The impact of middle leadership practices on student academic outcomes in New Zealand secondary schools

Camilla Highfield

Abstract

It has been reported in the literature that disparities in student achievement within schools is often larger than the disparity across schools. This study draws upon the existing literature on middle leadership and within-school variation to establish the relationship between effective middle leadership in secondary schools and student academic outcomes. Student academic outcomes are compared across and within 41 urban Auckland schools over a three year period to determine the extent of within- and across-school difference in English, mathematics and science for 15-year-olds. The academic results for students are then compared at department level for a subset of 10 schools where the middle leadership practices within each of those departments has been investigated.

In the New Zealand secondary school context, the public examination system for 15–17 year olds is the National Certificate of Educational Achievement (NCEA). This qualification system provides publicly available assessment data that can be used to compare student academic outcomes within and across schools and make comparisons to national norms. The New Zealand school 'decile' system, which groups schools on the basis of the socioeconomic status and student cultural background, is used for comparative purposes.

The first phase of this thesis describes the steps taken to determine the extent of within-school variation for 41 New Zealand secondary schools over a three year period. This was achieved by comparing the academic results for 15-year-olds from the English, mathematics and science departments within and across each of the schools and comparing those results to national averages. In some schools the results across departments with the same student cohort show considerable variation by subject and in other schools the results across departments are similar. The comparison of department results to national results categorised within a socioeconomic group provides insight into the extent of 'added value' at whole-school and department level. These data reinforce the argument that student achievement below expected results is not confined to schools working in poor socioeconomic circumstances.

In the second phase of the study, 10 of the 41 schools in Phase 1, self-select to take part in an investigation of middle leadership practices within their school. This involves members of the English, mathematics and science departments completing a quantitative questionnaire on the middle leadership practices that are known to impact positively on the quality of teaching that inevitably impacts on student achievement. The

results of these questionnaires produced five factors of middle leadership: *Collegial working environment, Goals and expectations, Focus on student academic results, Management of resources and Positive learning environment for students and teachers*. The questionnaire responses from each of the 30 departments across the 10 schools have been compared to the student academic outcomes data at three levels of the qualification to test the relationship between middle leadership factors and academic outcomes. Statistical tests were performed to substantiate evidence of the patterns and relationships and found that some factors have strong positive correlations with academic outcomes and others are negatively associated. Middle leadership practices had no relationship with the NCEA Level 1 (15-year-olds) results but had a strong relationship with the academic outcomes at the higher level of the qualification.

The multiple regression for Level 2 NCEA student achievement shows that decile of school can account for 62% of the variance in student achievement and when middle leadership practices are included in a simultaneous regression, 84% of the variance can be accounted for. Decile of school was a less important predictor for Level 3 NCEA student academic achievements (46%), but as with Level 2 when middle leadership practices were added the ability to predict student achievement rose to 62%. Therefore, a combination of school decile and middle leadership practices were a good overall predictor of student academic achievement at Levels 2 and 3 NCEA.

The positive predictors for student academic achievement were the factors *Goals and expectations*, *Management of resources* and *Positive learning environment for students and teachers*. Identifying the successful leadership practices evident in schools and departments that predict academic student outcomes will provide direction to policy makers and professional development practitioners working in the field.

Dedication

This thesis is dedicated to my immediate family. During the course of this project two people closest to me in my life have passed away. My only sibling and extraordinarily talented sister *Luci Grace Highfield* passed away on 6th March, 2009 aged 42, due to melanoma cancer. On 20th October, 2011 my divine and courageous husband *Ralph Hugh Toussaint* passed away aged 62. The burden and grief of their passing for myself, my parents Pam and Jim Highfield and my young daughters Arabella and Liberty Toussaint has been extraordinary. Despite these tragic and untimely deaths my family have never faltered in supporting me in completing this thesis. They have viewed it as a positive and important project that needed to be completed despite the barriers and obstacles to be overcome.

I will be ever humbled by their support and encouragement and I am constantly inspired by their courage and resilience.

Acknowledgements

Table of Contents

List of Figures

List of Tables

Chapter One: Introduction

The theoretical framework for investigation is discussed in this chapter and specific detail is provided about the New Zealand secondary school and qualifications context. The justification for the research in terms of global imperatives in education and the way these play out within the New Zealand context is described. The justification for the selection of the literature that has informed the study is explained and the research questions that arise from the literature review are attended to through the selection of relevant methodologies. The rationale for the measurement strategies used are established and informed by the existing literature.

An overview of the nine thesis chapters is provided at the end of this chapter.

Theoretical framework for this project

From the start of the new millennium, there has been an unprecedented public focus on student outcomes and increased school accountability, particularly in economically developed countries. The need for a competitive edge in the global economy, the changing skill requirements of a high technology work place, the demand for citizens who can cope with the problems of environmental change and degradation and population diversity are all strong drivers for increased performance of the public school system (2008, p. 3). League tables that rank schools, websites where students can comment on their teachers' performance ("nz.ratemyteachers.com," 2010) and publicly available evaluation reports carried out by government agencies (ERO, 2011a) have all put school leaders and teachers under increased pressure to ensure that their school is performing at or above expectation. Not only do teachers and leaders of schools compare their school results to each other, policy makers and researchers are comparing country level 'systems' and their ability to ensure students are performing at expectation, particularly in reading, writing and mathematics. The central challenge is to strengthen and maintain public confidence in a state-run education system for all so that people continue to send their children to public schools and provide their tax money for those schools to improve and develop (Levin, 2008). A coherent policy framework for school reform is critical in supporting ongoing improvement of the education system (Day et al., 2009) but the extra impetus will require horizontal and lateral ways of working where instructional leadership practices are enhanced at all levels of the system. As teacher effectiveness is such an important factor in determining student achievement and attainment, the extent that teacher leadership can act as a catalyst for

improving practice, changing beliefs and behaviours and positively impact on effectiveness (Harris & Muijs, 2005) in the classroom is central to this study. In the secondary school context it is the middle leaders that often have the strongest connection to the teaching in their department. The structural context of high schools means that these middle leaders were a relevant focus for a project aimed at identifying predictors of improved student academic outcomes.

During the last 2 decades a considerable body of evidence has accumulated to show that although the ability and socioeconomic background of students are a major determinant of achievement, schools can make a difference to students' levels of progress. School effectiveness research focuses on the achievement of all students and their progress over time. An effective school is one that adds value to student outcomes in comparison with other schools serving similar intakes (Sammons, Hillman, & Mortimore, 1994). At the secondary school level, much of the research to date has focussed on the organisation of effective secondary schools (Lee, Bryk, & Smith, 1993) and to some extent effective teaching within those schools. Much of this research highlights the fact that major sources of inequity lie within schools as well as between them. Decisions made within schools have been found to create substantial variability in teachers' conditions of work and students' opportunities to learn, even within the same school (Lee et al., 1993). Sammons, Hillman and Mortimore (1994) argue that effective schools are learning organisations, with teachers and leaders continuing to be learners, keeping up to date with their subjects and with advances in understanding about effective practice. The indirect empowerment of teachers by leaders when they create the conditions or provide the resources that directly impact on positive student outcomes is particularly relevant in the context of a secondary school department (Robinson, Hohepa, & Lloyd, 2009). Lee, Bryk and Smith (1993) state that academic organisation in the high school system is a critical influence on students' academic outcomes and that effective school personnel understand the relationship between social relations, teaching and instruction, and academic learning. When department staff are instructionally led and focussed on high standards of student achievement, learners will reap the rewards.

The theoretical framework for this research draws on the existing literature on within-school variation in student academic outcomes (Fitz-Gibbon, 1992; Goldstein, 2001; Tymms & Fitz-Gibbon, 2001), and the research on middle leadership in secondary schools (Dinham, 2007; Harris, 1999; Sammons, Thomas, & Mortimore, 1997; Siskin, 1994). By combining the evidence from these two bodies of knowledge, the theory for this thesis was developed and the methodology and procedures for the project were

established. Extensive literature and models that detail the theory and practice of successful school leadership have provided a strong foundation for this study as many of the features of effective principal leadership can be born out at middle leader level and vice versa. Accordingly, the methodological decisions evident in some of the empirical school leadership studies have been useful in shaping the design of this study.

The advantage of the 'department' as the unit of analysis to test for within-school variance was suggested by Lee et al. (1993), in a large review of studies on the organisation of effective secondary schools in the early 1990s. These authors propose that the bureaucratisation of secondary schools has led to the department as a device to efficiently organise teachers' work within the complexity of a modern secondary school. Role cultures in which responsibilities are defined around a set of positional roles within formal hierarchical structures (Feist, 2008; Thrupp & Willmott, 2003) are particularly dominant in large secondary schools in the urban New Zealand context. When the department can focus development on teachers' pedagogical skills and strengthen their commitment to teaching, staff are likely to behave more collegially. Evidence from the literature suggests that creating shared work practices and building a professional culture is a critical leadership role (Feist, 2008). Secondary school teachers describe themselves as subject matter specialists, with their ties primarily to their department as opposed to the school. There is also evidence that teachers in different departments hold substantially different views about the organisation of the secondary school within which they teach and that there are significant differences among departments in the same school in terms of teachers' beliefs, work commitments and social relations (Rowan, Raudenbush, & Kang, 1991).

The research studies that were of greatest interest and relevance to this project were those that came from the schooling effectiveness and improvement field because they were mostly concerned with measuring student outcomes and identifying the school level effects including departmental leadership. Establishing the links between the actual leadership behaviours and professional practices of middle leaders and the outcomes for students has similar constraints as those faced by researchers identifying the link between student outcomes and principal leadership. Robinson et al. (2009, p. 475) state that qualitative research supports the general belief that school leaders make a substantial difference to student outcomes. Establishing the quantitative links between school leadership and student outcomes has been more difficult. The findings of quantitative researchers indicate that school leaders have small and indirect effects on student outcomes, because those effects are mediated by teachers (Hallinger &

Heck, 1998) . It is well established in the literature that it is the quality of teaching that has the greatest impact on student learning (Alton-Lee 2003), so determining what middle leaders do that supports and promotes quality teaching and learning within secondary academic departments is central to investigating the impact of middle leaders on student academic outcomes.

The New Zealand senior secondary school context and the qualifications system

New Zealand secondary schools are structured like many others in western countries where core curriculum subjects are taught by a team of teachers organised into departments. In New Zealand secondary schools the eight essential curriculum areas of mathematics, English, languages, social sciences, the arts, science, health and physical education and technology are often used for organisational purposes. Secondary schools are generally divided into departments that are related to curriculum, and the staff within them are usually lead by a head of department and other designated leaders with various lines of accountability, such as Head of Biology or Head of Junior English. The term 'middle leader' in this project has been defined as those teachers with a designated leadership role in an English, mathematics or science department in a secondary school.

The compulsory national qualifications system for state-funded secondary schools in New Zealand is called the National Certificate of Educational Achievement (NCEA) and is characterised by its close assessment of the taught curriculum (Bendikson, Hattie, & Robinson, 2011). It is a standards-based assessment of student learning designed to provide three levels of national qualification for students studying subjects based on the New Zealand curriculum that are taught nationally at secondary level. The Ministry of Education is the government department responsible for the development of achievement standards (NZQA, 2011d). The qualification is recognised by employers and used for selection by universities and polytechnics, both in New Zealand and overseas (NZQA, 2011a).There are various features of the NCEA that are innovative and break with long-standing New Zealand tradition (Meyer, McClure, Walkey, McKenzie, & Weir, 2006). Unlike the norm-referenced and scaled examinations of the previous system, the standards-based assessments do not focus on comparisons between students but on how well they perform in relation to the standard. This system also offers opportunities for students to be assessed internally, by their teacher with interdepartment and school moderation, as well as externally in public examinations held at the end of the school year. NZQA has a formal quality assurance process to

ensure that the assessment of each standard is fair across all students, regardless of the school they attend (NZQA, 2011a).

There are three levels of NCEA certification depending on the difficulty of the standards achieved. At each level students must achieve a certain number of credits to gain an NCEA certificate (NZQA, 2011a). Eighty credits are required at any level and must include literacy and numeracy standards. In all secondary schools in New Zealand it is compulsory for students to take a course of study in English and mathematics until the end of Year 11, in order to gain the minimum literacy and numeracy requirements to gain the Level 1 certificate and get on the path for university entrance. At least one science subject is highly recommended as compulsory for students at this level. This means that performance across departments at Year 11 can be measured knowing that the greatest majority of students will be studying all three subjects. Schools and therefore departments can offer students opportunities to choose to study toward gaining achievement and/or unit standards at this level. Unit standards have a more vocational orientation at this level of the qualifications framework and achievement standards are closely aligned with national New Zealand curriculum levels. For the purposes of this study, only the results from achievement standards have been used for the analyses as these assessments are externally moderated and have an external examination component; therefore, the results can be compared. The achievement standards also provide the higher level of qualification required to enter university.

Standards are organised into levels of increasing difficulty. The standards assessed in schools are usually at Levels 1, 2 and 3. Most Year 11 students (15-year-olds) start at Level 1, and progress to Level 2 in Year 12 and Level 3 in Year 13 (NZQA, 2009). NCEAs are awarded according to the level of the standards achieved. For example, if a certain number of Level 1 standards are gained a Level 1 NCEA certificate is awarded. Achievement standards have three different grade descriptors: Achieved (A), Merit (M) and Excellence (E) (NZQA, 2011a). A typical Year 11 mathematics or English course would be made up of assessments towards five individual standards that cover the curriculum. Approximately half of these would be internally assessed and half externally assessed by public examination. This study uses data from the most commonly assessed internal and external achievement standards so that the results are comparable between schools. For example, a typical mathematics assessment programme, in a school cohort at Year 11 or NCEA Level 1, may look like that presented in Table 1.1. In New Zealand the external standards are assessed by public examination in November of every year.

Table 1.1

Typical Level 1, Year 11 Maths Course for 15-Year-Olds in 2008

Standard no.	Assessment	Name of standard	Credits earned
90147	External	Use straightforward algebraic methods and solve equations	4
90148	External	Sketch and interpret graphs	3
90149	Internal	Solve problems involving measurement of everyday objects	3
90150	Internal	Use geometric techniques to produce a pattern or object	2
90151	External	Solve straightforward number problems in context	3
90152	External	Solve right-angled triangle problems	2
90153	External	Use geometric reasoning to solve problems	2

The rationale for selecting the English, mathematics and science as subjects for this study was because the largest number of 15-year-old students in New Zealand would be studying these subjects because they are generally[1] compulsory at Level 1 for Year 11, (15-year-old students). This enabled individual student data and performance in each of the subjects to be measured because all students were completing courses in English, mathematics and science. In addition, these departments are likely to be similar in size, resource and status within a secondary school, with a range of teachers with Year 11 classes. The data were generally reliable because of the national accountability systems, which are publicly available and comparable for internally and externally assessed NCEA standards in these subjects.

Measuring socioeconomic status of students for comparative purposes

Every school in New Zealand — primary and secondary — is accorded a number from 1 through to 10 that denotes the average socioeconomic and ethnicity status of the students who attend the school. In England, researchers often use such indicators as the number of students who receive free school lunches but New Zealand researchers are able to make use of the decile system to make comparisons of student achievement in relation to their general socioeconomic status by using this descriptor. New Zealand schools receive funding from the government essentially on a per pupil basis. The decile system arose from a review of resourcing by the Ministry of Education, in order to determine a transparent, simple-to-calculate and valid formula for differentially funding schools (Hattie, 2002). Supplementary funding is distributed to schools based on decile rankings from 1–10. These rankings were introduced in 1995 (Ladd & Fiske, 2003). Schools with a decile 1 ranking serve the highest proportions of minority and economically disadvantaged students while decile 10 schools serve the more affluent

---

[1] Some New Zealand schools may create 'alternative' courses for a small minority of students who are unlikely to be able to achieve in these subjects, particularly science.

students and are often situated in prosperous urban suburbs. Hence, the New Zealand government funds lower decile schools at a higher per pupil rate than it does higher decile schools.

The New Zealand public has full access to school decile information and are able to see how each secondary school performs in relation to others that are similar by comparing percentage pass rates of students in the same decile band. Fiske and Ladd (2001) have documented the strong perception of parents that deciles can be used as the best short-hand index of school quality. In New Zealand, the decile of the school, rather than the performance of the students within the school, is the litmus test of acceptability (Hattie, 2002). The decile system enables all stakeholders to compare the performance of students to peers in schools similar to their own (Meagher-Lundberg, 2000). This perception derives from a belief that achievement scores, such as literacy and numeracy, are often related to critical characteristics of students and families beyond the control of teachers, and by expectations that 'similar' schools should add similar educational value (Hattie, 2002). The use of decile for comparative purposes is a central feature of the quantitative data analysis of this project. It has enabled schools to be compared in terms of 'like with like' and has helped determine which schools are 'adding value' given the demographic of the population intake of each of the schools.

Evidence of the within-school variance problem in New Zealand

Variation in the achievements of pupils within schools, rather than variation between schools has come into focus internationally due to the accumulation of evidence that teacher effects on pupils are greater than school effects (Reynolds, 2004). Identifying within-school variation within the New Zealand secondary school context was a critical first step in this project. The New Zealand system can be characterised at secondary level as one of high performance and low equity (McNaughton, Robinson, & Timperley, 2011), with Māori and Pasifika students overrepresented in the statistics for students performing below expectations. The results for New Zealand students are carefully compared to and analysed against those in other Organisation for Economic Co-operation and Development (OECD) countries. Achievement results for students are publicly reported by age, gender and ethnicity for both nationally and internationally organised tests such as The Programme for International Student Assessment (PISA). The analyses provided by the PISA testing programme identifies the across- and within-school variance in student results and compares the measures across OECD and non-OECD countries. This test was a particularly relevant comparison for this study because it assesses the ability of 15-year-olds in mathematics, reading and science, while disaggregating results for student socioeconomic and cultural background.

An examination of New Zealand student performance in PISA over the last 10 years helps to establish the extent of within-school variation in New Zealand secondary schools. The timing of the 2009 PISA results and subjects tested (reading, mathematics and science) provided an ideal opportunity to test and compare the findings for this project because the tests have been performed over the same time period as the NCEA results collected for this study. The PISA studies reveal that New Zealand schools have demonstrated a wide range of student outcomes since 2001 from students who perform at the very top of their academic age level to those at the very bottom as compared with other OECD countries. In 2009, New Zealand students performed statistically significantly above the OECD average on the overall reading, mathematics and science scale (OECD, 2010b).[2] However, there was also a large variation of student performance in reading within schools. The variation of student performance within schools is 75.8%, which is 14.4 % above the OECD average of 61.4 %. In contrast, the variance across schools was 10% less than the OECD average of 42% (OECD, 2010b).

Key findings from the most recent report (Telford & May, 2010) on the achievement of New Zealand students in PISA indicate that in reading New Zealand 15-year-old students' overall reading performance was substantially higher than the average for the 34[3] OECD countries. Of the 65 countries or economies participating in PISA 2009, only two OECD countries (Korea and Finland) and two non-OECD partner economies (Shanghai-China and Hong Kong-China) performed better[4] in reading literacy than New Zealand. Canada, Australia, Japan and Singapore's performance was similar and the other 56 countries performed at a significantly lower level. The 2009 PISA results show the diverse abilities of our students as close to one in six of New Zealand students were top-performing readers but the proportion of New Zealand students at the lowest level of reading proficiency was 4%, similar to that of Australia (4%) and Singapore (3%).

In PISA 2009, the reading proficiency levels were extended to introduce a higher level, Level 6 and a lower level, Level 1b. Level 2 has been established as the baseline level at which students can begin to have enough literacy to access the curriculum and participate actively in life situations. In New Zealand, 14% of students did not achieve Level 2 reading literacy in 2009 as compared to the highest ranked countries, Shanghai-China (4%), Finland (8%) and Hong Kong-China (8%).The 2009 PISA results show that New Zealand has a range of students from across ethnicities such as Asian, Māori, Pākehā/European and Pasifika students who performed at the highest level of

---

[2] Source: OECD, PISA 2009 Database. New Zealand ranked 7th on the overall reading scale with 521, 12th on the mathematics scale with 519 and 7th on the science scale for 15-year-olds.
[3] Chile, Estonia, Israel and Slovenia are new OECD member countries.
[4] Terms such as 'better', 'larger', 'weaker' or 'smaller' are used when results are statistically significant at the 0.05 level.

reading literacy. While Pākehā/European and Asian students were more likely to be at the higher end, Māori and Pasifika students were overrepresented at the lower end of performance. One of the greatest concerns is the strong relationship between socioeconomic background, ethnicity and achievement (McNaughton et al., 2011). Across the 34 participating OECD countries, 14% of the variance in achievement is attributed to socioeconomic background of students, and in New Zealand the figure is 19% (OECD, 2010a). This is evidenced in the PISA statistics that reveal 35% of Pasifika and 24% of Māori students did not show reading proficiency above Level 1a. This compared to 15% Asian and 9% Pākehā/European (Telford & May, 2010).

In mathematical literacy, New Zealand students' overall performance was significantly higher than the average for the OECD countries. Five OECD countries and six non-OECD partner countries or economies performed better than New Zealand in mathematics, four OECD countries were similar and the other 49 countries had a significantly lower performance. Fifteen percent of New Zealand students performed at the lowest level in mathematics as compared to Shanghai-China (5%), Finland (8%) and Hong Kong-China (9%). In scientific literacy, New Zealand students' overall performance was substantially higher than the average for the OECD countries. One OECD country and three non-OECD partner countries or economies achieved a higher mean scientific literacy score than New Zealand. Six OECD countries were similar and the other 54 countries performed significantly lower. Thirteen percent of New Zealand's students did not show scientific proficiency at the baseline Level 2. New Zealand's 15-year-olds' mean mathematical and scientific literacy performance did not change between 2003 and 2009.

The 2009 PISA findings clearly show the performance differences within and between schools (OECD, 2010a). The report writers state that differences can be attributed to the organisation of schooling systems or differences in the quality of effectiveness of instruction that those schools provide. The PISA results in 2006 and 2009 show New Zealand as being well above the OECD average for within-school variance but well below for variation between schools. The 2009 results reveal the extent to which differences in performance between schools and among students within schools can be attributed to differences in socioeconomic background between and within schools (OECD, 2010a, p. 86). The results show that regardless of their own socioeconomic background, students attending schools in which the average socioeconomic background is advantageous tend to perform better than when they are enrolled in a school with a disadvantaged socioeconomic intake. The within-school differences in socioeconomic background across students are less pronounced in relation to

performance than the between-school differences. New Zealand does have one of the highest variations in reading performance explained by student's socioeconomic background **within** schools (10%) in the OECD. There is also a large variation of student performance in reading within schools. The total variance in student performance within schools is 75.8%, which is 14.4 % above the OECD average of 61.4 %. In contrast, the variance across school is 10% less than the OECD average of 42% (OECD, 2010b). The PISA results provide a valuable national picture of student achievement in reading, mathematics and science for 15-year-olds and results that can be compared at NCEA Level 1.

The New Zealand school leaver statistics show that in 2009, Asian students had the highest proportion of school leavers attaining at least NCEA Level 1 (90.9%), which was 4% higher than Pākehā/Europeans (87.1%), and Pasifika (76.6%) and Māori (65.9%) had the lowest rates (Ministry of Education New Zealand, 2010a). At secondary level there is a large variation in the percentage of students from different socioeconomic groups gaining qualifications at Level 2 of the national qualifications framework, designed for 16-year-olds (Robinson et al., 2009). In 2009, 86.9% of students from schools in the highest deciles (serving socioeconomically advantaged students) left school with at least a Level 2 qualification. This compared to 51.2% of the school leavers of decile 1 and 2 schools. The results show that there is also a large variation in the percentage of school leavers attaining at least an NCEA Level 2 qualification amongst schools within each decile (Ministry of Education New Zealand, 2010b). While New Zealand has between-school differences in student performance, it has far greater within-school disparities, which suggests that the relevance of effective pedagogical leadership strategies for all New Zealand schools will be critical in addressing this problem (Robinson et al., 2009).

Measuring across- and within-school variance

Having established that there is evidence of disparity in outcomes for students in the senior secondary school in New Zealand, the next step was to determine where the variance was occurring and how it could be investigated and measured. The arguments for the reasons for disparity will be ongoing but the evidence shows that within-school variance in secondary schools is a continuing cause for concern. Overall school results can mask underperformance in some areas. Research concerning school improvement underlines the importance of focusing change efforts at different levels within the organisation (Fullan, 1999). Silins and Mulford (2002) claim that middle managers in secondary schools have an essential role to play in school effectiveness and improvement and the way that the leadership roles interact within the school is also

crucial. Some researchers argue that academic outcome measures are best considered at department level (Fitz-Gibbon, 1991; Smith & Tomlinson, 1989) where educationally specific decisions are made and a greater proportion of variation is explained than at school level. The largest study of differential school effectiveness in the United Kingdom highlighted the importance of differences between departments in explaining differences in school performance (Sammons et al., 1997). It provided evidence that both schools and departments are differentially effective with pupils of different abilities and of different social and ethnic backgrounds. This study also suggests a need to broaden the study of school leadership to include middle leaders. Glover, Gleeson, Gough and Johnson (1998) have argued that the distinction between middle and senior management remains blurred and leadership functions are still not adequately delineated or defined. Given the important influence of middle leaders on teacher pedagogical and content knowledge, potentially these education professionals are important players in the schooling improvement agenda.

Most school effectiveness studies show that 80% or more of student achievement can be explained by student background rather than schools (Teddlie & Reynolds, 2000). School effectiveness researchers believe that even with only 20% of achievement accounted for by schools they add value despite the strong influence of family background on children's achievement (Sammons et al., 1998; Silins & Mulford, 2002; Thomas, Sammons, & Mortimore, 1995). There is a considerable body of evidence that the quality of teachers in classrooms rather than the school and how it is organised or led makes a difference to student academic outcomes (Silins & Mulford, 2002). However, the choice of social or academic measures used within and across schools and the complexities of how to show evidence of impact on student outcomes is an ongoing debate within the community of schooling effectiveness researchers. Stoll and Fink (1996) have argued that schools should measure what they value. Sammons, Thomas and Mortimore (1997) assert that while academic outcomes are not the only ones that should be valued, they are of critical importance in OECD countries where "high stakes" testing is institutionalised via a public examination system. They also believe that any comparison of results of individual schools should be comparing "like with like," taking into account student intake because it is the value-added improvement and the schools' ability to promote progress while students are in the school that needs to be measured, not what inherent advantages they had before they attended the school. Sammons (2009) argues that the focus on high stakes testing and exam results to measure and compare performance of whole schools should shift to a focus on differential effectiveness within a school. In this study, consideration of the internal variation in academic results for subgroups of students, particularly by department,

gives an accurate measure of the subjects or curriculum areas where students are achieving.

Schooling effectiveness researchers in Britain and the Netherlands have produced some large scale studies using student examination data at secondary school department or subject level that can be used to compare within-school variance or difference in department and subject performance (Fitz-Gibbon, 1992; Hofman, Hofman, & Guldemond, 2001; Luyten, 1994; Sammons et al., 1997; Thomas et al., 1995; Tymms & Fitz-Gibbon, 1990). Luyten (1994) argues that school effectiveness is actually an artefact and effective schools are simply those with a high percentage of effective teachers or departments. Luyten's analysis of variance using large data sets to calculate the mean examination results for Dutch schools over three years shows that schools may be effective with respect to a certain subject one year but much less effective the next year. This produces a within-school, across-year variation. He asserts that effectiveness cannot be assumed to be a stable school characteristic and that the same school might produce diverging effects because within a school both more and less effective teachers and departments will be found. He questions the extent to which school effects can be considered stable or variable across subjects and across years.

When student examination data are used to determine school effects there are a number of factors that need to be considered in the design of a project. There is now substantial academic agreement as to the most appropriate methods of estimating school effects and the data required for valid comparisons to be made (Goldstein et al., 1993; Scheerens, 1992). Scheerens (1992) has developed criteria cited as necessary for an adequate study of school effectiveness. Some of these *school* effectiveness measures can be applied to the measurement of *department* effectiveness within schools. The secondary school department can be used as the unit of analysis that allows for data analyses of sufficient discriminative power and it is possible to use longitudinal data. Adequate techniques for data analysis (e.g., multilevel models) are possible when employing a range of qualitative and/or quantitative techniques. In education, there has been much debate about the so-called 'unit of analysis' problem. Before multilevel modelling became well developed as a research tool, the problems of ignoring hierarchical structures were reasonably well understood, but they were difficult to solve because powerful general purpose tools were unavailable. Special purpose software has allowed the analysis of generic data. In a school setting, researchers can explore the extent to which differences in average examination results between schools are accountable for by factors such as organisational practice or possibly in terms of other characteristics of the students. The software makes it possible to study the extent

to which schools differ for different kinds of students, for example, to see whether the variation between schools is greater for initially high-scoring students than for initially low-scoring students (Goldstein et al., 1993) and whether some factors are better at accounting for or 'explaining' the variation for the former students than for the latter. In *Forging Links* (Sammons et al., 1997), multilevel models were used to establish whether the items made a statistically significant contribution to the explanation of school and departmental variation in the General Certificate of Secondary Education (GCSE) performance over three years, after controlling for differences between schools and their student intake.

