeks	4	1	2	2	2	1	2	2	1	1.9
7	think	4	4	4	4	4	4	5	4	5	4.2
8	learn more	2		3	4	5	4	4	3	4	3.6
9	not under	4	4	3	4	4	1	4	5	5	3.8
10	instructions	4	4	3	4	2	2	3	4	2	3.1
11	successful	3	2	3	4	2	2	4	4	3	3.0
12	not challenge	5	3	3	5	3	2	4	4	3	3.6
13	lots feedback	4	3	2	4	3	2	5	5	5	3.7
14	quickly	4	2	4	5	4	4	4	4	5	4.0
15	non on ass	5	2	3	4	4	1	4	4	5	3.6
16	no guidance	4	2	2	5	5	2	4	5	5	3.8
17	learn more	4	1	2	2	3	1	4	2	5	2.7
18	late	5	1	4	5	4	3	4	4	5	3.9
19	social comp	5	1	3	4	2	4	4	3	3	3.2
20	understand	4	4	4	5	4	2	4	4	5	4.0
21	better	4	5	4	5	3	3	4	4	5	4.1
22	I know	4	5	4	5	4	2	4	4	4	4.0
23	don't under	4	2	3	4	3	1	4	4	4	3.2
24	can't see	4	1	3	4	4	4	4	4	4	3.6
25	read	5	5	4	4	4	4	4	4	4	4.2
26	review ass	4	5	4	4	5	4	4	3	3	4.0
27	not help	4	5	3	5	4	3	3	4	4	3.9
28	earlier	4	4	3	4	3	2	3	4	4	3.4
29	revsing	4	5	4	4	4	4	4	4	4	4.1
30	read marks	4	5	4	3	4	4	4	3	4	3.9
31	memory	2	1		2	3		4	5	3	2.9
32	links	3	4		4	4		2	3	4	3.4
33	learnt new	4	4		2	5		4	4	4	3.9
34	understand	3	4		3	5		4	4	4	3.9
35	forget	1	4		3	4		3	2	3	2.9
36	good marks	5	3		5	5		4	5	4	4.4
Factor Scores 1	Quality feedback	4.2	2.5	3.2	4.4	3.7	2.3	4.1	4.0	4.7	3.7
2	Use feedback	4.2	4.8	3.7	4.0	4.0	3.5	3.7	3.7	3.8	3.9
3	focus on assgts	2.3	3.8	3.0	4.3	3.3	2.3	3.8	3.0	3.0	3.2
4	learn from exam	3.3	4.0		3.0	4.7		3.3	3.7	4.0	3.7
5	dist of effort	3.4	2.5	3.0	3.8	3.2	2.6	3.6	3.2	3.0	3.1
6	apprch to exam	2.7	2.7		3.3	4.0		3.7	4.0	3.3	3.4

Appendix L: Students' results

Andrew

Code	Assessed	Mathematics - Level 2	Ver[18]	Credit Value	Result	Credits Achieved
5243[19]	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	A	3
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284[20]	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	A	4
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	M	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	N	
90289	Internal	Simulate probability situations, and apply the normal distribution (2.6)	1	2	M	2

Chelsea

Code	Assessed	Mathematics - Level 2	Ver	Credit Value	Result	Credits Achieved
5243	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	A	3
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	A	4
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	A	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	N	
90289	Internal	Simulate probability situations, and apply the normal distribution (2.6)	1	2	A	2

[18] Indicates the Version of the Standard assessed
[19] The 5000 series Standards are Unit Standards and either receive N: Not Achieved; A: Achieved.
[20] The 90,000 series Standards are Achievement Standards and receive N: Not Achieved; A: Achieve; M: Achieve with Merit; E: Achieved with Excellence

Hannah

Code	Assessed	Mathematics - Level 2	Ver	Credit Value	Result	Credits Achieved
5243	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	A	3
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	M	4
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	M	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	A	4
90289	Internal	Simulate probability situations, and apply the normal distribution (2.6)	1	2	E	2

James

Code	Assessed	Mathematics - Level 2	Ver	Credit Value	Result	Credits Achieved
5243	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	N	
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	N	
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	A	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	N	
90289	Internal	Simulate probability situations, and apply the normal distribution (2.6)	1	2	M	2

Kate

Code	Assessed	Mathematics - Level 2	Ver	Credit Value	Result	Credits Achieved
5243	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	A	3
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	E	4
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	E	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	M	4
90289	Internal	Simulate probability situations, and apply the normal distribution (2.6)	1	2	E	2

Megan

Code	Assessed	Mathematics - Level 2	Ver	Credit Value	Result	Credits Achieved
5243	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	A	3
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	A	4
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	A	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	A	4
90287	External	Use coordinate geometry methods (2.4)	2	2	A	2

Raewyn

Code	Assessed	Mathematics - Level 2	Ver	Credit Value	Result	Credits Achieved
5243	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	A	3
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	M	4
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	E	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	M	4
90289	Internal	Simulate probability situations, and apply the normal distribution (2.6)	1	2	E	2

Sally

Code	Assessed	Mathematics - Level 2	Ver	Credit Value	Result	Credits Achieved
5243	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	A	3
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	M	4
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	A	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	SNA[21]	
90289	Internal	Simulate probability situations, and apply the normal distribution (2.6)	1	2	E	2

[21] SNA indicates that although the students was entered for this achievement standard in the external examination, they did not write anything on the assessment task booklet and so the 'Standard was not assessed'

Tracy

Code	Assessed	Mathematics - Level 2	Ver	Credit Value	Result	Credits Achieved
5243	Internal	Apply mathematical processes and skills in problems	3	7	A	7
5245	Internal	Solve problems involving lines and points	3	2	A	2
5247	Internal	Make and evaluate statements about populations based on sample data	3	3	A	3
5248	Internal	Use sequences and series to solve problems	3	2	A	2
90284	External	Manipulate algebraic expressions and solve equations (2.1)	2	4	A	4
90285	External	Draw straightforward non-linear graphs (2.2)	2	3	A	3
90286	External	Find and use straightforward derivatives and integrals (2.3)	3	4	N	
90289	Internal	Simulate probability situations, and apply the normal distribution (2.6)	1	2	A	2

Appendix M; 'I can do' sheet for Algebra Standard

Achievement Standard 2.1 - 90284 **Manipulate algebraic expressions and solve equations**	I have Completed
Achievement: I can • Expand brackets up to three factors	
• Factorise expressions where there is a common factor	
• Factorise quadratics with coefficient of $x^2 = 1$	
• Factorise quadratics with coefficient of x^2 an integer	
• Understand the meaning of fractional indices	
• Understand the meaning of negative indices	
• Simplify expressions which have fractional indices	
• Simplify expressions which have negative indices	
• Change the subject of a formula	
• Understand the meaning of "logs"	
• Simplify expressions that include logs	
• Simplify rational expressions	
• Solve multi-step linear equations	
• Solve multi-step linear inequations	
• Solve quadratic equations that can be factorised	
• Solve polynomials in factorised form	
• Solve simple log equations	
• Solve simple exponential equations	
• Form and solve linear / linear simultaneous equations	
Merit: I can • Use the quadratic formula to solve equations	
• Solve linear / non linear simultaneous equations	
• Solve exponential equations like $13^{4x-5} = 6$	
• Solve problems using algebra skills	
Excellence: I Can	
• Understand the nature of roots of an equation	
• Use the roots of an equation to solve a problem in context	
• Choose and use algebra skills to solve problems	