Although the academic results for departments in schools are able to be disaggregated through the use of these multilevel models and various electronic student data management systems that exist within schools, most schools choose to publish their whole-school results, thereby 'masking' the within-school variation. The aim of the initial investigation for this project was to examine the publicly available student academic qualifications data for 15-year-olds — by subject — in order to investigate the extent of the differentiation within schools. Once the extent of that variation was established the middle leadership practices of the successful and not so successful departments were examined in order to establish the link between the middle leadership practices and the academic outcomes for students. It is the link between middle leadership practice and student outcomes that needs to be measured. Identifying the deliberate acts of teaching, leadership and organisational policies of middle leaders that promote student progress is one way to increase leadership capability within the heart of the secondary system.

Significance of the research

The impact of middle leadership on secondary student outcomes is an under researched area, particularly given the institution of the department in the secondary school setting. In 2007, Bennet et al. (2007) published a review of the empirical research in the English language that reported on the work of middle leaders in secondary schools mostly in the United Kingdom. The research team reviewed 101 studies completed between 1988 and 2005. High school departments were first identified by Siskin (1991) in the early 1990s as sites of importance for educational researchers when she conducted some in-depth case studies in high schools in California. Subject departments had become visible in ways they had not been before and the result had been a growing body of empirical evidence demonstrating the salience of subjects and revealing the unexpected variations in departments within the same school (Siskin & Little, 1995). Schooling improvement and effectiveness

researchers particularly in England and the Netherlands conducted a range of important larger longitudinal studies (Harris, Jamieson, & Russ, 1995; Hofman et al., 2001; Sammons et al., 1997) that involved large data sets in the 1990s to investigate the impact the middle leaders had on the quality of teaching and student engagement and academic development of students. More recently, there has also been considerable research and policy interest in the identification and causes of within-school variation with reports produced by The National College of School Leadership in England (NCSL, 2006). Qualitative research has also been conducted in Australia by Dinham (2007), which builds on previous theory in identifying characteristics of effective middle leaders in departments where students were achieving academically well.

The contribution of this project to the field includes a method for identifying the extent and depth of within-school variation in academic results for 15-year-olds in a large sample of urban Auckland secondary schools. The methods used and the results could be used for comparative purposes for other similar studies investigating within-school variation. A method for converting grades from a standards-based assessment system, developed by The University of Auckland (Shulruf, Hattie, & Tumen, 2008) into a quantifiable data set was employed and the systems and procedures used could be utilised by other New Zealand and international researchers for similar projects. The development of the middle leadership questionnaire that investigates key middle leadership behaviours at department level could be used and adapted in a range of research and professional development contexts both within New Zealand and internationally. Critically, the project has identified some of the specific middle leadership behaviours that predict positive and negative student academic outcomes while controlling for student socioeconomic differences. The study has also revealed at which levels of the education system middle leadership makes a difference and the subject differences that occur.

Overview of the thesis

The thesis is structured into nine chapters. Chapter One provides an introduction to the research through a brief examination of the theoretical framework for the project from both an international and New Zealand perspective. To provide the reader with the necessary background information, the current study is situated within the context of the wider educational reform agenda, particularly within the schooling improvement research theory. The significance of the research and its contribution to the existing theory in the area of within-school variation and the effectiveness of middle leadership and the relationships to student academic outcomes is summarised.

Chapter Two reviews the bodies of literature relevant to the current research. It begins with a broad description of the recent empirical evidence of effective leadership practice in school settings that impact on student outcomes. The ideas are developed to consider the literature on the context of secondary school departments as an important and unique site for research and inquiry. The literature on within-school variation is examined and the research studies that have shown a link between middle leadership effectiveness and secondary student academic outcomes are described. This chapter concludes with a matrix to categorise characteristics and practices of effective middle leadership and shows the relationship between the literature and the research questions for the study.

Chapter Three outlines the methodological approach to the whole project with a more detailed description of the rationale and design for Phase1. The selection of the 41 sample secondary schools and methods for collecting quantitative data were explained in order to test the extent of within-school variation in the sample schools. Chapter Four outlines the results for the Phase 1 study and describes the overall performance of 15-year-olds in each school in the sample. The within-school variation, the relationship to school decile and therefore general socioeconomic status of the students and analysis of the differences between subjects is discussed.

Chapter Five describes the development of the middle leadership questionnaire, how the items were developed, the trial of the questionnaire and the fieldwork. The ethical considerations and the selection of the 10 schools are also discussed. A description of the demographic data for the schools is provided. Chapter Six describes the results from Phase 2 of the project, which measures middle leadership effectiveness in 30 departments (three departments within 10 schools). The statistical analyses related to the questionnaire are provided, including the factor analysis and a comparison of means by school, department and across subjects.

Chapter Seven describes the third phase of the project, which is a statistical analysis of the academic results across three year levels and across three years for each of the 30 departments, for whom middle leadership effectiveness was measured in Phase 2. The results describe the whole-school performance and the extent of the within-school variation. High- and low-performing departments are identified in comparison to national norms and subject comparisons are made. Patterns for achievement in subjects and the interactions between school and department effects are discussed. Measures for determining department performance are explained and summarised results for the sample are produced for comparative purposes. Chapter Eight describes the statistical tests performed to analyse the student academic outcomes and middle leadership

effectiveness. The emerging patterns and relationships are described and leadership attributes that predict student outcomes are discussed.

Chapter Nine identifies the major findings from the project and answers to the research questions are addressed. The limitations of the project are identified and the implications for the profession in terms of policy and professional development are discussed.

The chapter that follows (Chapter Two) includes a literature review of the key school and middle leadership theory that has informed this project. An analysis of the within-school variation research is also surveyed as it is a critical theory base that has informed this research.

Chapter Two: Literature Review

Introduction

This chapter will set the broad context of the study and state the parameters of the theoretical investigation. There are a number of interrelated theoretical areas that have been drawn on to inform this review. The studies of significant interest originate within the context of schooling improvement research completed since the 1980s and have included the areas of educational leadership that impact on student outcomes, school effects, within- and across-school variation and specific research on the practice of middle leaders in secondary schools. The focus for the review is in the secondary school context and much of the research draws upon evidence from schools in the United Kingdom, the Netherlands, Australia and the United States of America. Research studies that have shown particular evidence of differentiated student academic outcomes as a result of an investigation and/or intervention were of particular interest. These types of studies are scarce, particularly in relation to middle leadership in secondary schools, so drawing on the theoretical contexts of within-school variation and combining that work with the evidence on the importance of middle leadership has been critical in order to synthesise the relevant information and knowledge required to develop the theoretical basis for this project. It is the convergence or bringing together of the literature in these two subfields that has supported the development of the methodology, tools and instruments that have produced the findings of this research project.

The scope and rationale for the review included three main areas of literature, drawing on research that has occurred from the mid-1980s to the present. The first general exploration was in the field of schooling improvement looking at the impact of school leaders on academic outcomes for students and the different types of leadership practices that empirical research has shown have a direct or indirect impact on the quality of teaching and learning in a school. The second area was a review of the literature on middle leadership in secondary schools, which has been scarce but consistent since the mid-1980s. Further discussion is included on the third and interrelated area of work done in the last 20 years on within-school variation and school effects on student achievement. A particular focus on the work completed in state secondary school contexts using the analysis of public examination data as evidence of academic student outcomes in OECD countries was relevant to this study.

Evidence about the impact of effective leadership practices that impact on student outcomes in the school setting

In order to understand the importance of the impact of middle leaders, it is important to consider the literature on effective leadership practices at school level that impacts positively on student outcomes, as it applies and is relevant to the role of middle leader. The school leadership research that directly links and impacts on middle leadership in secondary schools is discussed in the first part of this review. These theories are important in setting the context for a further discussion on effective middle leadership in secondary schools.

The schooling improvement research that describes the reasons why schools are differentially effective and how leadership can contribute to outcomes for students with both positive and negative consequences is fundamental in situating the research on middle leadership within the field. The effective schools research has been a driving force behind efforts in many OECD countries to improve outcomes for students in public education, suggesting that through strategic school organisation and strong principal leadership improved student achievement is possible, despite environmental constraints (Heck, 1992). Leithwood, Harris and Hopkins (2008) have completed a review of literature summarising the main findings from studies around what the authors call "strong claims" about successful school leadership. They make a number of claims that find support to varying degrees in the empirical research evidence. The two claims that attracted the largest amount of evidence are that school leadership is second only to classroom teaching as an influence on pupil learning and almost all successful leaders draw on the same repertoire of basic leadership practices. These include: building vision and setting direction; understanding and developing people; redesigning the organisation; and managing the teaching and learning programme. Leithwood et al. (2008) state that school leaders improve teaching and learning indirectly and most powerfully through their influence on staff motivation, commitment and working conditions. This is played out in the building of staff capacity with an ongoing schoolwide focus on the development of curriculum content and pedagogical knowledge. These responsibilities are an area where principals are likely to distribute responsibility to those middle leaders with the specialised curriculum and pedagogical expertise, required to drive and monitor the quality of teaching and learning within their department.

One of the most important leadership constructs that has increasingly directly impacted on the positioning and responsibilities of middle leaders in secondary schools in the last 20 years is that of *distributed* leadership. In contrast to traditional notions of leadership

premised upon an individual managing hierarchical systems and structures, distributed leadership is characterised as a form of collective leadership in which colleagues develop expertise by working together (Harris, 2004). Distributed leadership theory focuses on how leadership practice is distributed among formal and informal leaders in a school setting. It is a form of collective agency incorporating the activities of many individuals in which group members mobilise and guide other teachers and themselves to focus on continued improvement and upskilling for the purpose of improved student achievement. Brown, Rutherford and Boyle (2000) assert that school leadership has a greater influence on schools and students when it is widely distributed and that some patterns of distribution are more effective than others. In the Leithwood et al. (2008) analysis, the schools with the highest levels of student achievement attributed this to relatively high levels of influence from all sources of leadership. Leithwood et al. (2008) claim that there is no loss of power and influence on the part of the principal when the power and influence of many others in the school increases. This notion of distributed leadership would suggest that well-supported middle leaders are likely to have a positive impact on teaching and learning, although Leithwood et al. (2008) stress that there is limited verifiable evidence available to test this assertion. Most restructuring and educational reform initiatives include the decentralisation of decision-making and the empowerment of teachers (Murphy, 1994). The notion of distributed leadership or shared power and decision-making among managers at all levels of the school is widely promoted as a factor contributing to school effectiveness and school improvement (M. Brown et al., 2000). If we accept that distributed leadership is critical to school success, we need to understand more about how this is played out in effective schools and departments. The way that leadership practices are distributed to and by middle leaders in the secondary school context is fundamental in developing our understanding of which leadership practices make a difference for positive student outcomes.

The distributed leadership research identifies many specific characteristics and behaviours of school leaders relevant to teachers and leaders throughout the school system. Spillane, Halverson and Diamond (2001) and Spillane and Seashore (2002) propose that educational leadership involves the identification, acquisition, allocation, coordination and use of the social, material and cultural resources necessary to establish the conditions for the possibility of teaching and learning. They describe distributed leadership as a form of collective agency incorporating the activities of many individuals in a school who work at mobilising and guiding other teachers in the process of instructional change (Spillane et al., 2001; Spillane & Seashore, 2002). The argument for school principals to practise instructional leadership and use distributed leadership patterns to enhance school wide leadership and effectiveness is compelling

when the evidence of these practices is linked to improved student academic outcomes. In a recent report entitled *The Impact of School Leadership on Student Outcomes* commissioned by the National College for School Leadership in England, the researchers found that leaders with values and aspirations focussed on improved student outcomes, associated with wise, timely and contextually focussed strategic decisions were the most successful in terms of improving student outcomes in a relatively short time frame (Day et al., 2009). The researchers go on to state that there are positive associations between the increased distribution of leadership roles and responsibilities and the continuing improvement of pupil outcomes (Day et al., 2009, p. 4). In a review of successful school improvement efforts by Glickman, Gordon and Ross-Gordon (2001), the researchers identified varied sources of leadership, including distributed leadership as an important contributor to the positive shifts in the schools studied. Silins and Mulford (2002) have shown that student outcomes are more likely to improve where leadership sources are distributed throughout the school community and where teachers are empowered in areas where they have expertise and believe the task or responsibility is important. The specific distributed leadership practices enacted at all levels of the school that are found to impact positively on student academic outcomes have an inevitable impact on those who are carrying out the middle leadership function in a secondary school and are therefore relevant to this study.

There is now considerable evidence that leadership practices need to be focussed on improving the teaching and impact on learning opportunities of students to improve their social and academic outcomes. Robinson (2009) claims that whether leadership is distributed or not, the impact of pedagogical leadership is nearly four times that of transformational leadership. Pedagogical leadership involves practices such as leaders establishing clear educational goals, planning the curriculum and evaluating teaching as opposed to being exclusively focussed on developing vision and inspiration in conjunction with relationships. Robinson's argument is that leaders need to employ theories of transformation and distribution within their practice in a blended and synthesised way that is fit for purpose and context specific. When school leaders work effectively with middle leaders they are more likely to influence classroom practice in a positive way. In contrast to traditional notions of leadership based on an individual managing hierarchical systems and structures, principals or school leaders that employ distributed or collective leadership practices will influence the way teachers develop expertise by working together (Harris, 2004) rather like an effective department. The interrelationship between effective whole-school and middle leadership practices within a secondary context is an important aspect of this research project and is discussed within the results chapters of this thesis.

The Leithwood et al. (2008) findings can be compared to Robinson's *Iterative Best Evidence Synthesis* (2009) where she asserts that if leaders properly identify and focus on improving school and classroom practices then they are most likely to have a positive impact on student achievement. Table 2.1 describes these dimensions, their meaning and the mean effect size and the standard error calculated from a meta-analysis of 12 studies that describe the educational significance in terms of student outcomes.

Table 2.1

Leadership Dimensions With Effect-Size Estimates for Student Academic Outcomes

Dimension	Meaning of dimension	Effect-size estimate
1. Establishing goals and expectations	Includes the setting, communicating and monitoring of learning goals, standards and expectations, and the involvement of staff and others in the process so that there is clarity and consensus about goals.	ES = 0.42 (0.07) 49 effect sizes from 7 studies
2. Strategic resourcing	Involves aligning resource selection and allocation to priority teaching goals. Includes provision of appropriate expertise through staff recruitment.	ES = 0.31 (0.10) 11 effect sizes from 7 studies
3. Planning coordinating and evaluating teaching and curriculum	Direct involvement in the support and evaluation of teaching through regular classroom visits and the provision of formative and summative feedback to teachers. Direct oversight of curriculum through schoolwide coordination across classes and year levels and alignment to school goals.	ES = 0.42 (0.6) 80 effect sizes from 9 studies
4. Promoting and participating in teacher learning and development	Leadership that not only promotes, but directly participates with teachers in formal or informal professional learning.	ES = 0.84 (0.14) 17 effect sizes from 6 studies
5. Ensuring an orderly and supportive environment	Protecting time for teaching and learning by reducing external pressures and interruptions and establishing an orderly and supportive environment both inside and outside classrooms.	ES = 0.27 (0.9) 42 effect sizes from 8 studies

Source: (Robinson et al., 2009, p. 95)

The interpretation of effect sizes for this study is based on Hattie's benchmark where 0.20 is small, 0.50 is medium and 0.80 is large (Cohen, 1988). Hattie asserts that teachers should be seeking effects of greater than 0.4 over one school year with gains greater than 0.6 considered to be very good improvement. Based on Cohen's theories, Hattie asserts that these numbers provide guidelines to discuss aspects of teacher and leadership practice for gains and improvement in student performance (Hattie, 2009). The dimensions most closely associated with teaching and classroom practice appear to have the biggest effect size (Robinson et al., 2009), so consideration of the practices

that sit within these dimensions were critical when considering the middle leadership practices likely to impact most directly on outcomes for students.

Robinson (2009) does not list collegial relationships as one of her dimensions but says, "Relationship skills are embedded in every dimension….Effective leaders do not get the relationships right and then tackle the educational challenges — they incorporate both sets of constraints into their problem-solving" (p. 9).

The dimensions described by Robinson (2007) and the claims of Leithwood et al. (2008) provide an important framework in which to consider the competencies and leadership characteristics of middle leaders. The modelling of leadership behaviour, particularly when distributed throughout a school would suggest that these dimensions, behaviours or characteristics would be evident to teachers in successful schools at senior leadership and department or syndicate level. The way these attributes are exemplified in schools at middle leader level, how they are played out, or not, and which relate most closely to student academic outcomes are critical to this study.

The first dimension, *Establishing goals and expectations,* will only impact on student achievement if the goals and expectations are embedded in school and classroom routines and procedures (Robinson et al., 2009). In a study by Heck et al. (1992; Heck, Marcoulides, & Lang, 1991), there was some evidence that the degree of staff consensus about school goals was a significant discriminator between otherwise similar high- and low-performing schools. The logic follows that if middle leaders have some influence in the construction of goals and expectations, they are likely to be fully committed to the outcomes they are working towards. The extent to which middle managers set departmental goals and individual staff goals from the strategic goals of the whole school is an important part of this investigation. Latham and Locke (2006) have shown that if a person has adequate knowledge, skill and commitment there is a linear relationship between the degree of goal difficulty and subsequent job performance. They go as far as to say that goal-directed action is an essential aspect of human life. Without goal-directed action, people cannot attain the values that make their survival and happiness possible. In the recent British study carried out by Day et al. (2009), the data showed that one of the most powerful behaviours of effective leaders was the establishment of a clear sense of direction and purpose for the school. Establishing a clear vision that influenced the development of goals to guide their own and others' actions in the school was a central driver for leadership activity. The vision and goals need to be widely shared throughout the school, linked to student achievement and clearly understood so that they can be actively supported by all staff (Day et al., 2009; Robinson et al., 2009). The transformational concept of leadership

involves developing a vision for the school and maintaining its relevance for all concerned so that the school culture supports the school's vision and direction in order to nurture the capacity and commitment of staff (Duke & Leithwood, 1994). Goal setting with staff needs to focus on the specific and the difficult as opposed to being vague and generalised. Specific and difficult goals will lead to higher staff performance because they allow people to effectively judge their effectiveness. Setting goals appears to be effective because they prompt people to generate solutions to an impasse, implement new strategies and monitor their effectiveness. Witziers, Bosker and Kruger (2003) suggest in the reporting of their meta-analysis that the direction-setting role of the leaders had the greatest impact on student outcomes. However, the way that leaders are responsive to their particular school context, rather than dictated by it, is an important distinction. Leithwood et al. (2008) assert that successful leaders are sensitive to context but what is most important is how they apply context specific leadership practices such as building vision and setting directions with the correct timing, developing clear priorities and involving the right staff in crafting and revising the school's direction. In summary, the benefits of goal setting are that it affects our choices and gives direction to our pursuits, increases our effort and prolongs persistence. These behaviours can be born out in the school and department context as teachers and middle leaders search for appropriate strategies to attain goals, monitor and evaluate their progress, adjust behaviour, and celebrate the sense of accomplishment when targets and goals are achieved which in turn increases a sense of personal effectiveness.

The second dimension reported by Robinson et al. (2009) as being a significant leadership behaviour is *Resourcing strategically*. These behaviours are about strategically securing and aligning resources to pedagogical purposes and link very strongly to Harris (1999) and Siskin's (1994) findings that effective middle leaders manage the resources equitably to the mutual advantage of the staff and students in the department. Robinson et al. (2009) assert that more needs to be known about the specific knowledge and skills needed by school leadership to link the allocation of resources for recruitment to enhancing teaching and learning. While this dimension refers to both staffing and teaching resources, the most important resource that leaders manage is teachers, since the quality of teaching explains more of the variance in students' achievement than any other system variable (Alton-Lee & Rowe, 2007; Nye, Konstantopoulos, & Hedges, 2004).

Understanding the relevance and impact of this dimension at department level is critical given the likely direct influence on students' experience of teaching and learning at the classroom level.

McLaughlin and Talbert (2007) describe strategic resourcing as "finding time." They describe American high schools that reformed as those that found time for teacher collaboration by changing their work structure. The use of strategic resourcing is illustrated in this quote by the principal of a high school in the *Bay Area School Reform Collaborative* (BASRC) study who said:

> The main ingredient in school reform, in my opinion, is being able to buy teacher time to actually do the processing, to do the work. ….I don't think we spend enough on people…this allows us to have summer planning, to do the retreats, to buy release periods. The kinds of things that get people really engaged. (McLaughlin & Talbert, 2007, p. 163)

In the third dimension described by Robinson *Planning, coordinating and evaluating teaching and the curriculum* there were clear indicators that distinguished higher performing schools from their counterparts. Robinson (2007) refers to three interrelated subdimensions that need to be examined and tested further at secondary school department level. These include leadership by active oversight and coordination of the instructional programme, including the coordination of curriculum. An example is when leaders are involved in classroom observation and subsequent feedback that ensures staff systematically monitor student progress and that test results are used to improve classroom programmes. Day et al. (2009) support the view that effective leaders encourage teachers to go beyond their usual teaching models and try new or alternative approaches. Their research has shown that effective leaders encouraged teachers to be leaders in their own classrooms and to make informed decisions to extend their teaching approaches. Day et al. report that in effectively led schools classroom observation is used in a developmental way to provide support, guidance and advice about further improvements. Harris and Chapman (2002) support this view and describe effective principals as those that include a range of people in their decision-making, allocate important tasks to teachers and rotate leadership responsibilities in the school. They put systems and incentives in place to encourage staff to get involved in teacher-led initiatives and projects. Teachers in these schools view this professional practice as one of support rather than surveillance, where they have an opportunity for dialogue around teaching. In the secondary school context it is often the middle leader who is involved in the day-to-day coordination and evaluation of the teaching programme.

The fourth leadership dimension from Robinson's (2009) meta-analysis *Promoting and participating in teacher learning and development* yielded the largest effect size in student outcomes. Many professional development evaluations reveal that when the principal or school leader(s) are involved with their teachers in the professional development learning initiative, the professional learning is likely to have far-reaching impact and become sustainable in the school. Through the analysis of some core studies, the *Teacher Professional Learning and Development Best Evidence Synthesis* (Timperley, Wilson, Barrar, & Fung, 2008) identifies the important conditions for school-based interventions that make a difference to student outcomes. The core studies are described as those rating medium to high in terms of methodological adequacy as well as medium to high impact for a medium to broad range of outcomes. Timperley et al. (2008) found that professional learning had more impact on student outcomes when leaders provided a range of conditions. These included an actively organised, supportive environment to promote professional learning opportunities and support of the implementation of new practices in the classrooms. Interventions were more successful when leaders developed a learning culture within the school and engaged in the learning alongside the teachers in a school. Robinson's (2007) analysis showed that leaders who participate with staff to understand the difficulties and complexities of the change process required were able to provide the most useful resourcing and support in order to make the changes and embed them effectively.

The research of Robinson (2007) and Timperley et al. (2008) is supported by the work of McLaughlin and Talbert (2007) drawing on two years of studies on professional learning communities in high schools in the San Francisco Bay area from 1999–2001. They found that as part of the professional learning culture of the school, leaders who provided targets for student outcomes and monitoring (whether these were met or not) were also successful. Leaders who fostered internal standards, greater accountability, promoted situated professional learning and reinforced the importance of connecting assessment judgements to teachers' day-to-day teaching engendered strong professional learning communities in their schools. Through collaboration on common work, teachers in these high schools developed new knowledge to improve their instruction and were strategic in bringing in expert knowledge from outside the school. They reflected on their practice, individually and together, and used evidence of student learning to design and evaluate interventions to address learning differentials among student groups. Through this process, teachers created shared language and standards for their practice, within and across subjects (McLaughlin & Talbert, 2007, p. 156). All of the practices described in these reviews (McLaughlin & Talbert, 2007; Robinson et al., 2009; Timperley et al., 2008) as effective in terms of promoting a school culture of

teacher professional learning and development are relevant to the middle leader and department culture of secondary schools. Robinson (2007) asserts that more research is needed on the reasons why school leaders at both senior and middle management levels choose to participate or not in teacher learning activities. In the secondary school setting, the principal relies on the department leaders to provide the instructional advice and expertise required for effective leadership at department level. Hofman et al. (2001) assert that a great deal of the educational decision-making is being transferred to the department heads. Middle leaders have a role to play in modelling and encouraging senior leaders in the school to participate alongside themselves and their department colleagues in professional learning and development. Middle leaders who keep their school leader informed and involved in the professional learning decisions will empower everyone involved and are likely to make initiatives more sustainable.

The physical and social environment of a school has also been shown to have an impact on improved outcomes for students. The results of the meta-analysis by Robinson et al. (2009) demonstrate that the fifth dimension *Creating a safe, caring and orderly school environment* in which staff can teach and students can learn has an effect size of 0.27. Although this effect size is small this is an area where middle leaders can have an influence because they are in the front line in terms of developing collegial relationships, conflict resolution and student management. Factors such as cultural understanding and a respect for difference, having a safe orderly environment with a clear discipline code (includes trust), minimal interruptions to teaching time, protection of staff from undue pressure and effective resolution of conflict as well as effectively dealing with teacher competence issues are all tasks that fall within the realm of the middle leaders, particularly in large high schools. In the work of Heck (1992) and Heck et al. (1991), there was a particularly strong effect among the high school samples when teachers were protected from undue pressure from parents and officials. Further detailed research is required to gain empirical evidence of the way the physical and social features of a school are differentiated within and between departments in a secondary context and what relationships these have — both positive and negative — with student academic outcomes.

The link between school and department leadership

There are some important studies that evaluate the importance of school and department leadership. The work by Hofman et al. (2001) completed in the Netherlands probes the leadership and management of over 100 secondary schools using configuration theory. Central to configuration theory is the assumption that the effectiveness of schools depends upon the fit of several structural characteristics,

specifically the management elements that seem to produce a cohesive and coordinated school. Hofman et al. (2001) assert that research into the effectiveness of secondary schools should examine closely the relationship between the leadership style of the school leaders and those of the department heads. Cohesion, coordination and collaboration within departments and schools are considered crucial to the development of an effective school (Hofman et al., 2001, p. 118). Effective educational leadership is closely connected with the development of a cohesive and coordinated school in which attitudes and values are shared by the school staff, and collective efforts to improve teaching methods and materials are common (Creemers, 1994; Teddlie & Stringfield, 1993). Such a climate makes it obvious for students what the school staff stands for and which set of rules and norms students should comply with. A leadership style monitoring and showing high expectations for all students, promoting school-based staff development, stressing collegiality and cooperation between teachers with a shared value system positively influences student outcomes (Hofman et al., 2001, p. 116). Bosker, Kremers and Lugthart's (1990) research supports this claim as his constructed typology of secondary schools found that pupils in cohesive, goal-orientated and transparently organised schools reach higher achievement levels than other pupils.

The database used for the Hofman et al. (2001) study consisted of 91 schools (school leaders, department heads and teachers) and 5,110 students and their parents. The schools that participated in the research comprise approximately 11% of the 803 secondary schools in the Netherlands. The achievement level of pupils was determined by means of a national standardised mathematical test at the end of their third school year. Four variables were used as covariates: pupil's school recommendation, socioeconomic background, pupil intelligence and pupil motivation for achievement. Briefly, Hofman et al. used multilevel analysis to group secondary schools in the Netherlands into three different types. The Cluster 1 schools scored fairly low on the coordination mechanisms such as creating an orderly environment, department influence on school policy and decision-making processes were insignificant. This type of school scored low on collaboration and mutual agreement between teachers in departments. A modest focus on staff development and cohesion within the working environment typified this group. The Cluster 2 schools had a variation of indicators of effective school and department leadership and just 13% of the schools in the study were designated Cluster 3, which meant they combined effective school and effective department leadership. When the student outcomes in mathematics were analysed with pupil covariates and the school cluster scores, Hofman et al. found that more decentralised processes for decision-making in secondary schools and the sharing of

power among managers at different school levels impacted positively on student outcomes. This outcome corroborates Murphy's (1992) statement that the better schools are themselves more tightly linked, structurally, symbolically and culturally. Murphy (1992) states that effective schools operate more as a whole and less as a loose collection of disparate subsystems. Thus, effective secondary schools seem not so much 'loosely coupled' but rather 'tightly linked' (Hofman & Hofman, 2000; Weick, 1976) and distinguish themselves from less effective schools by revealing a much stronger social and educational emphasis. This supports the claims of Harris et al. (1995) that for departments to be really effective they need to be 'nested' inside schools that are themselves managed effectively. The strength and relevance of the Hofman et al. study is the focus not just on school leadership but the way the school and department leadership link and interface. When the whole school is well coordinated and effectively linked this strengthens and enhances department leadership, which in turn impacts positively on student outcomes.

Many of the studies discussed in this section support the claim that leaders who set and communicate goals and expectations about quality teaching and learning, recognise academic achievements, and promote and participate in teacher learning and development are influential in raising student achievement (Robinson, 2007). The empirical evidence suggests that leaders who are able to develop cohesion in the school and create achievement orientated school policy, which includes regular monitoring of student progress and teacher function, are able to effectively raise student achievement. This work is support by the research of Hofman et al. (2001) as it reveals that when the school and middle leadership practices are effective and well aligned there is a greater positive impact on student academic outcomes in mathematics. These core leadership practices need to be modelled and distributed throughout the school because there is evidence, to be discussed in the forthcoming sections, that suggests that in many schools a great deal of the educational decision-making is being transferred to middle leaders. This is a positive development, given the evidence of the impact of distributed leadership but these department heads and middle managers need to be knowledgeable about the management and leadership practices that will support them to be effective because a substantial proportion of the variation in effectiveness among schools is due to variation within schools, particularly at department level (Creemers, 1994).

The importance and functions of departments and middle managers in the secondary school setting

This section describes the role of middle leaders, how the departments they lead are structured, their professional and collegial relationships and how departments fit within the secondary school physical and structural hierarchy. The organisation of high schools into the 'realms of knowledge' of subject departments is now nearly a universal feature of secondary schools in many western countries despite the widely varying location, size, vision and governance style of schools (Siskin, 1994). Highly standardised departmental labels divide teachers and courses along academic lines and in the New Zealand secondary school context the eight essential curriculum areas of mathematics, English, languages, social sciences, the arts, science, health and physical education and technology are often used for organisational and structural purposes.

First appearing at the turn of the twentieth century, primarily as a term for programmatic divisions, departments designating subject content had become a highly standardised arrangement by the 1930s.These functional divisions connected intimately to the university system and the organisation of knowledge in academia. In sharp contrast to the "mothering plan" of a single teacher for all subjects in a given elementary or primary classroom (Kilpatrick, 1905, p. 475), departmentalising content along disciplinary lines helped to reconfigure secondary schools as "the people's college" (Tyack, 1974, p. 57). The resulting configuration lies between the model of the elementary school (where teachers are teachers, students are students, and the key identifier is grade) and that of the college or secondary school, where key identifiers for both faculty and students are subjects. In high school, students are organised by grade or year level across subjects, but teachers by subject across grades or year levels. Metcalfe and Russell (1997) likened secondary schools to a production line in which the students moved along a conveyor belt from workstation to workstation as they move from classroom to classroom, from subject to subject. This description corresponds with the structural organisation of the large urban secondary schools included in this study, as the students gain their 'dose' of curriculum content in each classroom to prepare them for the next internal or external assessment point.

The head of department is a middle manager and in a large school could be managing between 10 and 18 staff. In the New Zealand context, a subject department would usually comprise a head of department as part of a leadership team of people with 'management units' or delegated responsibility for supporting different aspects of the curriculum such as biology within the science department or junior mathematics in the

mathematics department. Typically, in a large department of 18 teachers there would be at least three additional individuals with leadership roles. The head of department is generally not part of the senior management team or responsible for the overall strategic development of a school, but someone responsible for the operational work of others, namely classroom teachers within a subject area (Busher & Harris, 1999). In some schools these organisational hierarchical distinctions are not neatly delineated. Many staff will be involved in a complex switching of roles and lines of accountability between different aspects of their work. For example, some teachers will be responsible to both academic and pastoral heads of department for different aspects of their work. The demands of these two areas may conflict. Heads of academic departments will also be classroom teachers in their own or other subject areas. Heads of pastoral departments are often referred to as deans[5] in the New Zealand context, and senior staff will work in subject areas and be accountable for this aspect of their work to academic heads of department (Busher & Harris, 1999). Some departmental heads may be members of the school senior management team for reasons other than their departmental leadership.

The investigation of departments as they exist as a sub-unit within a secondary school has been the subject of qualitative research since the early 1990s. In her seminal work entitled *Realms of Knowledge: Academic Departments in Secondary Schools*, Siskin (1994) asserts that four critical aspects of the department emerge. The first is that departments represent a strong boundary in dividing the school; the second, that it provides a primary site for social interaction and for professional identity and community; third is that it functions as an administrative unit with considerable discretion over the micropolitical decisions affecting what and how teachers teach; and the fourth is that as a knowledge category it influences the decisions and shapes the actions of those who inhabit its realm (Siskin, 1994, p. 5). James and Aubrey-Hopkins (2003) found that the tendency of secondary schools to be physically organised into department and faculty blocks could strengthen the view of departments as mini empires that are culturally fragmented from other departments and the school as a whole, with few professional collegial relationships that cross the boundaries of subjects. Wise (2001) describes department collegiality as being bounded by their location in a hierarchical school structure, that coexists within a wider hierarchy. The location of departments as separate units can lead to a strong sense of territorialism, which is not conducive to collegiality between department leaders from different

---

[5] A dean in a New Zealand secondary school context has the pastoral care responsibility for a year level within a school.

curriculum areas and results in strong departmental subcultures within a school (Bennett et al., 2007).

Researchers have found many similarities across countries and jurisdictions in the ways that secondary teachers have described their department structure and professional significance. In Siskin's (1994) studies, teachers reported that the department was important to who they were, consequential in affecting what they did and largely determined how their work was perceived. The literature reveals common descriptions by teachers and middle leaders in terms of their perceptions of subject departments in relation to the structural, social, political and subject identity matters. Over the course of the teacher interviews that Siskin conducted in the early '90s teachers talked about their departments in three substantially different ways:

Social — as "we," the department represents the colleagues with whom they work most closely, whose individual efforts and group norms influence the ways teachers think about and conduct their practice;

Political — as "the department," it plays a primary role in acquiring and distributing resources and responsibilities among teachers; and

Subject — as "English" or "Science," it is the subject matter central to who they are, what they do and how they go about doing it (Siskin, 1994, p. 185).

In their analysis, Busher and Harris (1999) added that the structural organisation and configuration of the department (i.e., size, membership, location, subject affiliation) is important. They add that the status or esteem with which the department is held by the wider organisation and power in social situations and the strategies through which leaders and followers exercise power (Blase & Anderson, 1995; Busher, 1992) is critical to department success. Department power can be recognised within a school through formal authority or informal influence. Subject departments are not just smaller pieces of the same school social environment, or bureaucratic labels (Bennett, 1995), but as Siskin (1994) points out they are separate worlds, with their own "ethnocentric way of looking at things. They are sites where distinct groups of people come together and share in and reinforce the distinctive agreements on perspectives, rules and norms which make up subject cultures and communities" (Siskin, 1994, p. 81).

The professional regard teachers have for their department leaders is critical to the success of a department. The views on subject knowledge and pedagogy held by heads of department are of major concern to staff. Although interpersonal skills are crucial elements of middle leadership practice, this is not sufficient to give these subject

leaders the authority to underpin their status. Their subject knowledge and expertise as a teacher is the basis of the idea that middle leaders should lead by example (Bennett et al., 2007; Sammons et al., 1997). Brown and Rutherford (1999) found that heads of department believed that they were managers of curriculum rather than colleagues. They believed their role should focus on monitoring what was taught and the resulting student outcomes, rather than observing teachers in action in the classroom. Brown and Rutherford (1998) found that middle leaders did not necessarily need to be the best teachers in the department but they did needed to be capable of modelling good practice. This is supported by Fletcher-Campbell (2003) whose study of new heads of department found that they saw their role as an opportunity to develop the teaching of the subject within the department. The literature suggests that the inherent subject and pedagogical knowledge of the department middle leader is an important aspect of an individual's ability to lead and develop the teachers in a department, however they choose to enact that influence.

The leadership practices employed by middle leaders and the way that the staff experience them has been investigated by researchers (Bennett et al., 2007; Harris, 1999; Siskin, 1994).The way that middle managers act as transformational leaders and exercise interpersonal skills will affect the extent to which they can build a genuine collaborative culture and empower teachers is indicative of healthy departments and schools that are likely to bring about improvement in practice (Blase, 1995; Stoll & Fink, 1996). In a review of the empirical research of middle leadership in secondary schools, Bennett, Woods, Wise and Newton (2007) comment that teachers preferred a collegial team-based approach to leadership and management. A culture that combines a positive working environment with professional challenges and teamwork motivates teachers to work together to achieve some common goals. Schein (1990, p. 5) suggests that "the only thing of real importance that leaders do is to create and manage culture," (p. 5) and Siskin (1994) makes a similar claim for departmental leaders. In this respect, middle leaders have points of reference and influence that transcend many of the formal structures within the school. Consequently, their leadership style is fundamentally important in shaping the direction and cohesiveness of the department because the direction and examples they set are experienced day-to-day by the teachers that they lead and the students who they can influence.

The leadership and negotiating skills of the head of department are important to lead the people within the department but equally as important in linking the department to the rest of the school and the overall organisational goals, processes and school culture. Busher and Harris (1999) assert that heads of department who know how to

make contact and bargain with colleagues through an array of micropolitical processes are in a stronger position to implement or defend the policies of their departments than are those who do not. Subject departments therefore play important mediating roles between the demands of a subject area on staff and the demands of the school on them. The extent to which the head of department or middle leaders can arbitrate the demands of the wider school leadership and relate them to the existing culture and practice of the teachers in the department is considered crucial in their success (Bennett et al., 2007). Glover, Miller, Gambling, Gough and Johnson (1999) referred to the middle leader in this context as the "buffer and bridge" and saw this as involving both upward communication of departmental opinion into the wider school hierarchy and downward communication and mediation of external demands. Bennett et al. (2007) assert that in order for a subject leader to play a role in a wider school setting and to contribute to a collegial culture of collaboration both within the department and across the wider school they must first feel comfortable with their own subject colleagues and have earned their professional respect as a supporter and protector of their particular curriculum area.

A further tension for middle leaders and teachers that report to them, is the monitoring of colleagues' teaching performance (Glover et al., 1998). In an accountability culture focussed on student performance, middle leaders must take appropriate actions to ensure quality control and quality assurance within their departments (Bennett et al., 2007). Accountability and data are at the heart of contemporary reform efforts worldwide. Accountability has become the watchword of education, with data holding a central place in the current wave of large scale reform. Policy makers are demanding that schools focus on achieving high standards for all students, and they are requiring evidence of progress from schools that is conceived explicitly in the language of data (Fullan, 1999). Metcalfe and Russell (1997) found that the monitoring of colleagues' teaching quality by middle leaders was only acceptable in a culture of collegiality. They found that middle leaders would not accept the concept of monitoring teachers' work for quality control but would accept the idea of collective learning from each other. In California, teachers in the BASRC study schools (McLaughlin & Talbert, 2007) used evidence of student learning and achievement gaps to focus instructional improvement efforts. They shared responsibility for ensuring that all students met the grade levels in subject areas and designed strategies, materials and practices to achieve this goal. Witziers et al. (1999) describe departments in the Netherlands as frequently collegial in terms of discussing teaching content, but gave little attention to teaching strategies, teacher development or problems in the classrooms. The practice of observing the teachers in a department is not a consistently applied practice at department level.

Glover et al. (1998; 1999) report that some senior staff used the excuse of excessive administration to avoid the awkwardness of entering the classroom of another teacher to engage in monitoring. In four out of the seven schools they studied, formal monitoring and evaluation of classroom work was evident but the middle leaders expressed a tension between collegiality and positive team culture and the monitoring role they exercised to evaluate staff capability. Wise (2001) suggests that monitoring through classroom observation is seen by many departmental members as putting middle leaders in a role of surveillance as opposed to professional support. The research suggests that there appears to be a tension between building a department culture of critical discussion and critique of classroom practice and creating a collegial environment with a positive social atmosphere for teachers.

A review of the existing research suggests that the department can make a critical difference in the working conditions of teachers. Some departments develop norms of social inclusion, mutual support and commitment to students, while others are fragmented and concerned with individual survival. One department may readily command and distribute the resources that teachers need, while another in the same school may not distribute the resources equitably. Department planning for curriculum delivery could be tightly sequenced and coordinated while others might be more fluid and adaptable. These variations can be understood as local and unique differences but it is important to understand the patterns of difference and investigate which combinations of practices improve teaching and learning and therefore make a difference to student outcomes.

Evidence of effective departments in secondary schools

This section is focussed on discussion of the evidence of the specific practices of middle leaders in secondary schools that make a difference to student outcomes. The work of Siskin (1994) is important in terms of its innovation and depth. Her studies are essentially qualitative and ethnographic and concerned with the social, political and subject differences in the schools she studied in depth over four years. Further research more directly linked to student academic outcomes can be found in the work of Harris, Jamieson and Russ (1995), Harris (1998), Sammons et al. (1997) and Dinham (2007). They have approached the field by identifying the characteristics of effective and ineffective departments in secondary schools and linking these characteristics to student outcomes. These studies confirm that the department is the crucial "working unit" and that school performance and departmental performance are not inextricably linked. That is, poor performing schools can contain high-performing departments and the reverse can also be true.

For the purposes of this project it has been critical to understand the methodology and findings of the studies that have linked academic student academic outcomes data to specific leadership characteristics at department level. The findings of a British study of effective departments in secondary schools completed in 1994 by Harris et al. (1995) revealed that departments within schools had different performance in relation to student outcome measures. The research team used six academic performance indicators based on GCSE attainment that allowed them to analyse individual subject areas within the pilot schools for their study. The design of the study allowed the researchers to focus on a small number of departments that showed that on the basis of GCSE results the students studying these subjects were progressing further than might have been expected from consideration of their intake. Apart from the initial quantitative analysis of student academic results, the study was largely qualitative relying primarily on interview data from the senior management team, department members and pupils in the schools. These effective departments were broadly successful because of a collegial management style and a strong vision of the subject effectively translated at classroom level with a focus on teaching and learning. Well-organised resources, assessment and record keeping with effective systems for monitoring evaluation and feedback also featured strongly. These effective departments had a focus on autonomous pupil-centred learning with a syllabus that met their needs and rewarded their efforts. Harris et al. (1995) described specific characteristics of teaching and learning or pedagogical approaches that were evidenced in effective departments. They included an emphasis on students being involved in the learning process and teachers who provided a variety of tasks catering to individual, small-group and large-group situations. Teachers encouraged cooperative learning where pupils work together as part of a team sharing experiences as well as being given different roles and developing their own self-esteem. Students were actively involved in review and reflection of the learning process and teachers developed meaningful, formative and motivational forms of assessment that reinforced and built confidence.

The largest study that considers school and department leadership with student academic outcomes at secondary school level was completed between 1990–1992, by Sammons et al. (1997). The intention of this study was to establish the extent to which information about variations in school and departmental processes were systematically related to schools' effects on specific student outcomes. The report of their research in 94 London secondary schools involved collecting extensive student achievement data from GCSE exams and then interviewing a subset of 30 subject heads, five in each school. The heads of department were asked to identify factors they thought should be taken into account in judging the effectiveness of any department. Mathematics and

English heads of department gave similar responses and over 85% noted the quality of teaching in the department as a critical success factor. The English subject leaders most commonly (over 70%) commented on the extent to which departmental staff worked as a team and the commitment and enthusiasm of staff. The mathematics leaders gave equal stress to departmental staff teamwork and the enjoyment and interest of students in the subject. In addition, sixty percent of maths heads accentuated the importance of the prior attainment of students before their intake into the school, having an important impact on their achievement, compared to 44% of English heads. A quarter of the heads of department thought the development of students' study skills, the stability and experience of teaching staff and the uptake of GCSE and A level courses by students should be used to judge departmental effectiveness. Half of those interviewed believed that examination results were a reasonable measure of effectiveness (Sammons et al., 1997, p. 112). Middle leaders stated heavy workload as a barrier to departmental effectiveness and high levels of staff stress linked with implementing new initiatives and changing syllabi and assessment (Sammons et al., 1997, p. 119).

The results of the study show that after taking account of student intake variance there were a number of factors that had a positive correlation with student academic results in the English departments. Promoting the ability of students to learn independently, a caring pastoral environment, a strong cohesive senior leadership team, low levels of staff shortages and the uptake of the subject at GCSE and A level qualifications all had a positive impact on student outcomes for that subject. Aspects that rated negatively were difficulties in getting parent and community support and too little emphasis on homework by staff. The researchers have drawn attention to the importance of an academic emphasis in the department and the importance of staffing and quality teaching impacting on student performance in the English examination (Sammons et al., 1997, p. 150). Items identified as significant in accounting for positive differences in academic performance for mathematics, included the staff and students' shared belief that the school is primarily a place for teaching and learning. The rare occurrence of staff shortages, like English, also correlated positively. Items that accounted for negative impact on a mathematics department were low student motivation, the head of department's leadership approach being one of coercion as opposed to persuasion and insufficient quality teaching in some classes (Sammons et al., 1997, p. 152). When the quantitative analyses of student achievement data at school and department level were examined with the qualitative in-depth work on departmental leadership the findings of this study showed that both school and departmental processes are significant predictors of a school's academic effectiveness. The strength of the *Forging Links* study

is the link between department leadership and academic outcomes with a focus on English and mathematics, so the design and methodology have been particularly influential on the current project.

A more recent study and one more closely related to the New Zealand context, is found in an Australian study focussed on the practices of middle leaders in 38 Australian state-funded secondary schools where there was evidence of exceptional student outcomes (2007). The definition of educational achievement for this project was defined in terms of both academic and social outcomes and was constructed from a rubric from three interrelated domains outlined in state policy documents[6] related to developing the talents, knowledge, skills and values of students. The project employed a case study approach whereby quantitative data were initially gathered that included public examination data where there was an ability to determine the 'value add.' The qualitative data was used to determine and select schools where there was potentially exceptional educational achievement in a combination of academic departments and cross-curricular areas. Researchers visited each of the 38 schools and recorded their findings through various interviews with key staff, teaching staff forums, recording student and parent 'voice' and document analysis. The analysis of these data using the grounded theory technique of open coding revealed some key concepts of successful heads of department, who were leading teams where student outcomes were higher than in other similar schools. These influences on success included personal qualities and relationships where commitment, energy and enthusiasm for teaching motivated those around them (Dinham, 2007, p. 67). These department leaders were found to be experienced and effective teachers who possessed depth and breadth of knowledge, sound understanding of curricula and were well informed about current developments in their professional field. They were effective advocates for their department and well-respected members of the school community. Their ability to network helped them to operate politically within the school culture and secure resources for their department. They influenced evidence-based department planning and organisation, ensured resources were well used and took a leading role in programming. They facilitated the development of policy with their staff, which aided effective communication. Evaluation and reporting was always well documented and accessible. Effective student discipline and welfare strategies were developed and instigated and adequately followed up. Decision-making was collaborative with an implicit understanding that there would be universal and consistent application of group decisions. This supported collaborative team building and the development of a common purpose or goals. High value was

---

[6] *The Adelaide Declaration on National Goals for Schooling in the Twenty-First* Century (Ministerial Council on Education, Employment, Training and Youth Affairs, 1999)

placed on a wide variety of both in-house and out-of-school professional learning. Staff who reported to these middle leaders were supported to identify their strengths and weaknesses and encouraged to address these. Overall, the prime consideration of these middle leaders was students and their learning because they believed that when students' personal and social needs were met they would be academically successful.

There are considerable ethical limitations in attempting to research ineffective departments within a secondary school; however, one such study has been found to contribute to this literature review. Given the range of effective characteristics described in this paper from the work of Harris (1999; 1995), Sammons et al. (1997) and Hofman (2001), it is not unexpected that Harris found that ineffective departments have a combination of characteristics that collectively result in a dysfunctional departmental culture (Harris, 1998). The methodology employed in this study was qualitative in nature and involved an in-depth case study approach in eight departments across four schools. The participating schools were considered to be relatively effective in terms of overall examination performance but the departments were selected because the academic student achievement data that was analysed showed that they were less successful on this measure. Harris (1998) lists characteristics such as lack of a clear and shared sense of departmental vision, either laissez-faire or authoritarian leadership, a lack of professional and collegial formal and informal communication, poor organisation and inequitable deployment of resources, record keeping and assessment. These departments did not have agreed and detailed schemes of work and where they did exist they had evolved in an ad hoc way. An important finding was that these departments did not simply have ineffective teachers. Within each of the departments studied there was at least one teacher who was considered by staff and pupils alike to be an outstanding teacher. The remaining teachers within the department were in the main not considered to be ineffective or weak teachers by either pupils or teachers who were interviewed. In many cases, the converse was true and the major problem appeared to be that they taught in isolation of each other and did not function effectively as a team (Harris, 1998, p. 273). Siskin's (1994) observations support this as she describes fragmented departments as those with low professional commitment and low social inclusion where decision-making is unclear. Although unique, the Harris (1998) study is important as it supports the theory that ineffective departments can coexist alongside effective departments in generally effective schools and that teachers in generally ineffective departments can continue to perform effectively despite the poor leadership of the department in which they teach.

The extent to which departments and middle leaders can operate successfully in weakly led or failing schools is still being debated. The effective schooling literature suggests that for departments to be really effective they need to function within an effectively managed school (Hofman et al., 2001; Luyten, 1994). In their study of effective departments, Harris et al. (1995) describe some positive effects from the senior management team on departments in the school. A whole-school focus on raising expectations of both pupils and staff and rewarding positive student behaviour was considered important by staff as it increased the motivation of the students in their classrooms. Another factor of whole-school management that encouraged departments to become more effective was the scrutiny of examination and test results at department level by the senior management team. A formative approach to reflecting on these results with a view to ongoing improvement was viewed by effective department leaders as being justifiably held to account for the results of students being taught in their department. However, in general the effective departments did *not* feel that they were operating in whole-school structures and cultures that were particularly supportive of their specialist endeavours and that the senior management team were not sufficiently collegial. In some schools, the management style of the senior leadership team seemed to be in stark contrast to the style chosen by the successful heads of department (Harris et al., 1995, p. 287). Sammons et al. (1997) asked heads of department to rate their satisfaction and effectiveness of the head teacher of their school. In an effective school with all departments categorised as effective, all the department leaders gave the head teacher the highest effectiveness rating. In mixed and less effective schools, the senior management team were rated as ineffective:

> In general, the Senior Management Team (SMT) of the more effective schools appeared to function as coherent and complimentary teams in a way which was not evident in the other schools. Indeed, in the two less effective schools conflict in the SMT was a strong feature. (Sammons et al., 1997, p. 75)

The case study findings showed that heads of department in effective and not so effective schools all wanted more day-to-day contact and communication with the senior leadership team and an opportunity for more consultation and greater involvement in decision-making. Sammons et al. (1997) stress the importance of effective schools benefiting from stability and continuity in the senior management team with credible people in these positions viewed by heads of departments as having the commitment, ability and creativity to do the job. Hofman et al. (2001), Harris et al. (1999) and Sammons et al. (1997) all describe evidence of effective departments thriving in effective schools. Whole-school leadership, policies, systems and processes

can have a very positive impact at department level but this is not crucial for department success. Effective department leaders can impact positively in their cross-school influence as well. Dinham (2007) suggests that the most effective heads of department seek out best practice from outside their department and school and yet they often are neither recognised or utilised to any great degree and in some respects are "hidden treasures" (Dinham, 2007, p. 77).

In conclusion, the evidence suggests that there are clearly identifiable characteristics, attitudes, behaviours and beliefs of middle leaders in charge of successful departments that are evident in a number of studies (Dinham, 2007; Harris, 1998; Harris et al., 1995; Hofman et al., 2001; Sammons et al., 1997). These studies have linked specific dispositions and capacities to positive academic outcomes because when middle leaders make leadership decisions that have "students at the centre" (Ministry of Education New Zealand, 2007) there is an inevitable impact on the quality of the teaching in the department. When the key leadership practices are missing at whole-school and department level or are dysfunctional, students are not so well served by their teachers. The specific practices employed by middle leaders to create the working conditions that motivate teachers in ways that impact positively on student outcomes are a critical focus for this research project, and the results sections of this thesis show which practices are being demonstrated in departments and how they link to student outcomes.

Linking within-school variation and the department as the unit of analysis

Since the 1966 study, *Equality of Educational Opportunity* (Coleman et al., 1966), researchers have continued to pursue questions around school variance and the 'effect' of attending a particular school. Raudenbush and Willms (1995) and Willms (1992) refer to the term "school effect" to identify the differences between a school's average level of performance after adjusting statistically for the intake characteristics of the school. In the past three decades, researchers have paid more attention to whether schools differ in their effects on students with varying ability, gender, family socioeconomic status and ethnicity. This research has provided some important findings relevant to monitoring school performance (Willms, 2000). Willms (2000) described the within-school variation in a Canadian study based on 1996 data from an elementary school climate study for New Brunswick: "Student scores within classes are relatively heterogeneous compared with the variation from class to class, or school to school. The pressure and support for change needs to be directed at particular teachers within schools, not simply at entire schools" (p. 241).

In a study of the relative effectiveness of Post-6 Institutions in England (including schools offering assisted places to underprivileged students), Tymms (1992) found considerable variation between schools in their performance in individual subjects but less variation for whole-school measures. Some differences were found for student outcomes in relation to the type of institution pupils had attended, but the effects were generally quite small and varied across outcome and curriculum measures. Tymms explored differences between schools by comparing the performance of similar students in different types of institution but found the within-school variation more significant. Both Tymms and Fitz-Gibbon (1990) found that the order of influence on student achievement was determined most strongly by the individual pupil, then by the department responsible for a particular performance, then by the school as a whole and finally by the type of school.

There are a number of British researchers who have been investigating the within-school variation problem in secondary schools over the past 10 years and using student examination results for the analysis. Based on quantitative analyses of Key Stage 3 exams in England, Reynolds (2004, 2007) argued that in secondary schools where progress is in line with expectations, there is clear evidence of within-school variation. He stated that if the results of six groups of boys and girls in English, maths and science were analysed, that 80% of schools show 'value added' significantly higher or lower than expected in at least one group. This shows that the within-school variation is by department or subject. Over a 3 year period, 50% of schools have at least one subject in which progress would put them in the top 20% nationally of the subject concerned. Reynolds (2004) contended that the causal factors for variation in performance by subject or department included variation in teacher competence, unreliable implementation of national strategies and the maximised impact of improvement strategies that widened the student achievement gap. A recently completed report that evaluated a professional development initiative tackling within-school variation in 20 training schools in England, offered some useful insights into the behaviours and practices of teachers that are key enablers for reducing variation (Chapman, Mongon, Roxby, & Manns, 2009). In summary, the researchers found that successful schools invest time in identifying and understanding the evidence of variation within the school and then develop specific strategies to work towards a vision for change and improvement. Staff work strategically and collaboratively to identify and implement strategies that strengthen teaching and learning in the school. School leaders are fully involved and lead the introduction of new concepts to staff and ensure that the plan devised to reach the goals is adhered to. In addition, student contributions are sought after, used and valued (Chapman et al., 2009). As schools put more effort

into identifying where the variation in student outcomes is and identify the reasons for these variations, they will be able to develop more cohesive overall school results, where the results are less dependent on a subject or department in which a student was taught.

There is an ongoing argument within the school effects debate regarding the size of teacher effects in terms of the impact of quality teaching on outcomes for students in the classrooms of effective practitioners. This debate is strongly related to the argument around subject or department variance within a school because there is considerable evidence that middle leaders can impact positively on quality teaching within a department. Research on teacher allocation to schools has documented that schools with high proportions of low-income minority students often have difficulty recruiting and retaining high-quality teachers (Darling-Hammond, 1995). In a comparative analysis using measurement of mathematics and reading scores, Nye, Konstantopoulos and Hedges (2004) found that the between classroom within-school teacher effect is always larger in the lower socioeconomic schools. However, the researchers found the within-school teacher variance to be greater than across school for the whole sample of results from the 79 primary schools they studied. They state their data showed that the pattern of between-classroom variation is more pronounced in lower socioeconomic settings.

Supporting that view are those researchers who have found that almost 40% of the variation in achievement in mathematics was due to a difference between classrooms, explained by the quality and effectiveness of teacher pedagogical content knowledge (P. Hill, 1998; P. Hill, Holmes-Smith, & Rowe, 1993; Rowe & Hill, 1997). In an alternative argument, Lamb and Fullarton (2000) found that the variation in mathematics achievement in high schools they studied was 57% within classrooms and 28% between classrooms and 15 % between schools. They found the differences between classrooms and schools was related strongly to the socioeconomic background of the students, student attitude toward mathematics and the student grouping practices of schools and classrooms, as opposed to the quality of the teaching. Willms and Cuttance (1985) simultaneously contradict and support this argument in their study on school effects in Scottish secondary schools. Their analysis suggests that schools with superior levels of performance in one curricular subject tend to have superior levels in other subjects. The finding could stem from a tendency of superior teachers to be attracted to and remain in certain schools (Willms & Cuttance, 1985, p. 303), a theory also suggested by Nye et al. (2004) but they also suggest that

school structures and environment are significantly impacting across all departments in the school.

The department as the unit of analysis in secondary schools

There is a growing realisation that there is a need to examine student achievement data at class and department level as opposed to whole-school analysis, because the evidence is that individual schools are not uniformly effective. Subject departments provide a unit of analysis for within-school variation that can be measured for consistency over time. An extensive data set of student achievement data has now been established in the United Kingdom called the A Level Information System (ALIS) (Fitz-Gibbon, 1991, 1992), which is an indicator system that has been in operation since 1983 and was previously known as the Confidential Measurement-Based, Self–Evaluation (COMBSE) project. Studies of these data by Smith and Tomlinson (1989) and Fitz-Gibbon (1991) show that more variance was accounted for by department level than by school level and the proportion of variance accounted for at the class level was more than for the department level, even though in many schools the class was the department.

In the late 80s and early 90s, Fitz-Gibbon (1992) found that schools with A level grades in the English department were not necessarily achieving the same grades in the mathematics department, or vice versa. Her analysis of O and A level student achievement data of 1,157 pupils located in one local education authority covering the north of England from 1983–86 found that with mean O level grade as a covariate, the effect of subject (English or mathematics) was highly significant. The interaction of subject within a school was also highly significant, indicating that different schools obtained good results in different subjects. Fitz-Gibbon (1992) asserted that the results of the COMBSE (1983) indicated that parents or researchers should not be looking for the best *school* in which to educate their children, but the best department. She suggested that reform efforts should also focus on efforts to improve education *within* schools, department by department, rather than a focus on schools in competition with each other.

In a large study of differential effectiveness of secondary schools in London, Sammons et al. (1997) found that 32% of the schools recorded significant negative effects in some subjects and significant positive effects in others, taking account of prior attainment and background. These schools had mixed effects at GCSE and marked within-school variation, but these results were masked by a reliance on a single whole-school measure of a total GCSE score. When the researchers looked at schools' effects on

total GCSE performance over three years from 1990–92 the whole-school performance score was more stable than those for specific subjects. This study by Sammons et al. (1997) and others by Harris (1999), Busher and Harris (1999) and Brown et al.,(2000) in the late 1990s used mixed methods approaches to identify the features of schools and departments that were similarly and differentially effective in terms of student outcomes.

If school leadership is second only to classroom teaching in terms of impacting on student outcomes (Day & Leithwood, 2008) then middle leaders, and the departments within which they work, are likely to be able to have a more direct impact on the curriculum and pedagogies of their staff than the school leader. Using the department as the unit for analysis to identify differential effects within a school provides a useful measure of effectiveness in terms of academic outcomes, and collecting this data allows measurement of consistency across subjects and stability over time.

The development of a matrix to categorise characteristics and practices of effective middle leadership in departments in secondary schools

In this section, a matrix is established to map the department leadership practices against the general school leadership practices described in the literature as improving student outcomes. This framework has been used to develop the items for the questionnaire that was administered with middle leaders and teachers as part of this project.

Table 2.2 below maps the literature on middle manager leadership effectiveness against the *Best Evidence Synthesis* dimensions determined by Robinson et al. (2009) and related empirical school leadership research of Leithwood et al. (2008) and Day et al. (2009). In many cases, the themes are very closely linked but there are specific leadership practices of middle leaders that are intrinsic to that role that have been described in detail in the right-hand column of Table 2.2.

The list of effective middle leadership characteristics in the matrix is drawn from the work of Bennett et al. (2007), Busher and Harris (1999), Dinham et al. (2007), Glover et al. (1998), Harris (1997, 1998; 1999; 1995), Sammons et al. (1997) and Siskin (1991, 1994). These characteristics were critical in constructing a questionnaire that would be robust enough to gather the information required to test the hypothesis of this project and answer the research questions. This matrix provided the theoretical framework for the development of the items for the questionnaire that are detailed in the right-hand column of Table 2.2.

It is recognised that some of the individual items could relate to more than one aspect of middle leadership and many items are not exclusive to one dimension. For example, *Students who are struggling to achieve* could relate to dimension two about strategic resourcing to recruit staff with the dispositions and capacities required or dimension three about planning, coordinating and evaluating teaching and curriculum. The matrix provides an organising framework to ensure that the items are linked to existing theory and that the main ideas that exist within that theory are assessed by items within the questionnaire. The last category, not evidenced in the school leadership literature but in the middle leadership research, has been labelled *Collegially focussed on staff and student welfare*. The items related to this middle leadership area were further developed and adjusted when the questionnaire was piloted.

Table 2.2

Middle Leadership Matrix: School Leadership Dimensions From Direct Evidence Mapped Against Characteristics of Effective Middle Leaders to Inform the Development of the instrument

BES Leadership Dimensions (Robinson et al., 2009)	Direct evidence from the literature on schools with effective school leadership	Evidence of characteristics of effective departments	Items developed for the questionnaire, derived from the existing literature (item numbers relate to the order in the questionnaire)
1. Establishing goals and expectations	Includes the setting, communicating and monitoring of learning goals, standards and expectations, and the involvement of staff and others in the process so that there is clarity and consensus about goals. Building vision and setting directions, which includes building a shared vision, fostering the acceptance of group goals, demonstrating high-performance expectations. Redesigning the organisation, improving the quality of the communication throughout the organisation.	Effective middle leaders: • have a clear and shared sense of vision which is evident in the constant and ongoing professional talk at both a formal and informal level. • involve departmental members in the shaping of departmental policy that are in line with the goals of the school. • translate the perspectives and policies of senior staff into the individual classrooms.	The leaders in this department ensure that: 1. School goals are translated into clear department goals. 2. Teachers have performance appraisal goals and plans that relate to the department and school goals. 3. Teachers understand how their classroom work contributes to the department and school goals. 4. The department has clear and measurable academic goals.
2. Strategic resourcing	Evidence of effective school leadership includes: Leaders aligning resource selection and allocation to priority teaching goals. Includes provision of appropriate expertise through staff recruitment. Providing the flexibility to recruit staff with the dispositions and capacities required.	Effective middle leaders: • manage their resources equitably to the mutual enhancement of the whole department and to advantage the students. • ensure decisions and problem-solving is focussed on improving the teaching and learning for all pupils. • liaise and seek information from other important areas of the school and represent the views of the department to the senior managers. • are well connected and networked externally to other subject and/or industry experts whose expertise could support the work of the department. • ensure that quality teachers are employed and staffing shortages are quickly resolved.	The leaders in this department ensure that: 5. The department budget is aligned to priority teaching goals. 6. Staff have equitable access to department resources. 7. Resources are prioritised to build a culture focussed on student learning. 8. Meeting time is allocated to support the achievement of priority teaching and learning goals. 9. The department provides a safe, supportive, and well-organised environment for teaching and learning. 15. They are involved in the

BES Leadership Dimensions (Robinson et al., 2009)	Direct evidence from the literature on schools with effective school leadership	Evidence of characteristics of effective departments	Items developed for the questionnaire, derived from the existing literature (item numbers relate to the order in the questionnaire)
			recruitment of high-quality staff for the department.
3. Planning coordinating and evaluating teaching and curriculum	Evidence of effective school leadership includes: Direct involvement in the support and evaluation of teaching through regular classroom visits and the provision of formative and summative feedback to teachers. Direct oversight of curriculum through schoolwide coordination across classes and year levels and alignment to school goals. Managing the teaching and learning programme, which includes staffing the teaching programme, providing teaching support, monitoring school activity and buffering staff against distractions from their work. Setting the stage for new cultural norms related to performance, distributed forms of leadership to sustain high levels of performance	Effective middle leaders: • are role models who are viewed as expert practitioners who are up-to-date with developments in their field. • develop effective and systematic mechanisms for evaluating department and teacher performance, while monitoring pupil progress. This information is collected using a variety of means. Data are used to evaluate department performance and includes soliciting the views of students. • have as their prime concern students and their learning. They set protocols in relation to teaching and learning supporting formative assessment and feedback to students.	The leaders in this department ensure that: 13. Students experience the classroom environment as supportive of their learning 14. There is an expectation of consistent high-quality teaching. 17. Students who are struggling to achieve are taught by the most effective teachers. 19. All students have the opportunity to attempt NCEA standards that will provide access to university. 20. Assessment data are used to plan teaching. 21. Teachers set academic goals that are based on analysis of their students' achievement data. 22. They lead discussion about students' results with staff. 24. Students receive high-quality, timely feedback that helps them to improve their work. 25. Teachers are provided with effective and timely advice about any concerns they have about their teaching. 29. Student feedback is regularly collected, analysed and acted upon.

47

BES Leadership Dimensions (Robinson et al., 2009)	Direct evidence from the literature on schools with effective school leadership	Evidence of characteristics of effective departments	Items developed for the questionnaire, derived from the existing literature (item numbers relate to the order in the questionnaire)
4. Promoting and participating in teacher learning and development	Evidence of effective school leadership includes: Leadership that not only promotes, but directly participates with teachers in formal or informal professional learning.	Effective middle leaders: • support staff in a wide variety of both in-house and wider professional learning that will broaden and deepen their skills and knowledge and therefore build department capacity. • support staff to identify and address their strengths and weaknesses. • foster a climate for improvement that means that staff are willing to change existing practices. • mentor colleagues and encourage professional learning and development.	The leaders in this department ensure that: 18. They make time to observe department teachers at work in their classrooms. 26. They take care of their own professional development needs and share their new learning with colleagues. 28. Professional development ideas are implemented and used in the classroom.
5. Ensuring an orderly and supportive environment	Evidence of effective school leadership includes: Protecting time for teaching and learning by reducing external pressures and interruptions and establishing an orderly and supportive environment both inside and outside classrooms. Understanding and developing people, which includes providing individualised support and consideration, fostering intellectual stimulation, modelling appropriate values and behaviours	Effective middle leaders: • ensure that management and organisation is carried out so that key elements of teaching and learning are organised in an optimal way. • are highly organised and work collaboratively to generate detailed and collectively agreed schemes of work. • ensure that student discipline is effectively managed and any conflict in the department is quickly and effectively resolved. • set regular meetings with a relevant agenda and outcomes. • provide a caring pastoral environment leadership style which empowers others.	The leaders in this department ensure that: 10. Staff handle student discipline problems with fairness and equity 11. Problems that staff raise about the work environment are addressed in a timely and effective way 12.Conflict within the department is quickly identified and resolved 23. Record keeping processes are efficient and effective

BES Leadership Dimensions (Robinson et al., 2009)	Direct evidence from the literature on schools with effective school leadership	Evidence of characteristics of effective departments	Items developed for the questionnaire, derived from the existing literature (item numbers relate to the order in the questionnaire)
Additional Characteristics Collegially focussed on staff and student welfare	Evidence of effective school leadership includes: School leaders redesigning the organisation, which includes building collaborative cultures, restructuring and re-culturing the organisation, building productive relations with parents and community, connecting the school to its wider environment. Indirect evidence of leaders who create educationally powerful connections, particularly between school and home, engage in constructive problem talk and select and develop well designed smart tools based on valid theories can make a difference to student outcomes depending on the context.	Effective middle leaders: • lead by modelling professional, collegial and co operative ways of working • are described as having personal qualities and relationships where commitment, energy and enthusiasm for teaching motivates those around them • are effective advocates for their department are politically astute and provide a collegial 'bridge' between the department and the school • ensure that the personal and social needs of students are met in order to underpin their academic success.	The leaders in this department ensure that: 16. They take responsibility for the learning of all students in the department 27. There are opportunities for teachers to give feedback to departmental leaders and vice versa. 30. The staff work together as a team and support each other. 31. Leadership roles are shared among the department's teaching staff according to individual strengths and interests. 32. The department is a collegial place to work where staff support each other. 33. There are robust discussions about successes and challenges. 34. Staff talents and achievements are recognised and celebrated.

Conclusion

This chapter reviewed the school leadership and within-school variation literature to reflect the substantial research and synthesis of that material that links strongly to school and student effects. The empirical evidence of the effectiveness of middle leaders in secondary schools is less robust and extensive and the largest study in the area was conducted in the early 1990s. The best of this work in the last 20 years has identified critical leadership dimensions of middle leaders, how effective departments function and the behaviours and practices that can be linked to improved student outcomes. Using the department in a secondary school as the unit of analysis has strong support from school effectiveness researchers where there is evidence to suggest that the outcome measures can be equitably compared. The literature review reveals the importance of further work being undertaken in establishing the link between the observed and reported behaviours of middle leaders and how these impact on departmental effectiveness and academic outcomes.

Research questions

The key questions that have guided the development of the research methods and tools used in this project.

1. What evidence is there in large urban Auckland secondary schools that there is variation in academic student performance between English, mathematics and science departments?

    a. Are the student academic results between the departments within a school variable? Are the results consistently variable or stable by year?

    b. What is the extent of within-school variation for schools in this sample?

    c. How do student academic results in a subject compare across schools?

2. To what extent do middle leaders and the teaching staff in English, mathematics and science departments report the occurrence of effective instructional and organisational practices?

3. What are the relationships between academic outcomes of students in English, mathematics and science departments and the middle leadership practices that have been reported as occurring in those departments?

Chapter Three: Methodology

Introduction – Outline of the methodological approach and research activity

In this chapter a description of the methodological approach to the three phases that comprise this thesis will be explained and the processes and components of the project will be described. The methodology sets out to identify the within- and between-school variance problem in terms of academic achievement, the difference in department and subject performance in urban secondary schools and the middle leadership attributes that could support effectiveness and success. The rationale for the selection of methods used throughout the phases will be described and the subsequent chapters will provide further detail of how the methods were employed. The methodology selected required a design that would enable the student academic outcome data collected to be correlated with evidence of middle leadership effectiveness, therefore establishing the patterns and links between the practices of middle leaders and academic student outcomes. Raudenbush (2005) argued that education research was strong when inferences based on one form of inquiry were checked or "triangulated" against inferences based on an alternative form of inquiry characterised by complementary strengths. This project involved inquiry into large publicly available data sets of student academic records at school and department level and then examining the relationships between the academic outcome measures and the reported middle leadership practices in those subject departments.

The design of this project involved first establishing the extent of the within-school variance problem through quantitative data analyses of NCEA Level 1 results and second collecting evidence of the effectiveness of middle leadership in secondary school departments. Further, academic student achievement data at NCEA Levels 2 and 3 were required to investigate relevant patterns and relationships between the grades students achieved and the practices of the middle leaders in the departments in which they were studying. There were three methodological phases within the project. Phase 1 involved a quantitative analyses of student achievement data in NCEA Level 1 (15-year-olds) in English, mathematics and science from 41 urban Auckland secondary schools. In Phase 2, a middle leadership questionnaire was developed and administered to middle leaders and the teachers that report to them in 30 departments across 10 schools. Phase 3 involved further quantitative data analyses of student achievement of NCEA results at Level 2, 3 and scholarship (16–18-year-olds) in

English, mathematics and science in the 10 Phase 2 schools in order to extend the understanding of student academic achievement in each of those departments.

Phase 1 was designed to address the first research question regarding the evidence that exists of differences in academic outcomes between English, mathematics and science departments. Student academic outcomes in these departments were collated over a period of three years (2008, 2009 and 2010) with a focus on Level 1 NCEA (15-year-olds). Goldstein (1997) remarked that the judgment of teacher performance should not be judged by a single cohort of students but rather on performance over time and Sammons (1996) commented that change in school and departmental effectiveness is likely over time periods of one or two years due to changes in staff, pupil intake and ethos. In the current study, results from within the same school were compared and subject performance across schools with similar student cohorts were compared to national norms. In Phase 2, the second research question that interrogates the effectiveness of leadership practices occurring in departments was addressed through the administration of a questionnaire and analyses of the data. The items in the questionnaire were developed from the key theories developed from the literature review and the questionnaire was administered to middle leaders and the teachers that report to them in 30 departments within 10 large urban secondary schools. The analysis of further student academic achievement data in Phase 3 supported the interrogation of the third research question that sought to establish the extent of the relationship between the academic outcome data and the reported leadership practices in the departments. The questionnaire results were subsequently correlated with the academic student achievement results for NCEA Levels 1, 2 and 3 in English, mathematics and science departments within the 10 schools to determine the effective leadership features of departments that were linked to relatively higher or lower academically performing departments.

Selection of the sample schools for Phase 1

The literature shows that research aimed at elucidating school and class level effects implies sampling designs that are sufficiently large to allow for the simultaneous estimation of effects at the school, class/teacher and student levels. Goldstein (1995) and Tymms (1993) argue that at secondary level departmental effects are also likely to be significant and should be built into the design of research on school effectiveness. The rationale for selection of schools was based on the need to compare schools of a similar size. Therefore, state-funded, large urban schools from one city area with a

school roll of over 1,000[7] students were the criteria for selection. Goldstein (1987) argues that with a department of 60 or more pupils, performance indicators become reliable and meaningful. All schools selected for the quantitative data analysis had a minimum of 100 student results represented within any academic department. For the first phase of the quantitative analysis the data being used is at the student cohort and department level in the same year so there are no cohort differences by subject to impact on the overall findings. In order to compare department and whole-school results, the grades for the most commonly assessed internal and external achievement standards for each department were collated. The national results for all schools were also collated by decile so that comparisons could be made with national norms. For the purposes of this study, the decile rankings are useful to control for economic disadvantage in the analyses and relevant comparison of student achievement data. Decile ranking is important in the organisation of the school and department data because it assists the comparability of similar schools to ascertain value-added information. For example, a decile 4 science department could have student achievement data that are comparable to a decile 8 school. When compared to national data it could be performing far in advance of schools with a similar student population and decile rating.

The state-funded schools selected include very diverse student populations. The sample includes single-sex boys' and girls' schools, co-educational schools, catholic schools and many of the schools include students from a wide range of cultural and religious backgrounds. The literature about sample size and selection warns about the problems of selecting schools with student attrition and high transience rates (P. Hill, W & Rowe, 1996; Stringfield, 1994), which would be experienced by many of the schools in the sample. In this study, the data being used were at the student cohort and department level for each of the three years. Individual student data were accessed in terms of the numbers of results for each department for each year. Individual student's results for each of the subjects were not accessed and compared. For example, an individual's performance in English, mathematics and science was not being compared but the student cohort results for the department for that year were analysed. Three years of data were used to get reliable indicators of department performance. These strategies were designed to mitigate for the problems associated with student and cohort variability. It is also worth noting that the examination results analysed for this study have not been norm referenced. There is no statistical moderation between subjects. The same group of students have completed examinations in English,

---

[7] Schools were selected for the study in 2007 and there were considerable fluctuations in school rolls throughout the four years of the project. Some school rolls fell below 1,000 pupils by 2010 and other school rolls increased.

mathematics and science and their achievement has been graded against pre-existing standards that have been determined by the curriculum.

An additional complexity in the qualifications landscape of New Zealand is that in some state-funded secondary schools, principals and boards have elected to offer Cambridge International Exams (CIE) at Years 11, 12 and 13 as well as, or instead of, the national NCEA qualification. The data from schools using CIE were removed from the original data set because the student cohort sitting the NCEA exams at Year 11 were not a representation of the full student population of that school so could not be compared to schools where all students were entered into one examination system. Three decile 10 schools were removed from the original data set of 44 schools and the results from the remaining 41 schools were analysed.

Method of generating quantitative data for this project

The performance of each department was determined by analysis of the quantitative National Certificate of Educational Achievement (NCEA) data which are publicly available via the New Zealand Qualifications Authority (NZQA) website. The student results were considered in relation to achievement against national norms for the decile ranking. Both internal and external NCEA assessment data for Year 11 (15-year-old) students in English, mathematics and science were collected from the NZQA website (NZQA, 2011c) and entered into Excel. Students in New Zealand are required by law to be enrolled in a secondary school until the age of 16 which means that the greatest majority of young people attempt the Level 1 qualification in Year 11. In Phase 3, additional student academic outcome data for NCEA Levels 2, 3 and scholarship were collected and analysed for the 30 departments whose teachers and middle leaders completed the questionnaires to build a deeper picture of department effectiveness. Table 3.1 details the number of results by year and subject that have been collated for the Phase 1 schools, giving a total of over 2.12 million results. A further 1.3 million results were collected in Phase 3 for Levels 2, 3 and scholarship.

The academic results of a standards-based assessment system can be quantified by school and school type in order to make relevant statistical comparisons. The Grade Point Average (GPA) was calculated by developing a numerical system for calculation purposes from standards that are assessed as Achieved, achieved with Merit or achieved with Excellence to a system that is currently employed by The University of Auckland.[8] At the introduction of NCEA in 2001, a GPA was introduced, based on the grade achieved for achievement standards. In calculating the GPA, the university

---

[8] Starpath: Project for Tertiary, Participation and Success

system assigned zero to 'Not Achieved,' 2 for 'Achieved,' 3 for 'Merit' and 4 for 'Excellence.' The difference between Achieved, Merit and Excellence is 1 but the grades 0, 2, 3, 4 do not represent a continuous scale. In the system employed in the project above, Merit is proportionally worth 1.5 more than Achieved (3/2) and Excellence is worth 1.3 more than Merit (4/3).

Once the GPA was calculated for the standards achieved by 15-year-olds in the school, a comparison of levels of achievement between schools, departments within one school and patterns of student achievement across higher and lower performing departments in relation to decile was achievable. The difference between the highest and lowest performing departments was determined by finding the mean for the GPA results in English, mathematics and science for 2008, 2009 and 2010. The performance within a subject across the sample of schools was also able to be established.

Table 3.1

Total Number of NCEA Level 1 Student Results for the 41 Schools in the Sample in English, Mathematics and Science Collated in 2008, 2009, 2010

Year	Total Level 1 NCEA Results			
	English	Mathematics	Science	Total
2008	391040	195179	273037	859256
2009	386488	191073	254637	832198
2010	197760	102218	130348	430326
Total	975288	488470	658022	2121780

The student results that have been analysed for this study are by standard or result. For example, school 10A has 1,906 results in English at Year 11 in 2008. This means that over 200 students will have completed internal and external assessments (exams) to assess their knowledge against pre-set standards in English. Each student will have attempted approximately seven achievement standards in each subject.[9] There are differences between schools in terms of the number of standards student work is assessed against. For example one of the larger high decile schools (10B) the total number of results available to analyse from the English department is 4,150 and in a smaller low decile school (1C) the number of standards entered by students studying Level 1 science is 161. By selecting schools with a roll of over 1,000 students there were enough results to make the analyses statistically reliable because at least 200 students would have been enrolled in Level 1 NCEA. In schools this size, if a proportion

---

[9] Schools and departments can make choices about the number of achievement standards offered to students. Table 1.1 shows a typical NCEA Level 1 mathematics course offering seven achievement standards with a total of 19 credits for the year's course.

of students do not sit or attempt the standard, as is the case in school 1C, there are still sufficient results for more than 60 students.

The numerical value for achieved grades was obtained by multiplying all of the achieved results in that school and in each of the departments by two. This process was repeated for Merit, which was multiplied by three, and Excellence, by four. Total GPA for each department was obtained by initially adding the product of Achieved, Merit and Excellence points to create a total of all points. Finally, the GPA for the department was obtained by dividing total points by the grand total of the number of student results in that subject. These results for GPA were transferred to another Excel spreadsheet, not as the raw number, but as a fraction, for example 2,617/1,769, where 2,617 is the total number of points and 1,769 is the total number of student results.

The GPA for each of the sample schools was compared by calculating a GPA for the national results for each decile ranking. This entailed extracting the national Not Achieved, Achieved, Merit and Excellence results from the National Qualifications Framework Statistics website for every domain that had previously been calculated at school level. In order to ensure exact comparison between the national results and that of the sample schools, each national result by decile ranking was treated as if it were another school. As GPA was calculated for each subject it was transferred as a fraction and calculated in the same way as the department results. This allowed each school to be compared with the national average in the same decile ranking. In a separate spreadsheet, the total results for Not Achieved, Achieved, Merit and Excellence were also converted to percentages for each department.

The process for collecting the Level 1 data for each of the standards in the subjects for the 41 schools was repeated in May 2009, 2010 and 2011 when the confirmed results from the previous year were released on the NZQA website. This enabled a 3-year analysis to be completed in the 2010/2011 years.

Establishing the extent of within-school variation

In order to test the hypothesis that there is within-school variation in student outcomes at department level, the effect size for the difference in student achievement between departments in each school was calculated, in order for meaningful comparisons to be made. The size of the relationship between the variables (department performance) needed to be tested to determine the significance in the difference between the highest and lowest performing departments.

Cohen's *d* (Cohen, 1990) was used to calculate the difference between the department results. Using the GPA the maximum difference between the highest and the lowest performing departments was calculated for each school. For example, schools showed one of the following six combinations: English high – Science low, Maths high – English low, Science high – Maths low, Maths high – Science low, Maths low-Science High, English high – Maths low. The national results were included as part of the calculation. A calculation was then made to determine the standard deviation for each department over all 41 schools plus the 10 sets of national results ($n = 51$) and then the average standard deviation between each pair of departments was calculated. This meant that the effect-size calculation for each school was relative to the national results.

The effect size for each school was achieved by calculating the difference between the highest and lowest performing departments for each school and then applying Cohen's formula (Cooper & Hedges, 1994, p. 238) to divide this number by the appropriate average standard deviation for the national results. The interpretation of effect sizes is generally based on Cohen's (1988) benchmark where 0.20 is small, 0.50 is medium and 0.80 is large (Hattie, 1992).

Tests performed on the data in Phase 1

A number of tests were performed in order to determine the extent of within-school variation of student academic outcomes over a 3-year period. This allowed comparability between departments within-school as well as across-school comparisons to be made over time. To determine the extent of within-school variance, the overall variance patterns of departments within schools and subjects across the sample, all results were compared to the national average results for decile.

Data collected included:

1. The NCEA results for every department for Level 1 for 2008, 2009 and 2010 in English, mathematics and science in order to calculate a GPA and determine an average GPA for the three years.

2. The percentage of students either not achieving or gaining Achieved, Merit or Excellence in Level 1 NCEA English, mathematics and science within each of the schools in 2008, 2009 and 2010. This measure was important for comparative purposes because the number of Merit and Excellence grades achieved per department was used as an indicator of quality teaching.

3. The decile for each school, which provided a broad socioeconomic indicator of the students attending each of the schools included in the study and was an additional measure for comparative purposes.

Tests were then performed to:

- Calculate the effect size for the variation within each school. This was calculated by determining the difference in the GPA between the highest and lowest performing department within the school compared to the national norms. An effect size for each school for each of the three years was calculated then averaged.

- Compare the performance of each school to other schools with students of similar socioeconomic background by using school decile. Each school's GPA was compared with the national norm for that decile, using an average GPA calculated over three years. This allowed an immediate indicator of how schools were performing on overall student academic performance in English, mathematics and science compared to other similar schools.

- Determine the level of academic student outcomes within each department by calculating the percentage of Excellence and Merit grades achieved by students within each department. This was then averaged over three years.

- Compare within-school variation by subject, across the three years. This identified the subjects that were the highest and lowest performing in each school and the consistency of that performance over three years.

The data were analysed between 2008–2011 using Excel and SPSS to determine the extent of the statistical significance of differences in the student academic outcomes between schools of different deciles.

Phase 2

In 2010, Phase 2 of the project involved gaining ethical approval from The University of Auckland Human Participants Ethics Committee to approach each principal of the 41 schools in the Phase 1 sample to request further research be conducted in each of the English, mathematics and science departments in their school. This phase of the research involved the administration of a measurement instrument designed to investigate the effectiveness of the middle leadership practices within each of the three departments in the school. Ten principals from the original sample of 41 schools agreed to participate in this phase. These 10 schools comprised a mix of decile, size and school type and provided a reasonable sample from the original 41 schools. The

questionnaire was administered in schools in November–December 2010.The detail of the development and distribution of the questionnaire is described in Chapter Five.

Phase 3

Once the 10 Phase 2 schools had been established, further information regarding student academic outcomes was compiled from the NZQA website. The Level 1 NCEA results were collated and the Level 2, 3 and scholarship results for the English, mathematics and science departments in each school were added. This enabled an extensive understanding and knowledge of student academic outcomes in each of the thirty departments in 2008, 2009 and 2010. Of interest were the number of Level 2 and 3 Merit and Excellence results that students achieved in each subject.

Specifically the data collected in Phase 3 included:

- NCEA Level 1 Not Achieved, Achieved, Merit and Excellence results and national norms for English, mathematics and science achievement standards in 2008, 2009 and 2010

- NCEA Level 2 Not Achieved, Achieved, Merit and Excellence results and national norms for English, mathematics and science in 2008, 2009 and 2010

- NCEA Level 3 Not Achieved, Achieved, Merit and Excellence results and national norms for English, mathematics and science in 2009 and 2010

- Scholarship results for 2009, 2010

The collation of this information enabled the data to reflect the achievement of the same cohort of students through their final three years of secondary schooling within each department in each school. By collecting 2009 data for Level 3 and scholarship the results for that group could be compared to the group in the previous year. It was also important to collect the student achievement data for 2010 as this was the year the questionnaire was administered.

Analysis phase

The results from the middle leader and teacher questionnaires were correlated to determine which leadership practices were most frequently evident in the department. A factor analysis from the questionnaire responses was completed using SPSS software. The items grouped themselves into six factors that are described in Chapter Six. Once the leadership factor scores were determined their relationship with student

achievement results in each of the 30 departments in the 10 schools was analysed to determine which middle leadership practices had predicted student academic achievement.

The key analyses for this project was identifying the intersection between the student achievement data and the middle leadership practices described in the questionnaires. An analysis was completed to determine the relationships between the middle leadership factor scores for each department and school and the scores for student academic outcomes. This system allowed departments and schools to be ranked and compared on the basis of their middle leadership and academic outcomes scores. Student academic outcomes were compared to national norms to provide value-added estimates that were critical for comparative purposes to identify departments and schools that were successful for a particular student cohort. The relationships between questionnaire responses and value-added estimates of whole-school effectiveness were examined to establish which school and departmental attributes were associated with positive outcomes for students.

This chapter has outlined the broad methodology for the whole project by describing the way student achievement data was collated and measured in order to determine the extent of within-school variation and link student academic achievement outcomes with identified middle leadership practices. The next chapter provides detail of the Phase 1 analysis of results from 15-year-olds in the 41 secondary schools over three years.

Chapter Four: Phase 1 Results

This chapter outlines the results of the analyses of three years of NCEA Level 1 results for English, mathematics and science in the 41 sample schools. Microsoft Excel and SPSS were used to enter the data and perform statistical tests in order to determine overall school performance and the extent of within-school variation. The data were used to determine whole-school performance in relation to national norms, how school academic performance was related to decile, within-school variation and subject variance within and across schools in the sample.

Overall school performance

An initial ranking of the schools in the sample was determined by using one GPA, calculated by determining a mean for each subject over three years. An overall mean GPA across the three subjects was determined for each school. Each school has been allocated a code and the number refers to the decile of the school. All schools with the first number 10 indicates that they are a decile 10 school; therefore, students are socioeconomically advantaged. Figure 4.1 shows that as would be expected the higher decile schools tended to be performing at an overall higher level than the lower decile schools, although there are some variations in that pattern. For example, school 8F, is ranked the 8th lowest performer over three years. It is ranked adjacent to the decile 2 and 4 schools, where students would generally be considerably more economically disadvantaged. Conversely, school 4B is ranked 13th from the top and above the national cohort of decile 9 schools, despite its lower decile. Hence in this school, students who are generally more economically disadvantaged are performing academically at the same level as students from a more advantaged socioeconomic background. In this ranking, there is a decile 10 school (10E) performing at a lower level than the national cohort of decile 5 secondary schools.

The figures in this chapter show the national results for each decile (black bar) that have been calculated from all schools of that decile across New Zealand. Inclusion of these results enables the individual schools in the sample group to be compared to a national norm of schools in the same decile so that the relative performance of the sample schools in the urban Auckland area can be compared. Figure 4.1 shows the largest national achievement gap is between the decile 9 and 10 schools and the decile 1 and 2 schools. The national results for decile 10 appear to be higher than for decile 9 but in this sample two decile 9 schools (9E and 9F) perform above the national average for decile 10 schools.
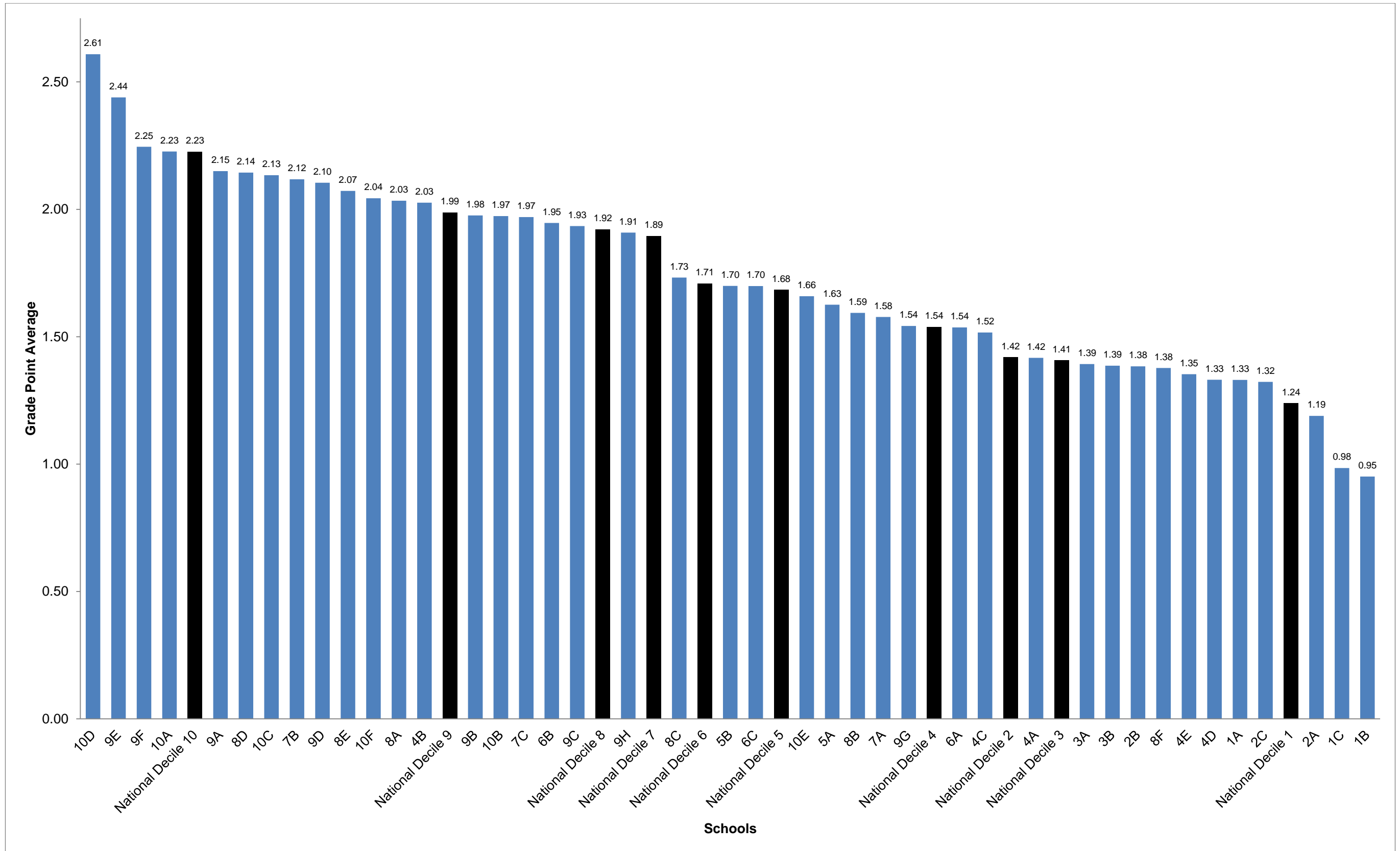
Figure 4.1. Ranking of schools based on Average Grade Point Average Level 1, English, mathematics and science 2008, 2009, 2010

The ranking of the 41 schools in Figure 4.1 is very similar to the ranking shown in Figure 4.2 of school results calculated in terms of the total percentage of Merit and Excellence results for English, mathematics and science over the three years. The percentages for each department's performance have been calculated over the three years and then a mean has been calculated across the departments to give a final percentage for the purposes of ranking. The accuracy and stability of these calculations is supported by the outcomes for both 'tests.' The top seven schools are the same in both graphs (10D, 9E, 9F, 10A, 9A, 8D, 10C). The ranking order is not identical but many schools remain in a similar position throughout the ranking. However, some schools do perform differently in this measure. Figure 4.2 shows that school 4B out-performs its decile by 15 places but is ranked lower in this test at 16th from the top. This would indicate that a lot of students achieve the standard but the proportion of Merits and Excellences is not at the same level of comparative performance. Conversely, school 8F climbs in ranking to 26th from the top in this measure as opposed to 34th from the top for GPA. Figure 4.2 shows how school results for the sample compare to national norms in this measure. Schools 8B and 9G are performing well below norms for their decile and school 10E performs at the same level as decile 8 schools whereas the GPA results for school 10E were comparable to decile 5 results.
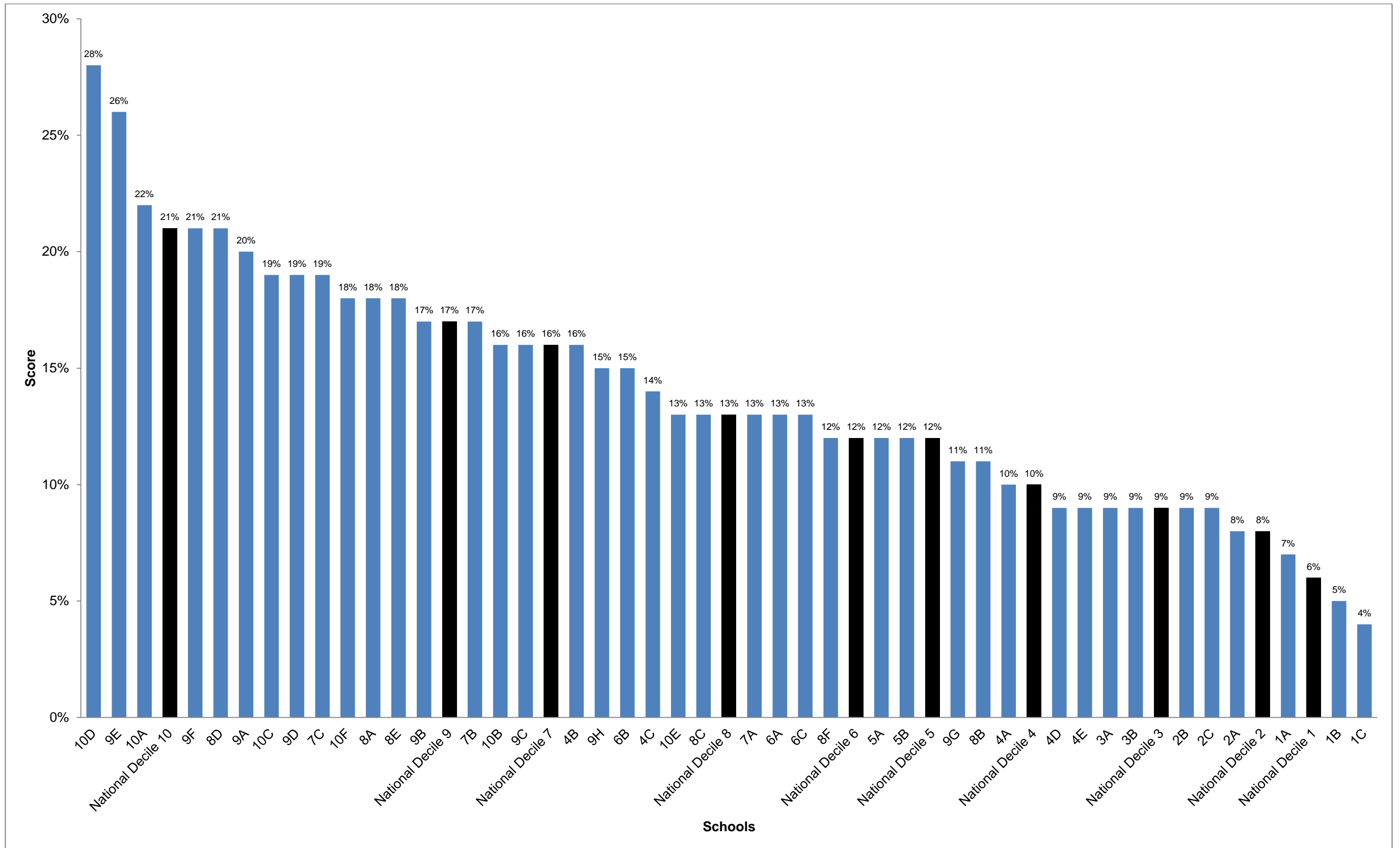
Figure 4.2. Total percentage of Merit and Excellence results for English, mathematics and science 2008, 2009, 2010

Figure 4.3 compares the overall Level 1 NCEA English, mathematics and science performance of students in schools in the sample in relation to decile of school, when the mean is calculated over three years. Figure 4.3 shows that when using this measure there is a considerable range of performance of schools within one decile. There is a clear overall trend for the higher decile schools to be higher performing but there is variation within the sample. In the decile 1–3 schools where students would be substantially economically disadvantaged, overall student performance in English, mathematics and science at Level 1 is below the GPA of 1.5.
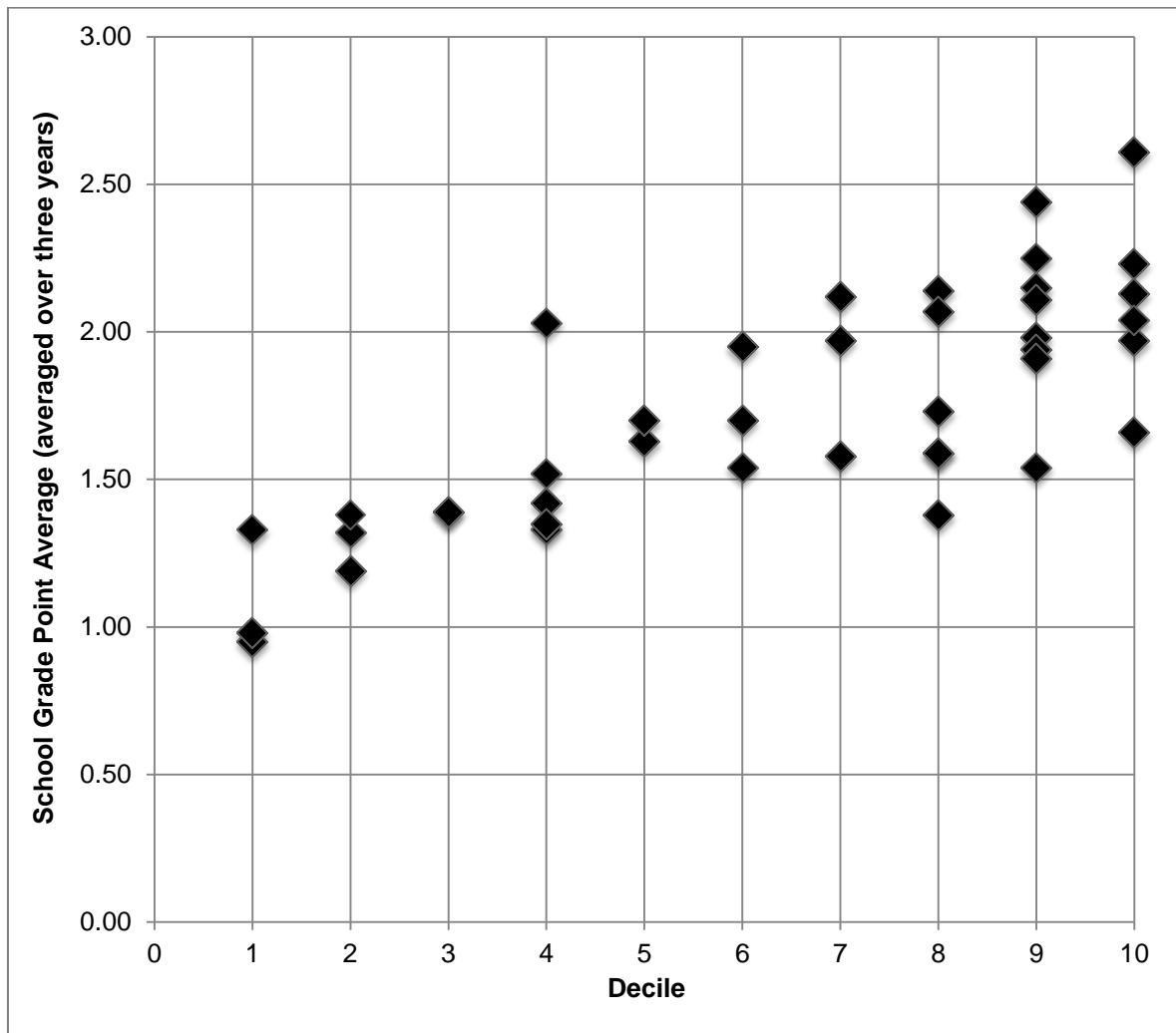
Figure 4.3. Grade Point Average, calculated from NCEA Level 1 English, mathematics and science results averaged over 2008, 2009 and 2010 related to decile of school (*n* = 41).

Figure 4.4 shows the number of schools in this sample that underperform or overperform compared to the national norm within each decile range. The central axis of 0.00 describes the average GPA for the school decile and is a benchmark for comparison. The graph results show how each of the schools perform above or below the national results for decile. Schools 10E (-0.57), 9G (-0.45) and 8F (-0.54)

underperform for their decile in English, mathematics and science for 15-year-olds and schools 4B (+0.49), 9E (+0.45) and 10D (+0.38) overperform compared to the national averages for their decile. This graph describes the variance of performance of schools within the same decile, as compared to the national average data. For example, there is a 0.95 difference between the performance level for 15-year-olds in English, mathematics and science between school 10D (top performer) and 10E (low performer). In the decile 9 schools, there is a 0.90 difference between school 9E and 9G. Even in the lower decile 4 schools, there is a difference of 0.70 between schools 4B and 4D. Four out of the six decile 10 schools underperform in this measure, one performs the same as the national norm and one decile 10 overperforms. There are eight decile 9 schools in the sample and four overperform and four underperform compared to the national norm for decile. In the lower decile schools in this sample, one school (1A) overperformed marginally compared to national norms. There are 27 schools in the sample of 41 that have underperformed in this measure when averaged over three years. Thirteen schools have performed above expectation and one school (10A) has performed at expectation.
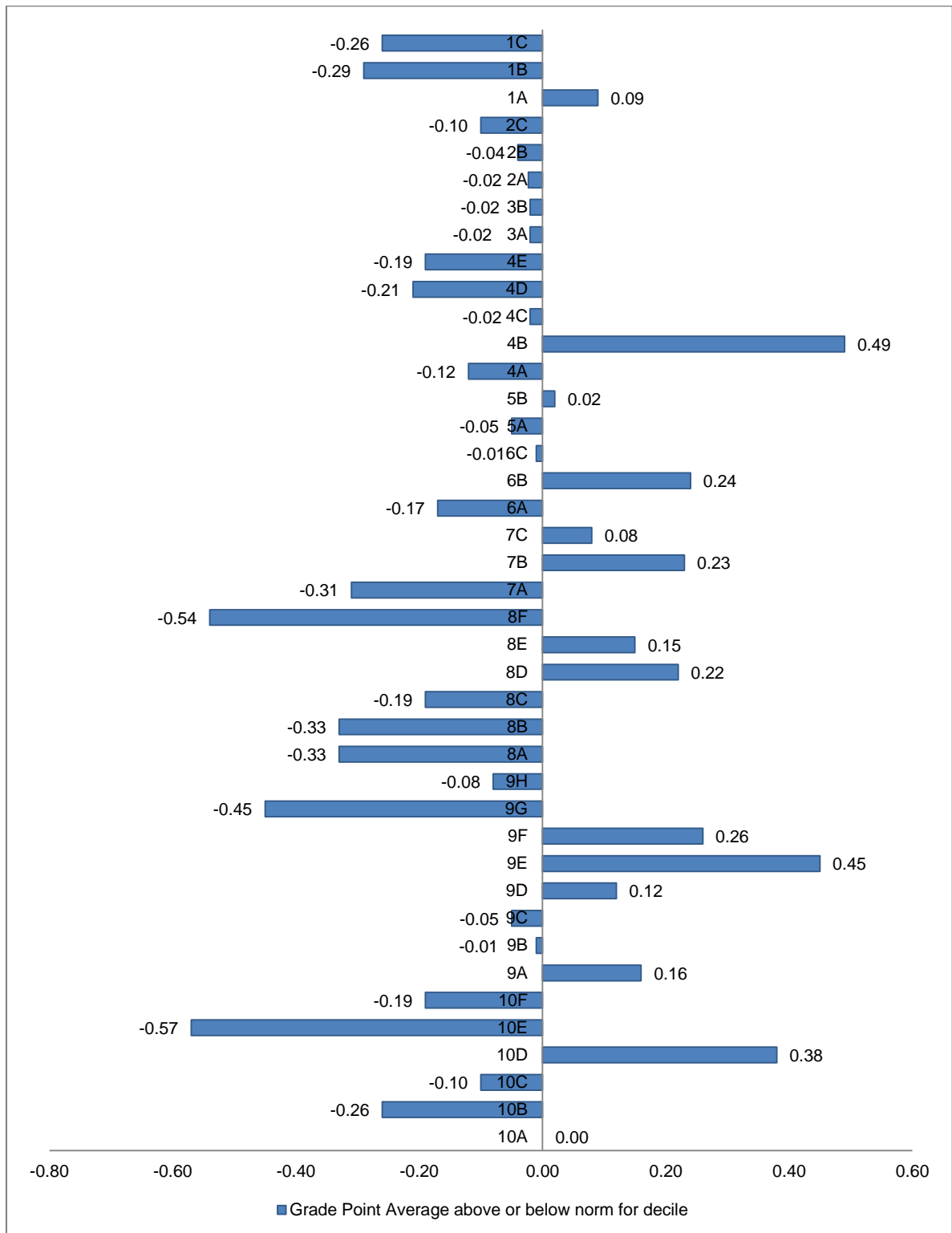
Figure 4.4. Difference between school average GPA in English, mathematics and science over three years compared to national norm for the decile

Statistically significant differences between subjects and school decile

A series of analyses of variance (ANOVAs) were conducted to assess whether there were any statistically significant differences between the decile of a school in the sample of 41 and the performance of students in Level 1 NCEA in English, mathematics and science. In this case it was appropriate to use ANOVA as opposed to *t*-tests because it can assess the effects of two or more independent variables simultaneously (Glass & Hopkins, 1984). The variables were decile of school from 1–10 and subject performance in English, mathematics and science. Subject performance was measured by using a GPA for student academic performance in Level 1 NCEA achievement standards, averaged from three years of results. Tables 5, 6 and 7 that show the results in this section are produced to show the statistically significant differences between the high- and low-decile schools because there was no statistical significance for either of the three subjects when comparing the decile 6, 7, 8, 9 and 10 schools with each other or the decile 1, 2, 3, 4 or 5 schools with each other.

Differences in the academic performance of students between English, mathematics and science departments were explored in relation to the decile of the school. A post hoc Tukey HSD test determined where the statistically significant differences lay (Table 4.1). There were statistically significant differences between English department academic results ($F(9, 41) = 10.18$, $p = .001$). The post hoc test revealed a pattern of statistically significant difference between decile 1 and 2 schools and decile 8 schools ($p = .003$), decile 9 schools ($p = .001$) and decile 10 schools ($p = .001$) for results in English for 15-year-olds at Level 1 NCEA. There was no statistically significant difference between decile 3 and 8 schools but a difference was seen between decile 3 and decile 9 schools in subject English ($p = .002$), and decile 10 schools ($p = .003$). There was a statistically significant difference between decile 4 schools in English and decile 9 ($p = .001$) and decile 10 ($p = .001$) schools. There was no statistically significant difference amongst decile 5, 6 and 7 schools in how their students performed in English at Level1 NCEA when averaged over three years of results.

Table 4.1

Statistically Significant Differences in Results for Level 1 NCEA English When Comparing Decile 1–5 Schools With Decile 6–10

School decile	6	7	8	9	10
1	.146	.14	.003	.001	.001
2	.249	.129	.003	.001	.001
3	.891	.739	.114	.002	.003
4	.976	.880	.110	.001	.001
5	1.00	1.00	.965	.223	.245

Table 4.2 shows that there were statistically significant differences between mathematics department academic results in the sample ($F(9, 41) = 8.27$, $p = .001$). The post hoc test revealed the differences between decile 1 and decile 6, 7, 8, 9 and 10 schools. There was a statistically significant difference between decile 1 and 6 ($p = .03$), 1 and 7 ($p = .1$), 1 and 8 ($p = .001$), 1 and 9 ($p = .001$) and 1 and 10 ($p = .001$) schools. The statistically significant difference for the decile 2 schools is between the decile 7 ($p = .03$), 9 ($p = .004$) and 10 ($p = .001$) schools. There was a statistically significant difference between the decile 3 and decile 9 ($p = .03$) and 10 ($p = .006$) school mathematics department student academic results. There was just one statistically significant difference for decile 4 schools in mathematics compared to decile 10 schools ($p = .03$). In this sample of mathematics departments there is no statistically significant difference between decile 2 and decile 1, 3, 4, 5, 6, and 8 school mathematics departments. There were no statistically significant differences between student mathematics academic results between decile 5 and 6 schools and any of the other deciles.

Table 4.2

Statistically Significant Differences in Results for Level 1 NCEA Mathematics When Comparing Decile 1–5 Schools With Decile 6–10 Schools

School decile	6	7	8	9	10
1	.027	.000	.001	.000	.000
2	.595	.035	.073	.004	.001
3	.834	.117	.243	.031	.006
4	1.000	.463	.774	.150	.027
5	1.000	.816	.976	.599	.230

Table 4.3 shows that there were statistically significant differences between science department academic results in the sample between schools for decile ($F(9, 41) = 6.46$ $p = .001$). The post hoc test revealed statistically significant differences between decile 1 and deciles 6, 7, 8, 9 and 10 schools. The results were similar to mathematics as

there were statistically significant differences in student academic performance in science between decile 1 and 6 ($p = .04$), 1 and 7 ($p = .005$), 1 and 8 ($p = .02$), 1 and 9 ($p = .001$) and 1 and 10 ($p = .001$) schools. The statistically significant differences between the decile 2 schools are with the decile 9 ($p = .04$) and 10 ($p = .008$) schools. There is just one statistically significant difference for decile 3 schools in science and that is with decile 10 schools ($p = .02$). In decile 4 schools the statistically significant difference in science is with decile 9 ($p = .03$) and 10 schools ($p = .005$).There was no statistically significant difference between the decile 5 and 6 schools and the academic results at Level 1 NCEA of science departments in other decile schools.

Overall this post hoc test is showing that school decile makes a difference in some subjects more than others. There are more statistically significant differences in mathematics and science between schools with different deciles than there are in English. Overall, if a student attends a decile 5, 6 or 7 school they have a statistically significant chance of getting similar results in English, mathematics and science in Level 1 NCEA to their peers in higher or lower decile schools.

Table 4.3

Statistically Significant Differences in Results for Level 1 NCEA Science When Comparing Decile 1–5 Schools With Decile 6–10 Schools

School decile	6	7	8	9	10
1	.027	.000	.001	.000	.000
2	.595	.035	.073	.004	.001
3	.834	.117	.243	.031	.006
4	1.000	.463	.774	.150	.027
5	1.000	.816	.976	.599	.230

Standard deviation and effect-size calculation

In order to test the hypothesis that there is within-school variation in student outcomes at department level, the effect size for the difference in academic student outcomes between the English, mathematics and science departments within each school was calculated. Using NCEA Level 1 student achievement data calculated as a GPA for each department, the standard deviation was calculated by identifying the mean of the whole sample and the mean variation for each subject for the data set across the 41 schools. The national results were not included as part of the means. The maximum difference between the highest and the lowest performing departments within each school across the three departments was calculated. For example, schools showed one of the following six combinations: English high – Science low, Maths high – English low,

Science high – Maths low, Maths high – Science low, Maths low – Science high, English high – Maths low.

The effect-size calculation (Cohen's *d*) provides an expression of the magnitude of the difference made to student outcomes by the variation in achievement of the English, mathematics and science departments. An effect size of 1.0 indicates an increase of one standard deviation, typically associated with advancing children's achievement by one year (Hattie, 1992, p. 3). The use of effect sizes in this project highlights the importance of the magnitude of differences, which in this case is between academic outcomes for the same cohort of students within one year in secondary schools. An effect size can be calculated for each school to show the extent of the within-school variation over time but the exact departments where the differences occur cannot be identified by viewing this graph.

When the effect size was calculated over the three years of 2008, 2009 and 2010, six schools out of the 41 had an average effect size less than 0.50, a further 14 had an effect size between 0.50 and 1.0, 18 schools had an effect size over 1.0 and three schools had an effect size over 2.0. These results can be seen in Figure 4.5. The variation in student outcomes between English, mathematics and science for these schools is considerable and the problem has not declined over the three years. The effect sizes for the national data for each of the deciles is considerably lower than the individual school effect sizes. Some schools (*n* = 9) did incrementally reduce the effect size over three years. In six schools, the variation incrementally increased and in the rest of the sample it was variable. The effect sizes for the decile 1, 3, 9 and 10 school national results were between 0.5 and 0.1 across three years but the other six national averages were below 0.5, which would suggest that there is limited variation across subjects nationally but the within-school variation is inconsistent. These data mirror the 2003 PISA results for New Zealand 15-year-olds in mathematics, where 90% of the variance was within school and only 20% was across schools (OECD, 2003).

Figure 4.5 ranks the schools in order of highest to lowest to show the effect sizes or degree of within-school variation across the sample in 2008, 2009 and 2010. The result shows a significantly different distribution than the ranking of average combined school performance in English, mathematics and science described in Figure 4.1 (p. 62). The decile of a school does not determine the schools' position on the bar graph. The top-ranked school for GPA (10D) is the third least likely school to have within-school variation. School 8D, which is ranked as having the third highest within-school variation in the sample, is an academically high-performing school, ranked sixth out of the 41 schools. The lowest ranked school for student achievement (1B) is ranked fourth out of

41 for within-school variation. The national results rank in the bottom half of the graph because the results across all three subjects when compared nationally are relatively even. Thirty schools in the sample have larger within-school variation between the three English, mathematics and science departments than any of the calculations for the national norms for decile. These results show that schools with consistently high levels of student achievement at Level 1 NCEA in either English, mathematics and science can also have high within-school variance and the reverse is also evident.
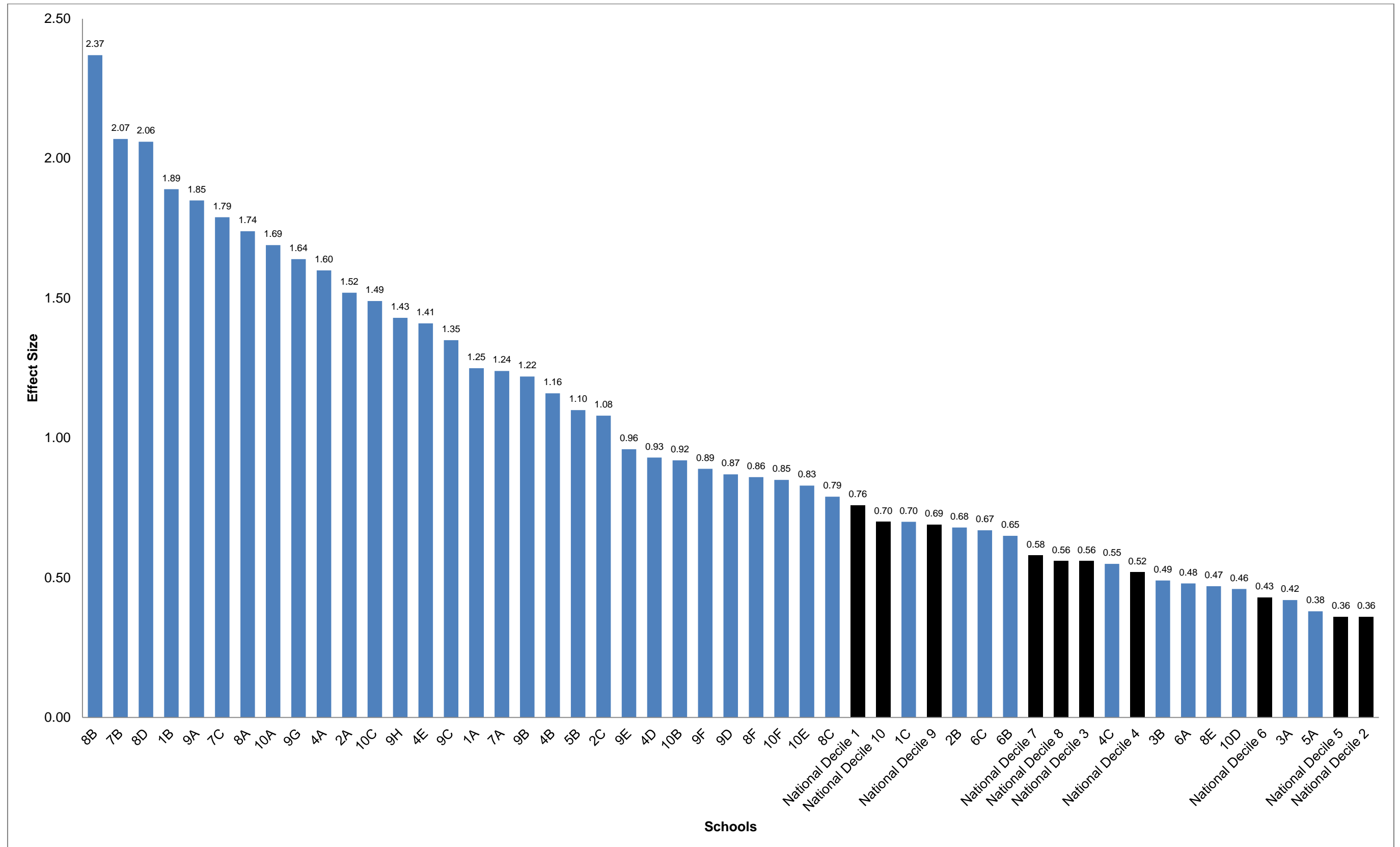
Figure 4.5: Within-school variation shown by effect size calculated and averaged over 2008, 2009, 2010 across English, mathematics and science for 15-year-olds

Figure 4.6 describes both the degree of within-school variation and the overall achievement level of a school at Level 1 in English, mathematics and science by measuring the average GPA. Figure 4.6 shows that some schools have a high GPA (red bar), showing the school is performing academically well in relation to others, while some also have a large effect size (blue bar) showing that there is considerable variation between the highest and lowest performing departments. Ideally, schools would be achieving a high GPA (red bar) and a low effect size (blue bar), such as school 10D. In some schools, the GPA and the effect size are almost identical, such as 8D and 7B and in some schools (9G, 8B, 2A, 1B) the effect size is larger than the total GPA.
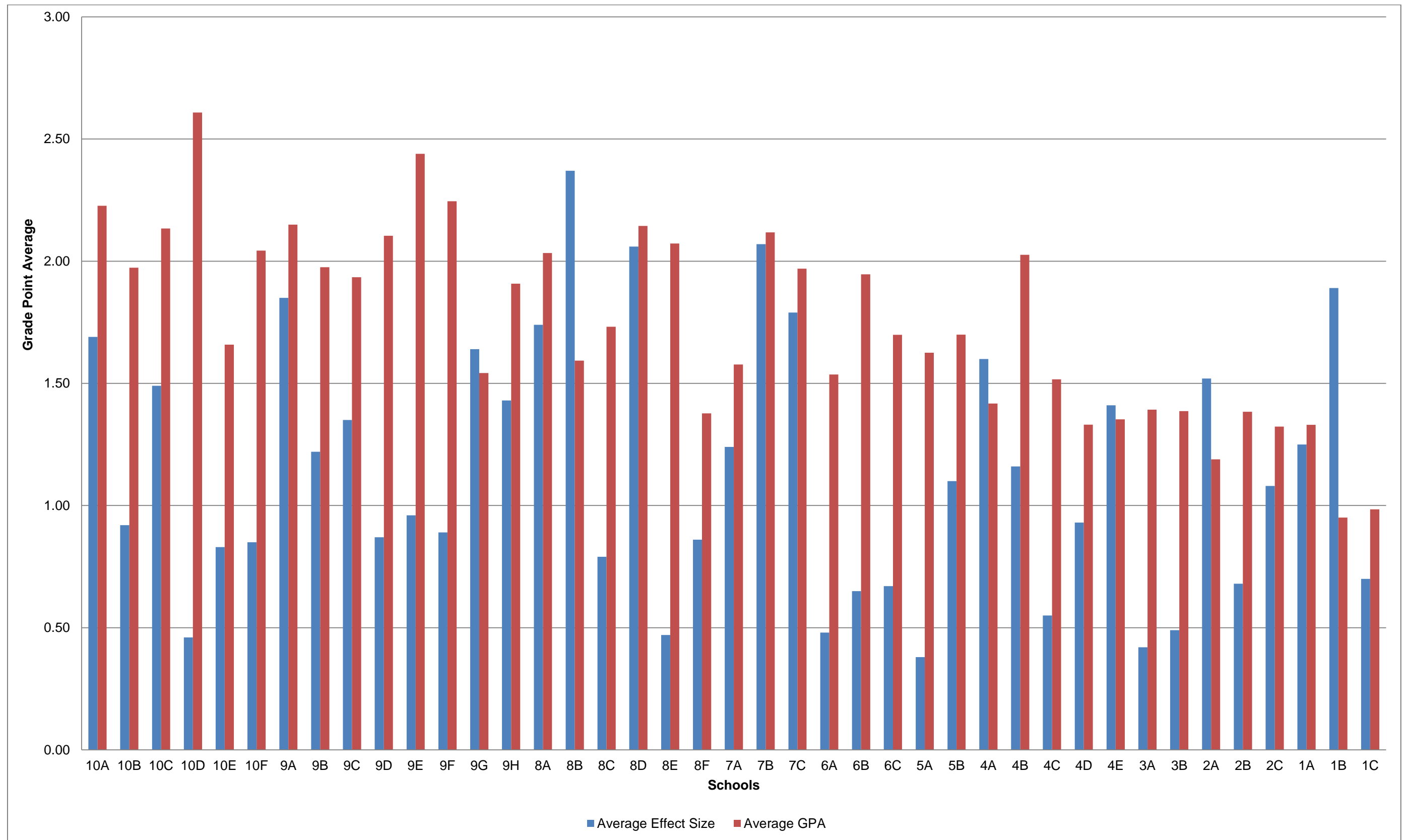
Figure 4.6. Grade Point Average and effect size for each school averaged over 2008, 2009 and 2010

Table 4.4 describes the subject average GPA results for the schools with the largest effect sizes, ranked in the top five. These schools all have an effect size over 1.85 when averaged over 2008, 2009 and 2010. There is one decile 9, two decile 8, one decile 7 and one decile 1 school in this group. The table shows the extent of the variance between departments within schools and the similarity in results across decile despite the difference in socioeconomic disadvantage. For example, the results for the students in the English department in school 1B are comparable with students in the science department in school 8B where students would be considerably more socioeconomically advantaged. Schools 9A, 8D and 7B have an overall high GPA for the whole school which is masking the within-school variation.

Table 4.4

Results for Schools With the Largest Effect Sizes (Over 1.85) Over Three Years

School	Average GPA of highest performing department		Average GPA of lowest performing department		Average GPA for school
9A	English	(2.2)	Science	(1.83)	2.15
8B	English	(1.95)	Science	(1.29)	1.59
8D	Maths	(2.57)	English	(1.90)	2.13
7B	Mathematics	(2.68)	English	(1.83)	2.12
1B	English	(1.28)	Mathematics	(0.66)	0.95

Comparing the percentage of Merit and Excellence grades in higher and lower decile schools

The following graphs (Figure 4.7 and 4.8) illustrate the extent of the difference between the percentage of Merit and Excellence grades achieved by students in the same school in English, mathematics and science across the 3-year period of the study. These higher grades are used as a measure of quality teaching and good academic student outcomes within departments and schools. Figure 4.2 (p. 64) shows the overall performance of all of the schools in the sample for this measure but in this analysis the department results for the higher and lower decile schools have been grouped for comparative purposes. Figure 4.7 shows the variation in achievement using this measure between the English, mathematics and science departments within each of the decile 8 and 9 schools, and compared to national norms for the decile. Figure 4.8 shows the same results for deciles 3, 4 and 5 schools. The graphs show considerable variation between departments in some schools, with 12 out of the 23 schools having more than a 10% difference between the highest and lowest performing department in the school. Eleven schools had less than a 10% department performance gap. Decile 9

schools were most likely to have a variation in department performance for this measure. These graphs also show how department and whole-school performance compares to national norms for the decile. Figure 4.7 shows that schools 8F, 9G and 9H do not perform at or above national norms in any of the three subjects English, mathematics and science. Schools 9A, 9E, 9F, 8D and 8E perform consistently above the national norms in all departments for this measure.
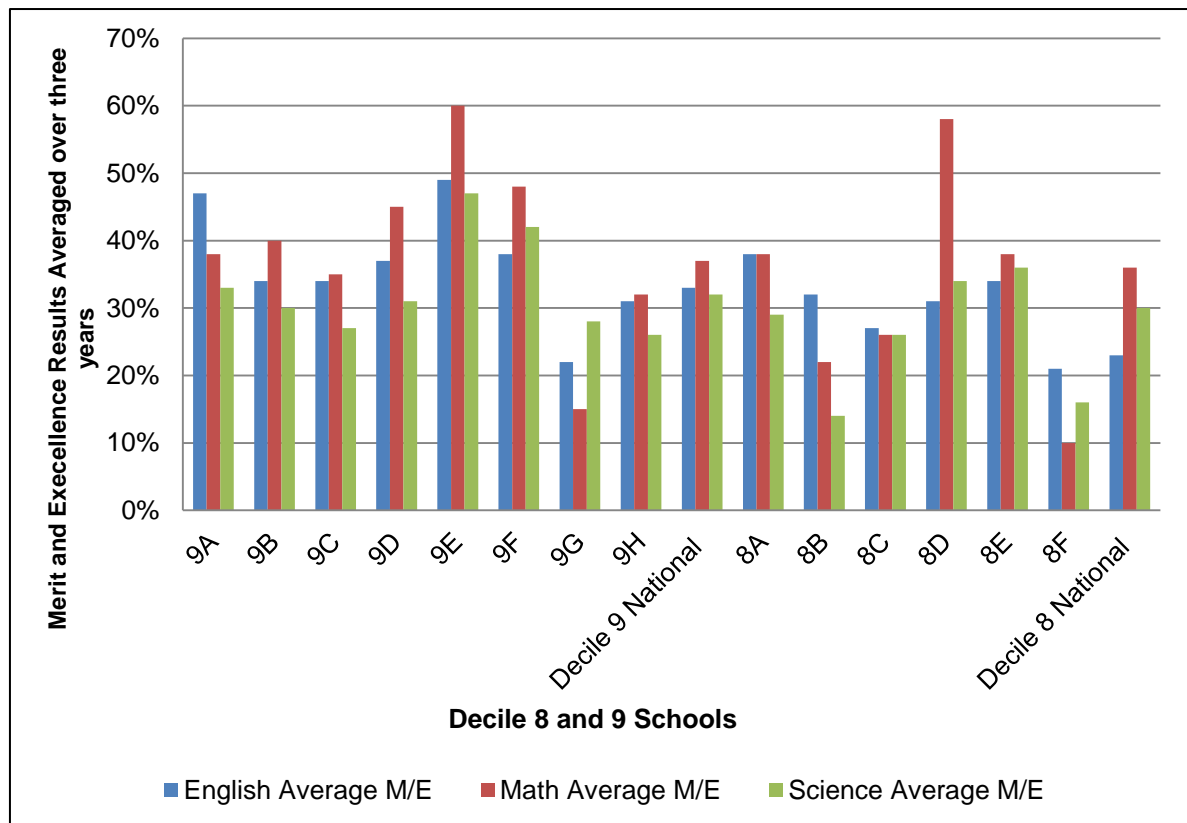
Figure 4.7. Level 1 NCEA Merit and Excellence grades averaged over 2008, 2009 and 2010 as a percentage for English, mathematics and science in Decile 8 and 9 schools, compared to national results for decile

When comparing the results of schools in Figure 4.8 with the results of schools in Figure 4.7 the overall pattern is that over half the departments in the higher decile schools are getting over 30% of students achieving a Merit or Excellence grade. Twenty-eight out of 42 higher decile departments achieved this as opposed to two departments out 27 in the decile 3–5 group. The results also show that some of the lower decile schools have students performing academically at a higher level than some higher decile schools at overall school level and in individual subjects. This is exemplified in school 4B where over 30% of students can expect to gain a Merit or Excellence grade in English, mathematics or science at Level 1 NCEA. This is a greater percentage than for students at some of the higher decile schools such as 9G, 8C and 8F.
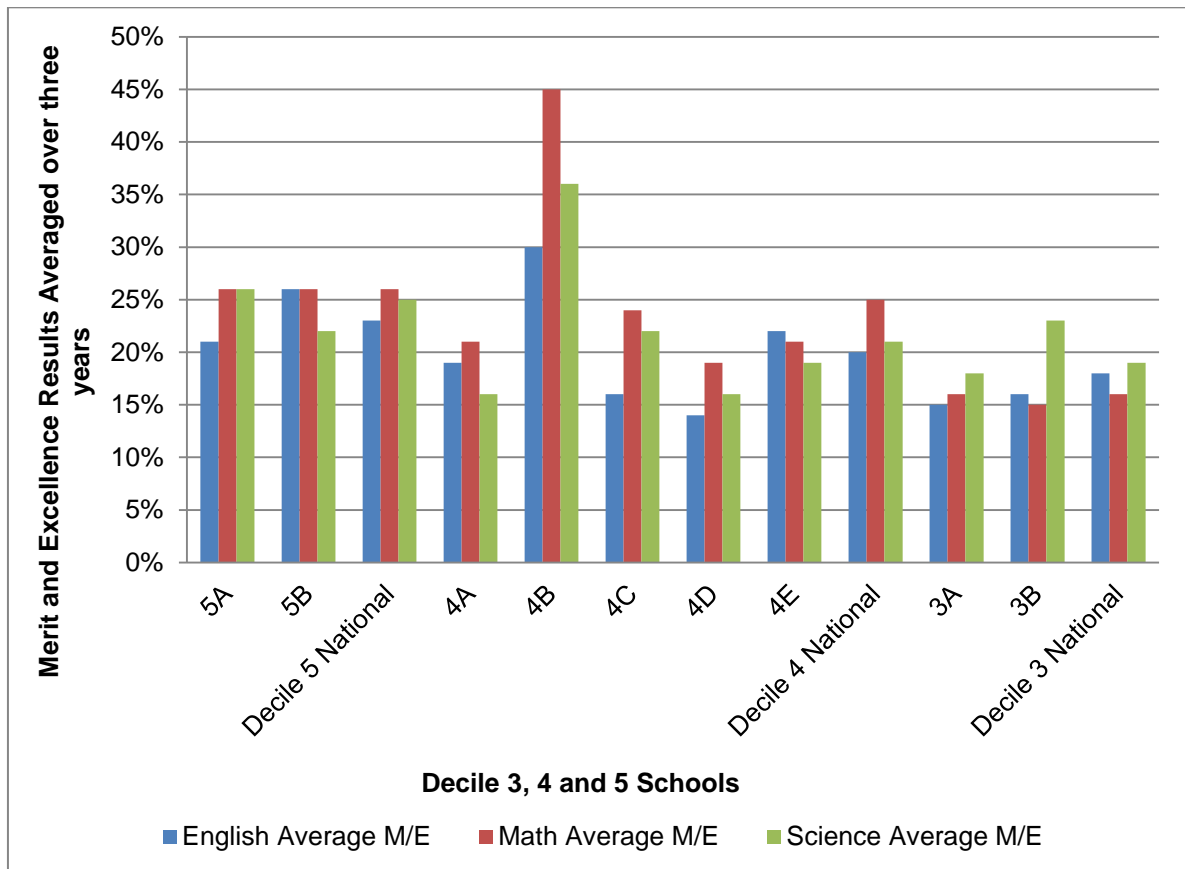
Figure 4.8. Level 1 NCEA Merit and Excellence grades averaged over 2008, 2009 and 2010 as a percentage for English, mathematics and science in decile 3, 4 and 5 schools, compared to national results for decile

Subject variation within and across schools

In order to establish which subjects across the sample performed consistently higher or lower the GPA for each department was calculated in the years 2008, 2009 and 2010. Patterns for consistently higher or lower performance were identified across the sample of 41 schools.

When the results are compared by department across the whole sample over three years, English was the highest performer in seven schools and the lowest performer in five. Mathematics was the highest performer in eight schools and the lowest in three schools. Science was the lowest performer in 10 schools and the highest performing department in two schools out of the 41 in the sample. Seventeen (42%) schools in the sample have the same subject as the highest performer at Level 1 for 15-year-olds and in 16 schools the same subject was the lowest performer (39%). In 10 (24%) schools, the highest and lowest performing department were the same for each of the three years and in 16 (39%) schools there was no pattern of performance. The department

GPA for each of the three years indicates that 31 (76%) of the schools in the sample have within-school variation that is also variable across three years.

Gender differences

The whole-school effect needs to be considered in relation to the gender of the students that attend the school. The sample of 41 includes 32 co-educational schools, four single-sex boys' schools and five single-sex girls' schools. The highest ranked schools for student achievement measured by GPA (10D and 9E) are single-sex girls' schools. Of all the decile 9 schools, the single-sex girls' schools rank higher than the co-educational schools in the category. School 8F that is ranked very low for its decile is a single-sex boys' school but the fourth ranked school (9F) is also a single-sex boys' school. In PISA 2009, girls outperformed boys in reading in every participating country and among the top- and high-performing countries, New Zealand had one of the largest differences between girls and boys achievement at age 15. In PISA 2009, New Zealand girls and boys achieved a similar mean for mathematical and science literacy performance (Telford & May, 2010). Table 4.5 shows the overall GPA for the three years for English, mathematics and science. The three top-performing highest decile single-sex boys' schools in this sample underperform compared to the girls' schools in English, but 9F ranks similarly or better than the decile 9 girls' schools in mathematics and science. The New Zealand school leaver statistics show that in 2009, 86.4% of all school leavers achieved NCEA Level 1 in literacy and numeracy. Female school leavers achieved the required literacy and numeracy credits at a higher rate (88.3%) than their male counterparts (84.5%) (Ministry of Education New Zealand, 2010a). The results of this current study reflect both the national and international statistics when considering gender differences for academic achievement in the senior secondary school.

Table 4.5

Overall GPA for English Matematics and Science Averaged Over Three Years (2008,2009,2010) for Single-Sex Girls' and Boys' Schools

School code and type	English GPA	Mathema tics GPA	Science GPA
10D	2.60	2.67	2.55
9A	2.42	2.14	1.89
9D	2.13	2.23	1.95
9E	2.42	2.62	2.28
9F	2.17	2.37	2.20
8A	2.15	2.17	1.78
8F	1.51	1.31	1.31
5B	1.84	1.72	1.54
4D	1.27	1.51	1.21

Note. Girls' schools are bold and boys' schools are italicised.

Discussion

The results of this analysis substantiate the 2003 and 2009 PISA results for New Zealand in terms of within-school variance (OECD, 2000, 2003, 2010a, 2010b). The PISA results present a broad and general picture of student academic achievement and the influence of socioeconomic factors at a national level. The results of this study uncover the extent of variance within and across this sample of urban Auckland schools. Socioeconomic indicators do not appear to be a strong determinant of within-school variation. The 2009 PISA results tell us that there are some New Zealand students who are performing at the highest proficiency level in mathematics (19%), which is above the OECD average (OECD, 2010b). This analysis shows that these students could be taught in high-performing maths departments, not necessarily in high-decile schools. Students that are lower performing in maths (2.3 %) according to PISA, could be taught in higher decile schools (Ministry of Education New Zealand, 2009b).

Socioeconomic indicators for academic achievement appear to predict an overall trend but variation occurs within the whole-school and within-school performance. The academic achievements of diverse learners in standards-based assessment systems have been mixed (Agnew, 2010). The research in America indicates that although minority students and those from low-socioeconomic backgrounds perform better in a standards-based system, there is still a significant gap with their middle class peers (Kannapel, Aargard, Coe, & Reeves, 2001; Madaus & Clarke, 2001; Ortiz, 2000). This study would generally support the literature but Figure 4.7 and 4.8 show that the difference in student achievement results for schools drawing on a cohort of students from a similar socioeconomic background is variable by department. Hattie (2002)

argues that the pervasiveness of the decile system as an indicator of school success, of teacher quality, of student achievement and of quality of the school is demeaning to the many teachers and students who excel in these schools (Hattie, 2002). The results of this study and analysis of student academic outcomes supports this statement. The simple rankings of schools in Figure 4.2 and 4.3 show that there is no doubting the power of socioeconomic status as a predictor of student success as the majority of the higher decile schools are ranked in the top half of both GPA and the percentage of Merit and Excellence grade indicators of success.

The relationship between decile and subject is also variable. The analysis of variance conducted between subjects and school decile show that generally the difference in results is between the decile 1, 2, 3 and 4 schools and the 8, 9 and 10 schools. In English there was no statistically significant difference between decile 5, 6 and 7 schools in Level 1 NCEA results. School deciles are differentially effective as a predictor of success in English, mathematics and science. In this sample of mathematics departments there was no statistically significant difference between the results of decile 2 and decile 1, 3, 4, 5, 6, and 8 school mathematics departments. There would be no difference in studying mathematics or science at a decile 5 or 6 school in this study. These distinct anomalies support Hattie's claims that decile is a crude indicator of school quality, with little evidence to support claims of teacher quality (Hattie, 2002).

Further evidence to support Hattie's claims are revealed in this study when considering the extent of the difference between the percentage of Merit and Excellence grades achieved by students in the higher and lower decile schools. There was some similarity in the results of some of the English, maths and science departments in the decile 8 and 9 schools compared with those in the lower decile 4, 5 and 6 schools indicating that lower decile schools can and do create the conditions for students to gain good academic outcomes (Figure 4.7 and 4.8). If the number of Merit and Excellence grades achieved by students is used as an as an indicator of quality teaching in the department, the considerable variation of teaching within and between schools that supports a high level of academic achievement is clearly implied in this study. There are three decile 8 and four decile 9 schools underperforming in relation to the national results averaged over three years (Figure 4.1, p. 62). These results concur with a recent study by Bendikson, Hattie and Robinson (2011) who, in a study of 102 secondary schools in New Zealand, found that neither high- nor low-decile schools were immune from lower performance levels and that schools regardless of their socioeconomic status or decile can impact positively on student academic outcomes. Understanding how the leadership and teaching practices in a department contribute to this increased

level of student achievement will reveal the specific practices that make a difference to the within-school variance problem.

The results of this study reflect other similar studies that show whole-school results mask within-school variation. This is consistent with the views of Thomas, Sammons, Mortimore and Smees (1997b) when they compared the GCSE results of 94 inner London secondary schools over three years (1990, 1991, 1992). They focussed on establishing whether school performance was both consistent over a range of six subjects and stable over time, across cohorts. They used a predefined set of criteria to identify groups of schools as being "effective" or "ineffective." The results are similar to this study, in that in the English study there was no clear cut picture in 70% of the schools, where there were mixed effects. There was marked within-school variation in student academic results but the high performance in some subjects and low performance in others masked the overall school results. Over three years they were able to classify just 9% of the schools in their study as "broadly more or less effective" (Thomas, Sammons, Mortimore, & Smees, 1997a, p. 191). If a similar measure was applied to this study, when considering Figure 4.6 that details the within-school variation and the total GPA, few schools have a high GPA and low effect size or within-school variation measure. Fifteen (36.5%) of the 41 schools have performed with a GPA of more than 0.2 above expectation for their decile; however, only six of these 15 schools have an effect size less than 1.0. This would show that 14.6% of these large urban secondary schools were consistently effective across English, mathematics and science over a 3-year period. The remaining nine schools are unquestionably high performing but students could expect differentiated results between the three departments.

The results of this study show that although the national results for each of English, mathematics and science show little variance, the variance of student outcomes across subjects within schools is considerable. This would support both the 2003 and 2006 and 2009 PISA findings for 15-year-olds in New Zealand. The variance in results for the same subject across schools indicates that student achievement may not be determined by the school a student attends but the department in which they have been taught. These results indicate that in some schools the 'department' in which the student is taught can greatly enhance or reduce chances of academic success. In this study, there is whole-school performance that is considerably above the national norm for the decile. For example, schools 9E and 9F perform above the national norm for decile 10 schools and schools 8D, 7B, 8E, 8A and 4B all perform above the norm for decile 9 schools. The converse is also apparent in Figure 4.1 (p. 62). School 8F

performs at a lower level than the national norms for decile 2 and 3 schools and school 10E, 8B, 9G and 7A all perform at a lower level than the decile 5 norm.

It is apparent that science is the lowest performing subject in the sample schools. There are also relatively few students studying science in decile 1 schools. This raises the question of why some subjects seem to have consistently lower results than others. Fitz-Gibbon and Vincent (1997) assert that more able students are attracted into mathematics and science subjects, so this does not explain the substantial difference of performance for these departments in this sample because these students were being assessed in all three subjects. Eight English departments were the consistently highest performer over three years in a multiracial city where for a proportion of students English is not the first language spoken at home.

When departments or subjects perform differently within a school, there are a number of arguments set out in the literature that link the variation in subject results to a range of other factors such as differences in 'subject difficulty.' In addition, more students achieving higher grades could be interpreted as the result of better teaching, more effort on the part of the students or use of private tutors. The results could also be skewed or accounted for by easier examinations and easier marking (Tymms & Fitz-Gibbon, 2001). The argument for lack of comparability within a subject and across years is mitigated in this project by generating data for national norms in each subject by decile and averaging results over three years. This has allowed comparison of subject performance within a school with subject performance in other similar schools. Fitz-Gibbon and Vincent (1997) argue that statistics alone cannot distinguish between unmotivated students, poor teaching, severely graded subjects or simply intrinsically difficult subjects. When the patterns are consistent rather than varying from school to school, the place to look for an explanation is in the subjects rather than in the schools. Some subjects are severely graded or allowed to be 'difficult' (Fitz-Gibbon & Vincent, 1997, p. 296). There has also been considerable debate about the comparability of different subjects in public examinations in England. Fitz-Gibbon and Vincent (1997) would support the claim that mathematics is 'harder' and therefore cannot be compared objectively with other subjects. This argument could contribute to the findings of this study where science was the lowest performer over three years in 10 schools out of 41. Perhaps science is just 'harder.' There was however a tendency for schools to have a pattern of one low-performing department despite the subject. The analysis shows that 42% of the sample had the same subject as the highest performer and 39% had the same subject as the lowest performer. In 24% of the schools, the highest and lowest

performing department were the same for each of the three years, which would indicate that subject difficulty may not be the cause of consistent underperformance.

Subject English was consistently the highest performer over three years in 7 out of 41 schools and was a top-performing department 43 times. This is an interesting result because the schools in urban Auckland have 32,079 Ministry of Education specifically targeted students who have English as a second language (ESOL). There are 164 ethnicities represented in the Auckland schools with 109 different languages spoken at home. Despite this, English being a top performer is consistent with the PISA results for 15-year-olds in reading (OECD, 2010b), and New Zealand students are also high performers in mathematics and science literacy (Telford & May, 2010). Mathematics is the consistently highest performing subject in eight schools and the lowest in three schools. Mathematics is the top-performing department 51 times in the 41 schools in either 2008, 2009 or 2010. For eight out of the 41 schools (19.5%) in this sample, mathematics was consistently the highest performing department in the school over the three years. Seventeen schools had the same subject as the consistently highest performer and 16 schools had the same subject as the consistently lowest performer over three years. Ten schools had mixed results across all three departments across three years and there was no consistently high or lower performing department.

The results in the low-decile schools of this study support the statement that Māori and Pasifika students are over represented in the statistics for students performing below expectations (McNaughton et al., 2011). Strathdee (2003) and Philips (2003) argue that one of the policy aims of the development of the NCEA was to remove barriers for low achievers and Māori and Pacific Island students and Rawlins et al. (2005) assert that standards-based assessment potentially provides schools with greater opportunity to adapt assessment tasks to meet the needs of diverse learners, while still assessing the set standards (Hipkins, Vaughan, Beals, & Ferral, 2004; Rawlins et al., 2005). Whatever the assessment method, the 2009 PISA findings clearly show the performance differences within and between schools (OECD, 2010a) and reveal the extent to which differences in performance between schools and among students within schools can be attributed to differences in socioeconomic background between and within schools (OECD, 2010a, p. 86). Across the 34 participating OECD countries, 14% of the variance in achievement is attributed to socioeconomic background of students, and in New Zealand the figure is 19% (OECD, 2010a). This is evidenced in the PISA statistics that reveal 35% of Pasifika and 24% of Māori students did not show reading proficiency above Level 1a. This compared to 15% Asian and 9% Pākehā/European (Telford & May, 2010). The two lowest ranked schools across the three years in this sample are

schools 1C and 1B, both serving large populations of economically deprived students. School 1C has 64% Pasifika students and 14% Māori students. School 1B has 76% Pasifika students and 14% Māori students (ERO, 2011b). School 1C had an overall GPA over the three years in English, mathematics and science of 0.97 and 1B had 1.05. The third and fourth lowest ranked schools (2A and 1A) had average GPAs of 1.24 and 1.27. These schools similarly serve communities with large Māori and Pasifika students. School 4D is ranked fourth lowest and is a single-sex boys' school. The ethnicities of the students are 18% Māori, 22% Pākehā/European and 45% Pasifika students (ERO, 2011b). School 4B, which has the same socioeconomic grading and is ranked overall 15th from the top of the sample, is a co-educational school with students from 70 nationalities. The student ethnicities are represented as 10% Pākehā/European, 5% Māori, 30% Indian, 24% Pasifika and 26% Asian. The socioeconomic status of these students is similar to the students in the low-ranked school 4D but the ethnic mix of the school is substantially different.

There does appear to be a whole-school effect in some schools where there is overall high performance across departments across three years but some schools exhibit inconsistency in results across departments and across years. Bendikson et al. (2011) argue that evidence of secondary school performance highlights the difficulty in ongoing improved performance, particularly over a period of more than three years. Understanding why it is that results can be so varied within a school between departments when they are serving the same student cohort and exist under similar senior leadership and governance conditions is critical to the enhancement and equity of student achievement outcomes in New Zealand. The variation of performance across schools that draw students from similar socioeconomic groups reveals that there could be a number of factors that are impacting at the school level on student outcomes. It could be that further evidence of background characteristics for student intake is needed to account for an accurate measurement of school effects (Sammons, 1996). The analyses of these results reveal some relevant and somewhat concerning disparities in student outcomes within and across the schools in the sample. These results show that the overall academic results for schools mask the disparity within schools and that students' ethnicity and gender needs to be taken into account when considering the value added of a particular department within a school.

This chapter has set out the academic results from the Phase 1 study indicating the complex and varied achievement patterns across the three years. The Phase 1 project has clearly shown considerable within-school variation and subject differences, particularly in relation to socioeconomic demographics, that need closer scrutiny. The

next chapter sets out the process for the development of the questionnaire that will be used to test the relationship between the middle leadership practices at department level and the academic student outcomes for English, mathematics and science.

Chapter Five: Development of the Questionnaire

The purpose of this stage of the study was to investigate the extent to which middle leaders in English, mathematics and science departments carried out leadership practices that have been described in the literature as effective. The second research question inquires about the extent to which the middle leaders and teachers in their departments were able to report their perception of the leadership practices known to be effective, occurring. Questions about how departments function and are led can only be answered by middle leaders and the teachers within their departments (Siskin, 1994). This chapter describes the rationale for the method chosen, the process for designing the questionnaire, conducting the field work, the ethical considerations in administering the questionnaire and the sample of schools who agreed to participate.

Instrument

The rationale to use a questionnaire to investigate the leadership practices of middle leaders in departments was two-fold. The instrument provided an opportunity to efficiently gather information from a large group of people simultaneously, without their ideas and opinions influencing each other. The format and timing for the administration of a hard copy questionnaire also allowed for responses to a wide range of questions concerning department leadership to be asked of all department teaching staff and responses standardised so that they were immediately comparable (Mann, 1998). By administering the questionnaire in person at a department meeting the researcher was able to maximise return rates and ensure participants were well informed about the purpose of the project.

The disadvantages of the written questionnaire format that needed to be mitigated were that they reflect the opinion, needs and views of the individual respondents who respond within a variety of contextual factors at one particular time when the questionnaire is distributed. This method includes no opportunity to probe as might occur in a focus group interview. A range of opinions of individual respondents has been gathered by requiring all department members to comment on identical items within the questionnaire at the same time so they could be checked and compared for variability (Mann, 1998). For these reasons, the questionnaire was developed with careful consideration of the existing literature on department effectiveness, ensuring the presentation and format was logical and trialled and tested with colleagues to make changes and alterations before it was used in the field.

In order to mitigate for variations in perceptions between respondents, the instructions on the front page of the questionnaire were trialled for clarity. The final questionnaire required middle leader participants to indicate how strongly they agreed with statements that described the leadership practices in the department in which they worked, with a special note in bold to explain that their responses related to both their own leadership practices and those of other department leaders. The teacher participants were requested to answer the questions considering the leadership of all the people who held formal leadership positions in the department in which they worked and to indicate the one response that came closest to describing their opinion for each item. At the top of the following pages of the questionnaire was the statement or stem *The leaders of this department ensure that.* This was a key design feature that served to remind participants of how they should be considering their response as they responded to the items on each page.

Developing the items for the questionnaire

Through the development of the literature review for this project it became evident that there was some empirical evidence of the practices of middle leaders that related to student outcomes. Although the research in the field is by no means extensive the work that has been done has been well described and was robust enough to develop some clear descriptors of effectiveness that could be applied to the secondary school context in New Zealand urban schools. *The Middle Leadership Matrix* (Table 2.2 on pages 46–49) describe how the literature in school and middle leadership has been categorised, to determine 24 broad themes or ideas regarding middle leadership behaviours that strongly informed the 34 items that were included in the final questionnaire. Spillane et al. (2001) assert that it is important to identify, analyse and understand the many 'micro' tasks that sit underneath all of the leadership practices in order for educators to understand the how as distinct from the what of school leadership (Spillane et al., 2001). They believe that it is the frequency, quality and deliberate carrying out of the microfunctions that will contribute to support and develop the essential functions. The 24 broad themes were developed into questionnaire items that would describe the recognisable micro tasks of middle leaders in secondary schools.

The development and design of the questionnaire for this project was informed by two recently developed questionnaires for school leaders. Both are 360 degree feedback models whereby teachers and school leaders comment on school leadership behaviours, as opposed to knowledge, dispositions or personal characteristics of leaders (Porter et al., 2008, p. 9). The first, a New Zealand-based tool called the Educational Leadership Practices Survey (ELP) was developed by Wylie and Robinson

in April 2009 (Wylie, 2010) for the purpose of providing formative information for principals about their instructional leadership practices. This survey was drawn from leadership dimensions developed by Robinson et al. (2009) and included items relevant to the leadership of teaching and learning. The items expressed desirable leadership practices and were linked to valued student outcomes (Wylie, 2010). The survey was trialled in a cross section of 36 volunteer schools and finalised to provide scales. The second informative instrument, known as the Vanderbilt Assessment of Leadership in Education scale (VAL-ED) is an evidence-based, multirater rating scale that assesses principals' learning-centred leadership behaviours known to directly influence teachers' performance, and in turn students' learning. The VAL-Ed measures critical learning-centred leadership behaviours for the purposes of diagnostic analyses, performance feedback, progress monitoring and professional development planning (Porter et al., 2008, p. 5). Both these instruments provided suitable and relevant models for wording and organising items cohesively.

In order to check that the questionnaire items covered the components of middle leadership behaviours, the questionnaire items from *Forging Links* (Sammons et al., 1997) were also available for cross examination and comparison. The questionnaire items, based on key aspects of effectiveness surfaced through the case study phase of that project. The heads of department were requested to fill in a hard copy postal survey judging departmental effectiveness, major successes and problems faced, factors contributing to effectiveness and barriers to achievement (Sammons et al., 1997, p. 102). The design of the *Forging Links* study required principals and heads of departments to fill in identical questionnaires so that the researchers were able to ascertain the difference between the views of the principal and the head of department on questions regarding aspects of school policy and the factors in the school that influenced effectiveness. This design feature relates to this current study because although the principals in the 10 schools were not questioned about the practices of their middle leaders, both the teachers and middle leaders in each department filled in identical questionnaires, therefore differences and similarities in views were able to be identified.

Categorising the items and developing the scale

The recent examples of school leadership questionnaires from ELP and VAL-Ed combined with the evidence of middle leadership effectiveness developed through the literature informed the decisions on the wording of the individual items, the structure of the questionnaire and the grouping of the items under broad headings. The questionnaire contained seven general headings under which the items were grouped.

These were *Goals*, *Managing resources*, *Ensuring an orderly and supportive environment*, *Effective teaching and student outcomes, Assessment and data*, *Culture of reflection and inquiry* and *Team approach*. Between four and six items were grouped under these general headings for participants to score the middle leadership practices occurring in their department.

The questionnaire items for both the teachers and middle leaders were identical and were presented with a stem at the top of each page that stated *The leaders of this department ensure that.* Participants were asked to tick one box from the positively packed rating scale (Lam & Klockers, 1982) described as: strongly disagree, mostly disagree, slightly disagree, slightly agree, moderately agree, mostly agree and strongly agree. These descriptors were converted to numbers when entered into SPSS for analysis purposes. A positively packed scale is a skewed response format with more response options on the positive end of the continuum (G. T. L. Brown, 2004). In this questionnaire, the options for respondents included four positive response points and three negative. Brown (2004) states the reason for utilising a positively packed scale as:

> Variance of participants' responses is necessary to measure accurately any psychological construct. It is likely that balanced response anchors will not provide variance when participants are inclined to respond positively to all items because they are deemed equally true or valuable. In other words, the statement being responded to may be so socially accepted it would be difficult to elicit variance in responses because respondents are likely to agree with it or have a generally positive affect toward the psychological construct. (p. 1016)

This argument is particularly relevant when considering the reporting of leadership practices. Middle leaders were more likely to self-report that they frequently encourage staff in this area than never or sometimes.

This leads us to consider the questionnaire response format that asks respondents to identify and rate their own behaviour and practices. The evidence of problems associated with how people respond to questions and how inaccurately their memory is in terms of remembering frequency is detailed in the literature (Oppenheim, 1998; Sudman, Bradburn, & Schwarz, 1996) and Mann (1998) asserts that while self-assessment is an ability that many professionals develop, many would consider that this ability varies widely. She stated that individuals' perceived needs or ideas may reflect an incomplete picture of desired abilities and outcomes and relationships between them. In this project, all participants were required to rate the leadership practices within their department. They were rating a set of practices that they experienced as opposed

to the performance of themselves or an individual. In the design stage of the study, it was anticipated that a separate analysis of teacher and middle leader responses would be carried out, so that comparisons between responses of teachers and their middle leaders could be made.

Trial of the questionnaire

Oppenheim (1998) states that it is essential to pilot every question, its sequence, scale, question layout, instructions and categories. Prior to ethics approval being sought, the questionnaire used in this project was trialled twice, in December 2009 and again in April 2010. The trial was completed with a group of 30 professional development facilitators who had previously held a range of leadership roles in secondary schools. Oppenheim (1998) suggests that respondents in the pilot study should be as similar as possible to those in the main enquiry and that they should be asked to help and give critical feedback. The selected pilot group for this study were asked to complete the draft questionnaire assuming the 'role' of a middle leader. They were able to provide robust and useful feedback about items that did not make sense, language that was not understandable and where they felt there were important aspects of middle leadership not stressed or clearly articulated through the questionnaire. These trial participants answered the questionnaire but also wrote comments on the hard copy where they provided further comments, questions and feedback. It was this feedback that helped develop some of the items around teamwork and collegiality as the trial participants believed these were important behaviours of middle leadership and were not adequately accounted for in the items of the trial questionnaire.

The trial provided important information about the order of the questions. After the trial, the questionnaire was reconstructed so that the demographic questions were asked at the end of the survey, after the most important items related to department leadership. It became apparent in the observation of the trial participants, that respondents were most focussed at the start of the task and were putting their greatest thought and energy into the completion of the demographic questions. The outcome of this reversal of the order of the questions was that when the questionnaire was administered in the field, the demographic data, being filled out last by respondents was completed erratically and the information was often incorrect, therefore limiting its use in the analyses.

An example of the final surveys used in the field for middle leaders and teachers are included as Appendix A and B.

Conducting the field work

Eligible participants were the middle leaders and teachers in the English, mathematics and science departments in the 41 schools selected for the student academic outcomes data analysis. Permission was requested from each of the principals of the 41 schools to administer surveys about the leadership and teaching practices within each of the three departments in the school. Ten principals from the original sample of schools agreed to participate in this phase, a response rate of 24.3%. The response rate was disappointing; however, there were a number of factors for refusal that were indicated by principals who declined their schools' involvement. A large contributing factor was the timing of the requests sent to the principals by the researcher. The Post Primary Teachers Association (PPTA), which is the union representing secondary school teachers, issued a 'work to rule' strike notice to their members at the same time as the researcher approached principals to participate, a risk to the project that could not have been anticipated. Some principals felt that they were unable to request staff to participate in this research project as it would be seen as additional to normal workload. Others stated that despite the PPTA strike notice the school had decided on a research ban due to the high numbers of requests they were receiving from research organisations and individuals. One principal was only prepared to participate if she were able to gain access to the results for the individual departments. This was denied due to The University of Auckland ethics committee recommendations. Ten of the 41 schools formally declined to be involved in the project and notified the researcher either by phoning or sending written confirmation via email. Ten school principals agreed to participate and 21 school principals failed to respond to the letters, emails and phone calls placed over a 6 week period encouraging them to participate. The 10 schools where agreement for participation was gained, represented a range of deciles and school type from the original 41 schools.

Once the principal had agreed to the staff of the English, mathematics and science department in their school participating, the heads of each department were contacted by the researcher and the project and time commitment fully explained to them in writing. There was full agreement by all of the middle leaders and teachers in the three departments in the 10 schools to participate. The questionnaire was administered in person by the researcher in schools between 10th November and 10th December, 2010.

A hard copy questionnaire was distributed to all middle leaders in the English, mathematics and science departments and the teachers that reported to them at a departmental meeting. A full description of the number of participants and percentage of respondents in each of the departments in each of the 10 schools is provided in Table

5.1. All participants were provided with an electronic version of the Participant Information Sheet (PIS) prior to the meeting with the researcher and these were also available in hard copy when the researcher met with the participants. These PISs are available in Appendix C. The questionnaires were completed by participants and then dropped into a sealed box that was later collected by the researcher. The process for dissemination of the questionnaire to middle leaders and teachers in the field can also enhance the validity. McCauley and Moxley (1996) assert that the practice of creating anonymity for raters, particularly those who are likely to have equal or less organisational power than the manager, is important for assuring more honest ratings so the measures taken to protect anonymity and privacy can contribute to the reliability of the data received.

Table 5.1

Survey Return Statistics

School Key	Total Middle Leaders Response	Middle Leader % response rate	Total Teachers Response	Teacher % response rate	Total School returns	% response rate total for school
3A	14/14	100%	20/28	81%	34	92%
5A	12/14	88%	43/65	66%	55	73%
6A	9/11	82%	22/42	52%	31	76%
8C	6/7	89%	13/20	64%	19	77%
8D	7/7	100%	19/32	60%	26	64%
8E	15/15	100%	30/39	83%	45	92%
9A	6/7	83%	9/17	65%	15	74%
9B	6/10	73%	15/27	56%	21	66%
9D	8/8	73%	34/44	83%	42	78%
9E	12/15	80%	27/58	39%	39	57%
Sample Total	95	87%	232	65%	327	75%

Note: Percentages are rounded to the nearest whole number

Ethical considerations

In order to carry out this phase of the project, approvals were obtained from The University of Auckland Human Participant Ethics Committee (refs 2010/303). A number of ethical considerations needed to be considered and mitigated in administering and analysing the results of the questionnaire. Requesting the views and opinions of teachers and middle leaders about the effectiveness of the leadership practices evident and experienced in the department in which they worked on a day-to-day basis needed careful consideration and sensitivity given the closeness of the collegial relationships. The questionnaire was designed to probe the professional knowledge culture and social

cohesion of the department and required participants to make judgements about their own and their department leaders in a way that they may not have been previously required to do. Siskin (1994) commented that teachers are also strongly connected collegially and have a loyalty to department members so it was important to ensure all participants felt assured of individual anonymity and security in relation to the data.

Informed consent, confidentiality and anonymity, benefits and risks to participants

Principals of the 10 schools that agreed to participate signed and returned consent forms. A copy of the principal PIS is attached in Appendix C. Heads of department were then approached in writing via email by the researcher and were provided with a PIS for middle leaders and asked to distribute the PIS to other middle leaders and teachers in their department to inform them of the project and invite them to the meeting where the questionnaires would be distributed. The principal was not informed which staff members participated in the project.

Participants were assured that the data regarding their department would be handled confidentially and not be provided to their principal or employer. The information required on the questionnaire was the school identification code and the subject in which the participant was a middle leader or teacher. Aggregated middle leader and teacher responses were analysed and the results for individual departments and schools were anonymised by the use of school codes. There were sufficient schools involved in this project that in any written publications about the work the schools and departments were not identified.

Data from the questionnaires have been stored electronically and securely on a disk at The University of Auckland for six years after which time they will be destroyed. An electronic copy of codified data from the questionnaires was created and used by the researcher for the duration of the write-up. Coded identifiers for schools, departments and participants known only to the researcher have been used in all publications related to this project and are password protected.

Sample

The schools that agreed to participate represented a range of deciles, school roll size and demographics from across the urban Auckland region. Seven co-educational schools from deciles 3–9 and three single-sex girls' schools agreed to participate. Evidence of within-school variation shown by an effect-size measure that describes the extent of the variation between department performance in English, mathematics and

science compared to national norms with a Level 1 GPA averaged over three years reflects a range of school academic success in these 10 schools. Table 5.2 details the demographic data of the schools and information about student ethnicity and the numbers of students most at risk of not achieving. The two right-hand columns include the percentage of Māori and Pasifika students enrolled in each school with statistics from the school's most recent ERO report (ERO, 2011a, 2011b) and the actual number of students the Ministry of Education specifically funds because they have English as a Second Language (ESOL) or they have refugee status.

Table 5.2

Profile of Participating Schools

School code with decile	Urban region in Auckland	School roll 2009/2010	School type	Level 1 GPA score averaged over three years	Effect-size score averaged over three years	% Māori & Pasifika students 2009/2010	No. of refugees & ESOL funded students**
3A	South	1779	Co-ed	1.39	0.42	35%	330
5A	West	2359	Co-ed	1.63	0.38	38%	187
6A	South	1763	Co-ed	1.54	0.48	25%	31
8C	East	826*	Co-ed	1.73	0.79	28%	82
8D	West	1110	Co-ed	2.14	2.06	17%	24
8E	East	2014	Co-ed	2.07	0.47	12%	55
9A	Central	947*	Integrated single-sex girls'	2.15	1.85	21%	0
9B	East	975*	Co-ed	1.98	1.22	9%	25
9D	North	2088	Single-sex girls'	2.10	0.87	9%	58
9E	Central	2003	Single-sex girls'	2.44	0.96	14%	142

Notes.

*School rolls fluctuate considerably and all schools originally selected in 2007 for the sample had a pupil roll over 1000. In three schools the roll has declined over three years.

**Figures provided by Ministry of Education via request and response from D. Haddock 11 April, 2011.

This chapter outlines the process for developing and administering the middle leadership questionnaire to participants. The theory that informed the development of the questionnaire items is detailed at the conclusion of the literature review in Table 2.2 of Chapter Two of the literature review. This chapter includes detail about the participants, the schools in which they work and the ethical considerations. The next chapter explains the results of the questionnaire and the implications for adding to the existing knowledge in regard to middle leadership.

Chapter Six: Phase 2 Results – Reported Effectiveness of Departmental Leadership

Results

This section details the results from the questionnaires completed by 327 middle leader and teacher respondents from the 10 schools. The data from the hard copy questionnaires were manually transposed into IBM SPSS 19 in order to conduct a number of statistical tests. The data was coded so that middle leaders and teachers in each department in each school could be grouped for statistical analysis purposes. Missing data were entered by calculating a mean response for the item across the sample (Marsh, 1990). The coding within the software program enabled an investigation into understanding the differences and similarities between departments, subjects and schools in their views of departmental leadership. Thirty department groups were created, three within each of the 10 schools, and within each of the 30 departments the teachers' and middle leaders' responses could be separated.

Factor analysis

An initial analysis of the responses showed that there was no statistically significant difference between the responses of middle leaders and teachers for each of the factors created from the items in the questionnaire, so the responses from both groups within a department were aggregated for analysis purposes. Using a principal component analysis in SPSS with oblimin rotation, a series of factors were identified from the items in the questionnaire. A factor is a dimension or construct that is a condensed statement of the relationships between a set of variables (Kline, 1994). Five factors emerged from the 34 items in the questionnaire. The items in the questionnaires for middle leaders and teachers were identical so all completed questionnaires by both middle leaders and teachers were able to be included in the factor analysis. In the design of this phase of the project, it was anticipated that the views of the teachers and middle leaders about the practices that were occurring in their departments would be significantly different from each other. For example, it was possible that the middle leaders may have believed that their practices were more effective than did the teachers they were leading. The statistical analysis of these results showed that there was in fact a strong level of agreement between middle leaders and teachers about what was occurring in each of the 30 departments, whether the leadership practice itself was strong or weak. This point is illustrated in Table 6.1 when comparing the similarity in the mean results for the teachers and middle leaders.

The analysis phase involved making comparisons between constant sets of subgroups within the sample to identify patterns and relationships that could inform further analysis and answer the research questions (Oppenheim, 1998). The questionnaire developed for this project had a high level of reliability and validity (presented in Appendix D). Reliability in evidence is built by showing that errors in ratings are minimised (McCauley & Moxley, 1996). In this case, the raters have not interpreted items differently and the leadership practices assessed are clearly important and valid for the participants. If a factor is defined as a construct or dimension that can account for the relationship between the variables (Kline, 1994), this is evidenced within this project in the way the items have grouped into factors that relate strongly to the empirical middle leadership literature.

The reliability indices in Appendix D shows the items from the questionnaire that were grouped together for each factor, the reliability and the response rates. Cronbach's Alpha was used to test the reliability of the results. Eleven of the 34 items were grouped into the factor with the largest weighting described as *Collegial working environment*. The reliability for this factor was the highest at $\alpha = .95$ with the largest number of items included. It is usual to regard factor loadings as high if they are greater than 0.6 and moderately high if they are above 0.3 (Kline, 1994). Item 32 with the highest reliability of $\alpha = .98$, *The department is a collegial place to work where staff support each other,* was tested by reconfirming the factor analysis by removing the item. When it was not included the Cronbach's alpha = .94; hence, a decision was made to leave the item in the scale. The items in this factor all related to participants' views of the effectiveness of department leaders to ensure that the department functioned well, and was a collegial and supportive environment including a culture of professionalism, feedback and advice. This factor also included all of the items related to the team approach within the department, including the environment being collegial with shared opportunities for leadership.

Five of the items grouped into a factor named *Goals and expectations* with a high reliability score of $\alpha = .85$. This factor was made up of items that related to goal setting within a department and how those goals related strategically to school and performance appraisal goals for individual teachers. Items related to department leaders ensuring consistent high-quality teaching and that the observation of teaching practice was occurring in the department were also included in this factor. Six items grouped into a factor entitled *Focus on student academic results* with a reliability score of $\alpha = .87$. The items in this factor largely related to questions about the use of student achievement data to inform teaching and learning and the importance of student

feedback. Three items were grouped together under *Management of resources* with a reliability factor of $\alpha = .80$. These items allowed participants to comment on the effectiveness and equity of the allocation of the department budget and how decision-making regarding the allocation of resources was focussed on the needs of the students. Eight items were grouped under *Positive learning environment for students and teachers* with a reliability factor of $\alpha = .87$. These items reflected participants' views on department leaders ensuring that students received high-quality feedback and the professional learning and development of staff was acted upon in the classroom. One item was not included in any factor as it did not correlate meaningfully with any of the other items. This was item 17 regarding students who are struggling to achieve being taught by the most effective teachers. This item was deleted from further analyses.

Appendix D shows that the items within each factor are highly correlated. Some items that were grouped together in the questionnaire have correlated strongly into a factor such as numbers 5, 6 and 7 in the factor *Management of resources.* Other questionnaire items have grouped together in more unexpected ways, such as the items within the factor *Positive learning environment for students and teachers,* where respondents have responded in similar ways to the questions about staff professional development and support as they have to the questions about student learning and support in the classroom.

Comparison of means

The mean results discussed in this section include a combination of the results from the teachers and middle leaders to determine a *department* mean which is then combined to produce a *school* mean. When splitting the files to investigate middle leaders' and teachers' responses, the data set for some departments was too small and therefore not useful for analyses purposes. Due to there being no significant differences between the responses of the middle leaders and teachers, a decision was made to combine the middle leader and teacher results into one group of responses for each department. Therefore, the mean for departmental leadership for each department represents the views of the teachers and middle leaders within the department. Table 6.1 shows the mean for each factor for the whole sample, middle leaders and teachers. When comparing the difference in mean between teachers and middle leaders, the teacher average mean for every factor is lower but similar to the middle leaders. The standard deviation for teachers' responses is also greater than that of the middle leaders.

Table 6.1

Means and Standard Deviations for Each Leadership Factor for the Whole Sample

Factors	Mean of whole sample	Middle Leaders Mean Score	Teachers Mean Score
	N = 330	*n* = 96	*n* = 234
Collegial working environment	5.54 *(1.21)*	5.70 *(1.02)*	5.48 *(1.28)*
Focus on student academic results	5.33 *(1.12)*	5.35 *(1.05)*	5.33 *(1.14)*
Management of resources	5.44 *(1.41)*	5.60 *(1.21)*	5.37 *(1.49)*
Goals and expectations	5.60 *(1.05)*	5.61 *(1.00)*	5.60 *(1.08)*
Positive learning environment	5.78 *(0.90)*	5.81 *(0.90)*	5.76 *(0.90)*

Note. Mean Scale = 1–7. Standard deviation in brackets and italicised

The means and standard deviation for the responses by subject across the whole sample of 10 schools for English, mathematics and science were also calculated and are shown in Table 6.2. This enabled an across-school comparison by subject and shows that mathematics departments have the lowest comparative mean score apart from the factor *Management of resources.*

Table 6.2

Means and Standard Deviations for Subject

	English teachers and middle leaders	Mathematics teachers and middle leaders	Science teachers and middle leaders
Numbers	*n* = 110	*n* = 113	*n* = 107
Factors			
Collegial working environment	5.52 *(1.37)*	5.31 *(1.33)*	5.81 *(0.78)*
Focus on student academic results	5.42 *(1.16)*	5.19 *(1.30)*	5.39 *(0.83)*
Management of resources	5.19 *(1.67)*	5.28 *(1.46)*	5.86 *(0.91)*
Goals and expectations	5.60 *(1.14)*	5.54 *(1.19)*	5.66 *(0.78)*
Positive learning environment for students and teachers	5.78 *(0.96)*	5.64 *(1.04)*	5.92 *(0.61)*

Table 6.3 ranks whole-school means for each of the factors and in the right-hand column shows an overall school rank from lowest to highest for leadership practices across all the factors. There is some inconsistency in school rankings across factors. School 9E ranks top in one factor and third, fourth, fifth and sixth in the other four factors. The overall top-ranked school (9B) ranks first in three of the five factors. The three bottom-ranked schools overall for leadership factors, 6A, 5A and 9D, are consistently ranked in the bottom places for all five factors. This overall ranking gives a general indication of middle leadership practices within a school, compared to urban

schools in the sample but it masks the within-school variation by department detailed in Figure 6.1.

Table 6.3

Schools Ranked by Means for Each Factor (Lowest to Highest)

School rank by mean	Collegial working environment	Focus on student academic results	Management of resources	Goals and expectations	Positive learning environment for students and teachers	Overall ranking
1	6A (4.79)	5A (4.88)	8C (4.72)	6A (5.05)	6A (5.14)	6A (4.85)
2	9D (5.06)	6A (5.01)	9D (5.08)	9D (5.19)	8C (5.54)	9D (5.18)
3	8C (5.22)	9D (5.01)	5A (5.23)	5A (5.47)	9A (5.57)	5A (5.34)
4	9A (5.44)	8E (5.41)	6A (5.33)	8E (5.55)	9D (5.57)	8C (5.41)
5	5A (5.56)	9E (5.5)	9A (5.64)	3A (5.65)	5A (5.59)	8E (5.61)
6	8E (5.63)	8C (5.7)	8E (5.67)	9E (5.79)	9B (5.89)	9A (5.71)
7	9E (5.75)	8D (5.72)	8D (5.76)	8C (5.89)	8E (5.95)	3A (5.84)
8	3A (5.79)	3A (5.73)	9E (6.01)	9A (6.04)	3A (6.01)	9E (5.86)
9	8D (6.04)	9A (5.89)	3A (6.04)	9B (6.11)	8D (6.1)	8D (5.97)
10	9B (6.05)	9B (5.95)	9B (6.32)	8D (6.16)	9E (6.24)	9B (6.07)

A critical part of the study was to identify the differences in leadership practices within departments within schools. The table in Appendix E shows the mean and standard deviation for each factor for each of the 30 departments in the sample. Figure 6.1 shows the mean results for all of the factors averaged for each department in each of the 10 schools. This graph shows the difference in the average results from the leadership questionnaire responses of department members so that within-school department leadership variance as a whole can be identified. The graph shows that in nine out of the 10 schools, mathematics is the middle ranked department within the school and school 8D is the only school where the mathematics department ranks highest. In five out of the 10 schools, English is the highest ranked out of the three departments and in four schools, the science department ranks the highest for middle leadership practices. There are three English departments in the sample performing well below other departments across the sample and compared to other departments in their school.
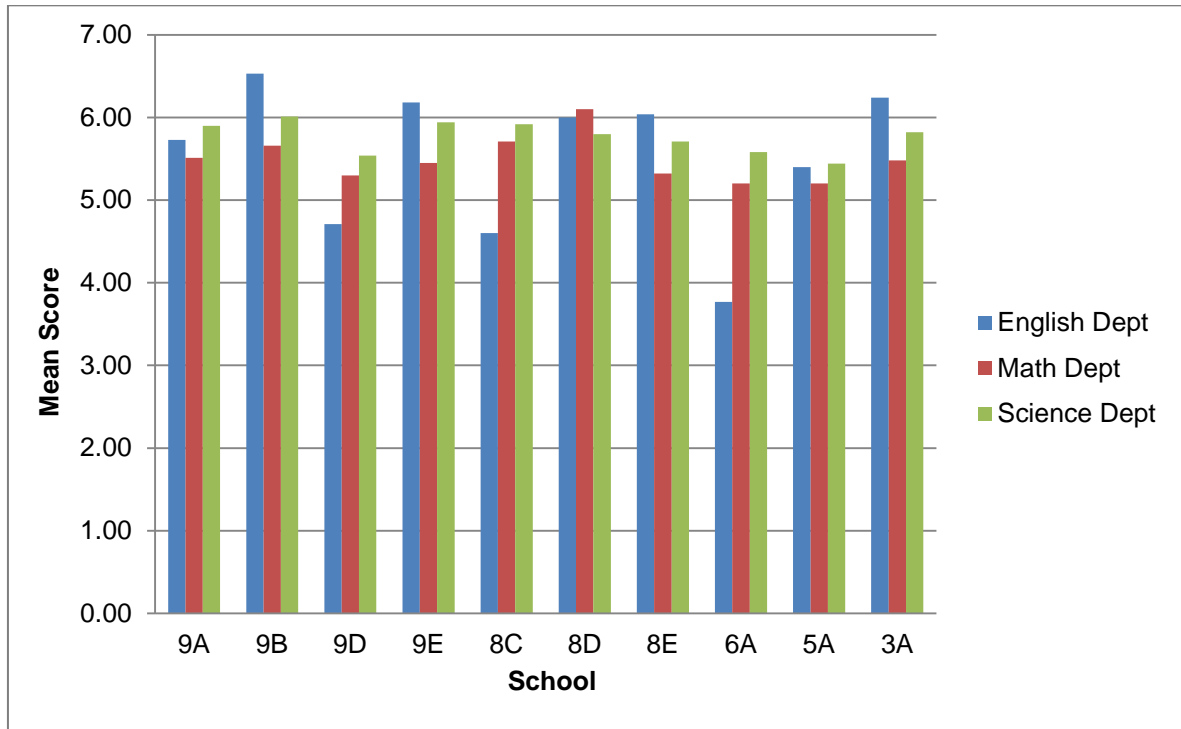
Figure 6.1. Mean middle leadership practices score for each department in each school

Further detail is presented in Figure 6.2 that shows the detail of the mean factor rating for each of the departments in the decile 8 and 9 schools. This graph clearly shows that the English department in 9D has lower mean scores in all factors compared to other English departments in higher decile schools and compared to other departments in the same school. The *Management of resources* in the English department in school 8C is the lowest ranked factor in the decile 8 and 9 schools. Conversely, the English department in school 9B has high scores in all factors and as can be seen in Figure 6.3 is the highest scoring department in the whole sample and is also the highest scoring school as can be seen in Table 6.4. Figure 6.2 clearly shows the English department in 8C and 9D as having lower mean scores in most factors compared to other English departments in similar decile schools and compared to other departments in the same school.
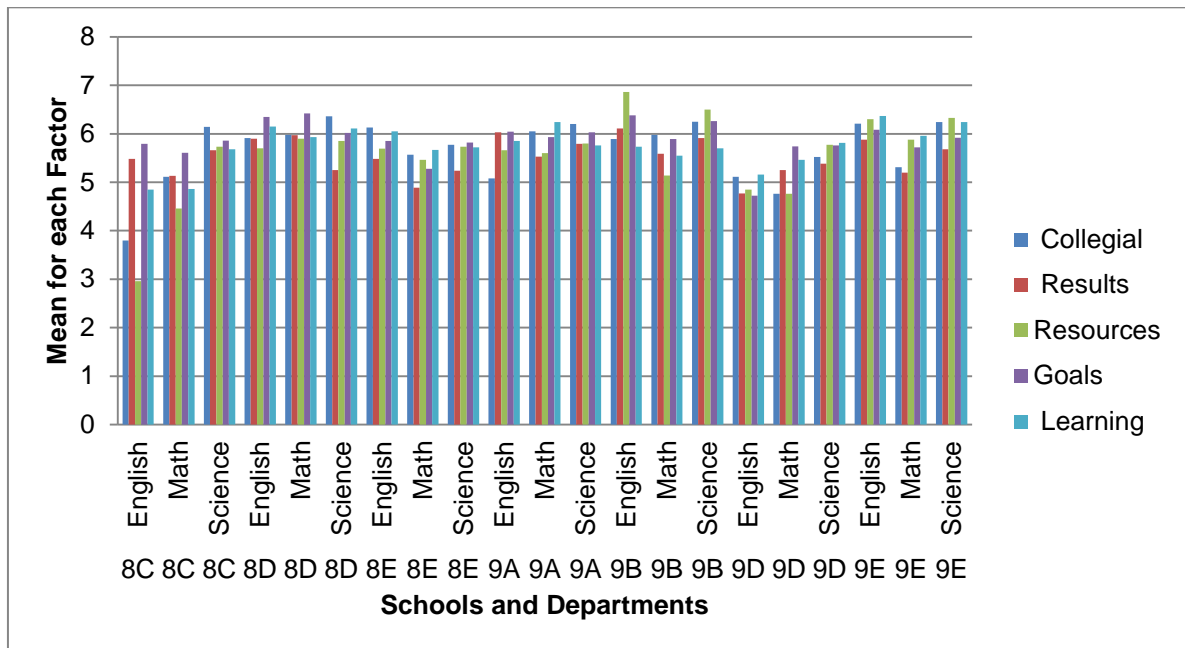
Figure 6.2. Leadership factor ratings within each department in the decile 8 and 9 schools

The graph presented in Figure 6.3 shows the detail of the mean factor rating for each of the departments in the decile 3, 5 and 6 schools. This graph clearly shows that the English department in 6A has lower mean scores in all factors compared to other English departments in lower decile schools and compared to other departments in the same school. The *Management of resources* in the English department in school 6A is the lowest ranked factor in the whole sample. Conversely, the English department in school 3A has high scores in all factors and as can be seen in Figure 6.2 as one of the highest scoring departments in the whole sample.
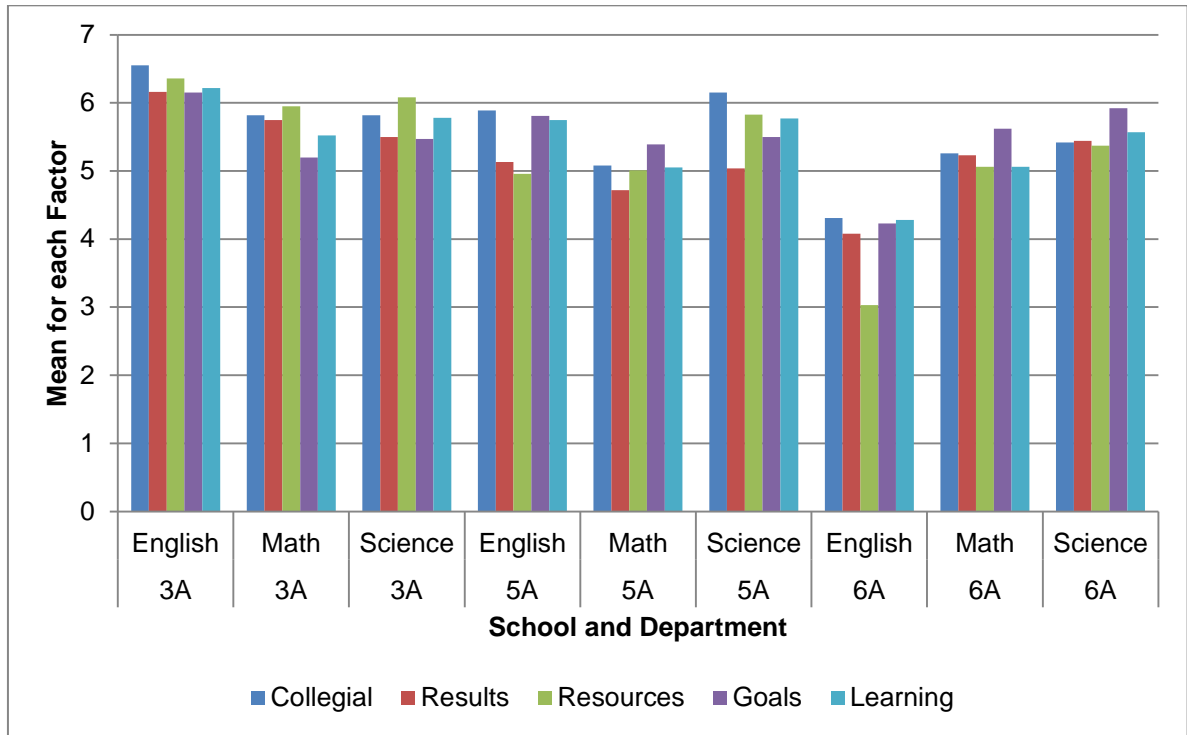
Figure 6.3. Leadership factor ratings within each department in the decile 3, 5 and 6 schools

Analyses of variance (ANOVAs)

A series of analyses of variance (ANOVAs) were conducted to assess whether there were any statistically significant differences among the views of middle leaders and teachers within departments in schools and across the whole sample. The analysis of variance is a statistical technique that evaluates whether there is any systematic (i.e., non-random) difference among the set of means. The first ANOVA was conducted to test for statistically significant differences for each factor by comparing subject departments (Table 6.4).

Table 6.4

Results of ANOVA for Each Factor by Subject Departments

Variable and source	Sum of squares	*df*	Mean square	*F*	Sig.
Collegial working environment				1.791	.169
Between groups	5.850	2	2.925		
Within groups	377.151	231	1.633		
Total	383.001	233			
Focus on student academic results				.432	.649
Between groups	1.148	2	.574		
Within groups	306.725	231	1.328		
Total	307.873	233			
Management of resources				6.646*	.002
Between groups	28.313	2	14.156		
Within groups	492.082	231	2.130		
Total	520.395	233			
Goals and expectations				.361	.698
Between groups	.850	2	.425		
Within groups	272.325	231	1.179		
Total	273.175	233			
Positive learning environment for students and teachers				1.122	.327
Between groups	1.847	2	.923		
Within groups	190.037	231	.823		
Total	191.884	233			

Note. $*p \leq .05$

Differences between English, mathematics and science departments were explored in relation to the factors: *Collegial working environment, Goals and expectations, Focus on student academic results, Management of resources and Positive learning environment for students and teachers.* Where statistically significant differences were found a post hoc Tukey HSD test was conducted to determine where the differences lay. There was a statistically significant difference between subject departments in how they viewed the

management of resources ($F(2, 231) = 6.65$, $p = .002$). To manage the possibility of Type 1 error, Bonferroni corrections were computed by taking 0.05 divided by 5. This yielded the value of 0.01. The *p* value of *Management of resources* was .002; therefore, it satisfies the Bonferroni requirement. There were statistically significant differences for *Management of resources* between the English and science departments ($p = .003$) and between the mathematics and science departments ($p = .01$). There were no statistically significant differences across subject departments in their views of collegiality ($F(2, 231) = 1.791$, $p > .05$), focus on student academic results ($F(2, 231) = 0.43$, $p > .05$) goals and expectations ($F(2, 231) = 0.36$, $p > .05$), or positive learning environment ($F(2, 231) = 1.12$, $p > .05$).

A second ANOVA was conducted to determine the statistically significant differences for each factor by comparing whole-school results (Table 6.5). There were statistically significant differences between *schools* as a whole in regard to how they viewed the collegial working environment *($F(9, 224) = 2.73$, $p = .005$)*. When the post hoc Tukey test was performed it showed that there was a statistically significant difference between school 6A and school 8D, which is supported by the descriptions of whole-school mean where school 6A has a mean of 4.79, which is considerably lower for this factor than that of school 8D, which is 6.04.

Table 6.5

Results of ANOVA for Each Factor by School

Variable and source	Sum of squares	df	Mean square	F	Sig.
Collegial working environment				2.725	.005
Between groups	37.793	9	4.199		
Within groups	345.208	224	1.541		
Total	383.001	233			
Focus on student academic results				2.749	.005
Between groups	30.624	9	3.403		
Within groups	277.249	224	1.238		
Total	307.873	233			
Management of resources				5.305	.000
Between groups	91.425	9	10.158		
Within groups	428.970	224	1.915		
Total	520.395	233			
Goals and expectations				2.609	.007
Between groups	25.923	9	2.880		
Within groups	247.252	224	1.104		
Total	273.175	233			
Positive learning environment for students and teachers				3.897	.000
Between groups	25.978	9	2.886		
Within groups	165.905	224	.741		
Total	191.884	233			

Note. p ≤ .05

There were statistically significant differences between 9 pairs of schools in the results for the factor *Management of resources* ($F(9, 224) = 5.30$, $p = .001$). The post hoc Tukey test showed that the differences existed between school 3A and schools 6A ($p = .001$) and 8C ($p = .006$) both with low means of 4.55 and 4.20 respectively. The difference in this factor also existed between school 6A and schools 8D ($p = .008$), 8E ($p = .004$), 9E ($p = .001$) and 9B ($p = .001$). There were a range of means for these schools with school 8C having the lowest mean of 4.72. There were differences in this factor between school 8C and 8E ($p = .04$), 9B ($p = .01$), 9E ($p = .004$) and a marginally significant difference in relation to school 8D ($p = .052$) mainly because the mean for school 8C was so low at 4.20.

There were statistically significant differences between one set of schools in regard to the factor *Positive learning environment for students and teachers* ($F(9, 224) = 3.90$, $p = .001$). The post hoc Tukey test showed the statistically significant difference was between the results for school 6A and schools 3A ($p= .01$), 8D ($p = .003$), 8E ($p = .01$) and 9E ($p = .001$).

Discussion

This section identifies the main findings from the questionnaire data that will inform further analysis when combined with department and school student achievement data. The first un-anticipated finding in this project was the level of agreement between middle leaders and teachers in the rating of their experience of the middle leadership practices in their department. Second, that there are some factors or aspects of middle leadership that respondents viewed positively by giving a high score to items and others they viewed negatively by giving a low score, sometimes within the same department. There were statistically significant differences at school and department level in the way middle leaders were perceived to manage resources. Third, there was a tendency for some schools to have overall poor middle leadership but there was also within-school variation. The trend appeared to be that if the English department in a school was a poor performer in relation to middle leadership practices, the school was a poor performer and high-performing English departments were in the overall high-performing middle leadership schools. The fourth finding is that the respondents from mathematics departments reported the lowest scores on all factors in comparison to English and science, apart from *Management of resources*. The fifth finding was that the quality of middle leadership performance at school level was not always decile related and some schools with socially disadvantaged students were shown in this study to score higher in items related to middle leadership practices than schools with socially advantaged students.

The finding about the level of agreement between teachers and middle leaders about the leadership practices within a department was surprising. McCauley and Moxley (1996) discuss the importance of gaining feedback from different perspectives. They state that the comparison between self-ratings and feedback from others can be so diverse that it challenges a manager's behavioural patterns and provides motivation to rethink the impact of behaviour on others. To some extent, the level of agreement between department members in this study is an important and positive result. If department members hold similar views about problem areas for development, then issues could be easier to address. The ethical considerations of this study did not allow for the middle leaders to directly receive the results from the responses from colleagues in their department, but many of the middle leaders who took part in the project were keen to re-use the instrument for department discussion as a result of participation in the research. In organisations that have a development orientation where continuous learning and growth is an expectation, instruments such as this questionnaire can provide an opportunity for reflection and improvement.

When the results for the whole sample were analysed in SPSS, there were some statistically significant differences between some subjects and schools in how they viewed different factors. The factor *Management of resources* was revealed as being statistically different between the subjects mathematics and science and English and science. This factor was also statistically significantly different between all 10 schools in the sample. The items that sit within this factor describe the equitable access to resources that are well aligned with priority teaching goals. These decisions are well within the control of the middle leaders and yet there seems to be considerable variation between participants' responses right across the sample in their perception of how these resources are being managed to best advantage student learning. These results have similarities with the existing literature. In a study that involved the analysis of interview data from 39 high school teachers, Johnson (1990) found that departments exerted considerable control over course offerings, tracking decisions, textbook selection and teacher and student assignment to particular courses. In a study by Ball and Lacey (1980) based on four case studies of departments in British comprehensive schools, the authors argued that the influence of a department over school policy and resource decisions is based on the number of subject specialists belonging to a department and on the cohesiveness and consensus within the department itself as well as the nature of the teaching and learning within a subject. Little (1993) supports this statement in her contention that departmental membership and leadership affect relative departmental strength within a school. She asserts that a department constituted of full-time subject specialists has an advantage in the competition over

school resources. Ball and Lacey (1980) warn that although strong departments create positive working environments they can sometimes dominate in gaining the greatest amount of the scarce resources within a school. The varied responses to this questionnaire from departments within and across schools would indicate that this is a highly contested area and a source of considerable dissatisfaction in some schools and high levels of agreement in others.

The finding that respondents from different subject areas viewed the management of resources in their department differently is supported by the existing literature on the content and context of school subjects in secondary schools (Grossman & Stodolsky, 1995). An additional complexity is that school subjects often require different resources for effective instruction. Science instruction with a laboratory component may use disposable materials that need to be regularly replaced while texts read in English have a longer useful life (Grossman, Stodolsky, & Knapp, 2004). The way the demand for equitable access to resources is played out within departments and across schools has been shown in this study to be widely variable with school decile making little difference to the way teachers and middle leaders score this factor. The English department in the lowest decile school (3A) gained one of the highest mean scores of 6.29 with higher decile English departments in school 6A gaining a very low mean of 2.72 and the English department in the single-sex girls' school in an affluent area (9D) gaining a mean score of 4.61.

The statistical analysis of the means by factor by department, school and subject show some degree of variance of practice both within schools and across subjects. Figure 6.2 showed that some schools (e.g., 9B) have overall high middle leadership performance in each of the English, mathematics and science departments, some are low (6A) and others have mixed results (8C). None of the 10 schools have three low-performing departments in terms of middle leadership, although school 6A performs lowest in relation to the other schools in most of the factors. The three lowest ranked schools overall in terms of middle leadership, 6A, 8C and 9D, all have English departments performing significantly below other departments in the school in terms of middle leadership and this impacts on the schools' overall rank. Grossman and Stodolsky (1994) argue that subject differences should not be ignored and account for differences in instructional practices and the way students experience the taught curriculum. An analysis of a range of studies has led them to claim that teachers in the humanities are more progressive and inclined toward less transmission modes of instruction with science and mathematics teachers focussed on the sequential acquisition of accurate knowledge. In the context of this study, it appears that when instructional leadership

practices in the English department are functioning well, that other departments may benefit. The results from the middle leadership questionnaire show that when English teachers are being well led, then there is a positive relationship to the middle leadership in the mathematics and science departments.

Where all three departments perform in the top half of the sample seen in Table 6.3 (8D, 9A, 9B, 9E, 3A), this could be attributed to a whole-school effect. There could be factors such as school leadership within the school that are positively impacting on each of the departments. Siskin (1995) would argue against this assertion because she considers that the microclimate or department level is where teachers *live* and experience the working environment. She regards leadership strategies aimed at the whole school as not worthwhile as they would actually miss almost every department and teacher within them. The interaction between department and school effects is discussed in more detail in the next chapter in relation to student academic outcomes and further conclusions are drawn in the results section of Chapter Eight.

The highest performing subjects or departments in terms of middle leadership practices differ throughout the 10 schools. Four schools have English departments as their highest performer and four schools have science departments as the highest performer. In school 5A, the whole-department middle leadership mean score for science and English are the same (5.4). The department least likely to score the highest in terms of middle leadership practices within a school is mathematics. Just one school (8D) can claim that the mathematics department is the most effectively led. There was a statistically significant difference between mathematics and science teachers and middle leaders in how collegial they viewed their departments. Grossman and Stodolsky (1995) argue the extent to which departments are composed of teachers from diverse disciplines may contribute to the degree of cohesiveness of the department. Their research of the perceptions of 399 teachers of five academic subjects in 16 schools found that subject subcultures may be characterised by both beliefs about the subject matter that bind teachers together and by norms regarding teaching practice, curricular autonomy and coordination. They state that the issues and concerns of the typical mathematics teacher are not the same as the typical English teacher, nor do they work under the same constraints. Grossman and Stodolsky's (1995) view is that secondary school teachers are socialised into a particular view of the world as seen through a disciplinary lens and that this socialisation relates to disciplinary ways of thinking. The findings in this project support this theory particularly in terms of the collegial experience of mathematics colleagues. The maths respondents in this sample gave lower ratings to items related to collegial environment, team work, opportunities for leadership, robust

opportunities for professional discussion and a safe, supportive and well-organised working environment, than their science colleagues who responded. Subject teachers from each of the English, mathematics and science departments across the sample scored the items and perceived their importance differently. Further follow-up observations and probing would be required to understand if in fact mathematics departments are less collegial or whether mathematics teacher perceive the importance of these types of leadership behaviours as less important than other characteristics described in the questionnaire.

The decile rating or socioeconomic advantages or disadvantages of the students in the school appear to have some bearing on the quality of middle leadership in the English, mathematics and science departments. Table 6.3 shows that generally the higher decile schools in this sample gained higher middle leadership scores but school 3A consistently has a score over 6 out of 7 in all factors. It could be argued that the teachers and middle leaders in this department work in a school where high proportions of students come from economically disadvantaged circumstances, where English is not a first language or spoken in the home. Conversely, Table 6.3 shows the English department in school 9D as scoring 5 or less in every factor and yet it is situated in a socially advantaged single-sex girls' school where English is the first language for the great majority of students.

This chapter has summarised the reported leadership practices of the 30 departments in the study and drawn together some summarised findings that are important in relation to identifying the leadership practices that have a relationship with academic outcomes for students. The next chapter details the student academic results in each of the 30 departments in the 10 schools and in Chapter Eight the middle leadership and academic outcomes scores are statistically tested and compared to determine patterns and relationships.

Chapter Seven: Phase 3: Results from the Analysis of the Student Academic Outcome Data in the 10 Phase 2 Schools

In this chapter, the results of the analysis of student academic results in the 10 schools that participated in the middle leadership questionnaire will be reported. The Level 1 NCEA student achievement results in English, mathematics and science were collected for 2008, 2009, 2010 in Phase 1 of this project. In Phase 3, further results for NCEA Levels 1, 2, 3 and scholarship in 2009 and 2010 were collected for each of the departments in the Phase 2 schools in order to gain an understanding of how departments were progressing over time. Many of the results reported in this chapter focus on 2010 because the questionnaire was administered in these schools in late 2010. It was important to collate the NCEA results for student academic performance in the same year that the questionnaire was conducted.

Results

The number of additional results collected for this phase have been summarised in Table 7.1. Information regarding the year and level of the results that have been collected is also summarised.

Table 7.1

Levels and Number of Student Academic Results Collected for Phase 2

	Total NCEA results collected from the 10 Phase 2 schools			
Year and Level	English	Mathematics	Science	Total
2008 Level 2	136474	124216	137790	398480
2009 Level3	54533	21841	72276	148650
2009 Scholarship	72	79	111	262
2010 Level 1	197760	102218	130348	430326
2010 Level 2	139077	122301	145756	407134
2010 Level 3	57285	23046	77182	157513
2010 Scholarship	65	73	264	402
Total	585266	393774	563727	1542767

Note. Results are the score recorded for an achievement standard by an individual student.

The collection of this data enabled analyses to occur that would identify the variation in academic outcomes within and across the 10 schools over a period of time. An identification of the extent of the within-school variation in each of the 10 schools over

the previous three years was important to identify the variation of performance across departments within the school. A measure of overall whole-school performance based on the academic student results from all three departments enabled a comparison between whole-school performance and within-school variation. An analysis of the academic success of students within each of the 30 departments and their progress over the previous three years enabled the identification of the highest and lowest performing departments in terms of academic progress to be clearly identified. This was important because the results from the middle leadership practices questionnaire needed close consideration for those departments identified as being highest and lowest performing. An analysis *by subject* for the 10 schools supported a close comparison of subject performance across the sample that could then be compared with the leadership practices' results.

The data collection for this phase included:

1. The collation of NCEA results for English, mathematics and science for Level 1 for 2008, 2009 and 2010 in order to calculate a GPA and then determine an average GPA for the three years.

2. The collation of NCEA results in English, mathematics and science for Level 2 in 2008, Level 3 in 2009 and 2010 and scholarship.[10]

3. The percentage of students either not achieving or gaining Achieved, Merit or Excellence in Level 1 NCEA English, mathematics and science within each of the schools in 2008, 2009 and 2010. This measure was important for comparative purposes because the number of Merit and Excellence grades achieved per department has been used as an indicator of academic success of a department.

4. The percentage of students either not achieving or gaining Achieved, Merit or Excellence in Level 1, 2, 3 and scholarship in NCEA English, mathematics and science within each of the schools in 2010.

Tests were then performed to compare the performance of each of the schools with each other and of departments within and across similar schools. This was achieved by calculating the GPA for the whole school over three years (2008, 2009, 2010) to determine the overall performance for the whole school in English, mathematics and science. The GPAs were used to calculate the effect size for the variation within each

---

[10] Scholarship exams provide recognition and monetary reward to top students in their last year of schooling. Students pay a fee to sit scholarship exams and they are assessed against challenging standards that are demanding for the most able candidates (NZQA, 2011b).

school in 2008, 2009 and 2010. This showed the degree of variation in academic performance of students in English, mathematics and science within the school. In order to determine the effectiveness of departments to support academic success, a calculation was made to show the percentage of Excellence and Merit grades achieved by students within each department in 2008, 2009, 2010.

Whole-school performance and within-school variation

The results described in this section have been aggregated from hundreds of student results from NCEA achievement standards at department and then school level. The results are not collected at individual student level and therefore cannot be used for statistical comparisons or correlations to compare individual student achievement. The aggregation of results for comparison does allow school and department results to be compared with each other and with national norms.

In order to establish an understanding of academic performance in the 10 schools, two measures were used and compared to national norms for decile (Figure 7.1). The first measure (blue bar) shows the overall GPA for Level 1 English, mathematics and science averaged for 2008, 2009 and 2010 in each school. The red bar shows an average GPA for whole-school performance in 2010 across all three levels of the senior secondary qualification, calculated and averaged from the English, mathematics and science results.
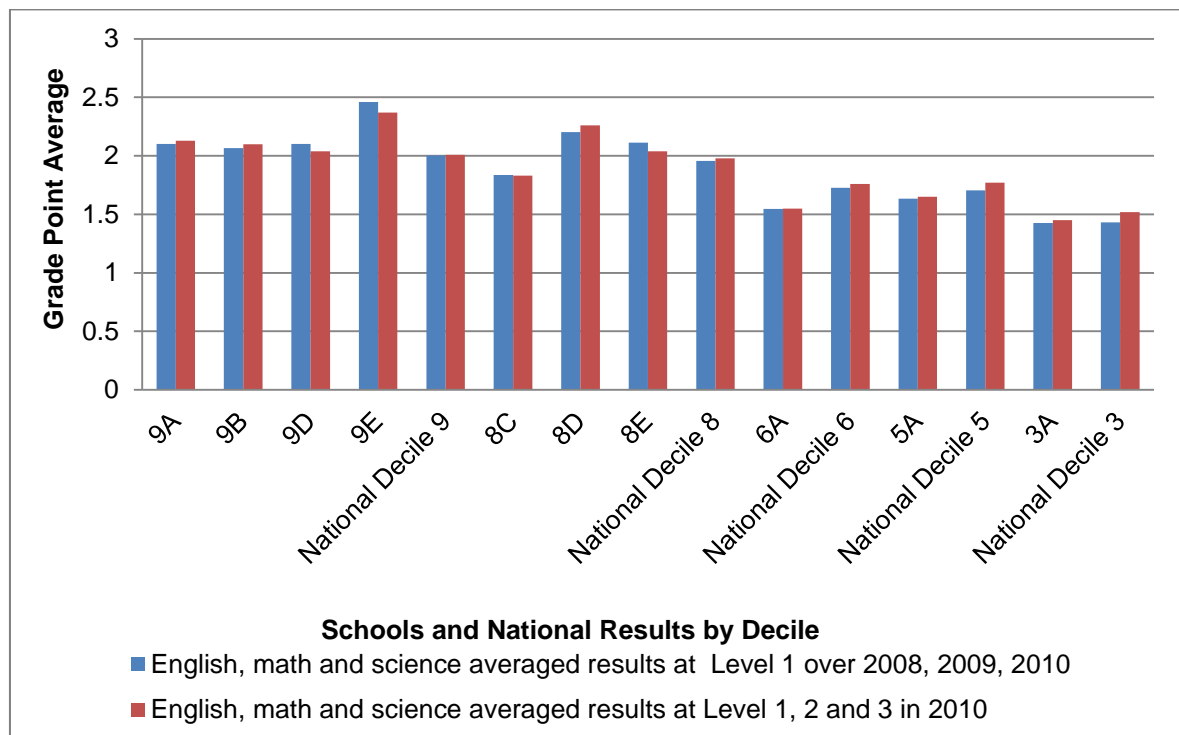

Figure 7.1. Whole-school performance of Phase 2 schools compared to national norms by decile

When comparing the two measures for whole-school performance (red bar and blue bar), there are strong similarities in outcomes when comparing student achievement at Level 1 over three years and the results averaged across all levels of NCEA in 2010. Figure 7.2 shows that three of the four decile 9 schools are performing above expectation for whole-school measures compared to the norm for that decile. Two of the decile 8 schools are performing above national norm and one is not. The decile 3, 5 and 6 schools are not performing above the norm for their decile when using these measures. Figure 7.2 shows the schools that are performing above and below in relation to national norms when the GPA is averaged over three years. This figure shows that six of the schools in Phase 2 are performing better than the national norm when the English, mathematics and science results are aggregated and four schools are performing below national norms on this measure.
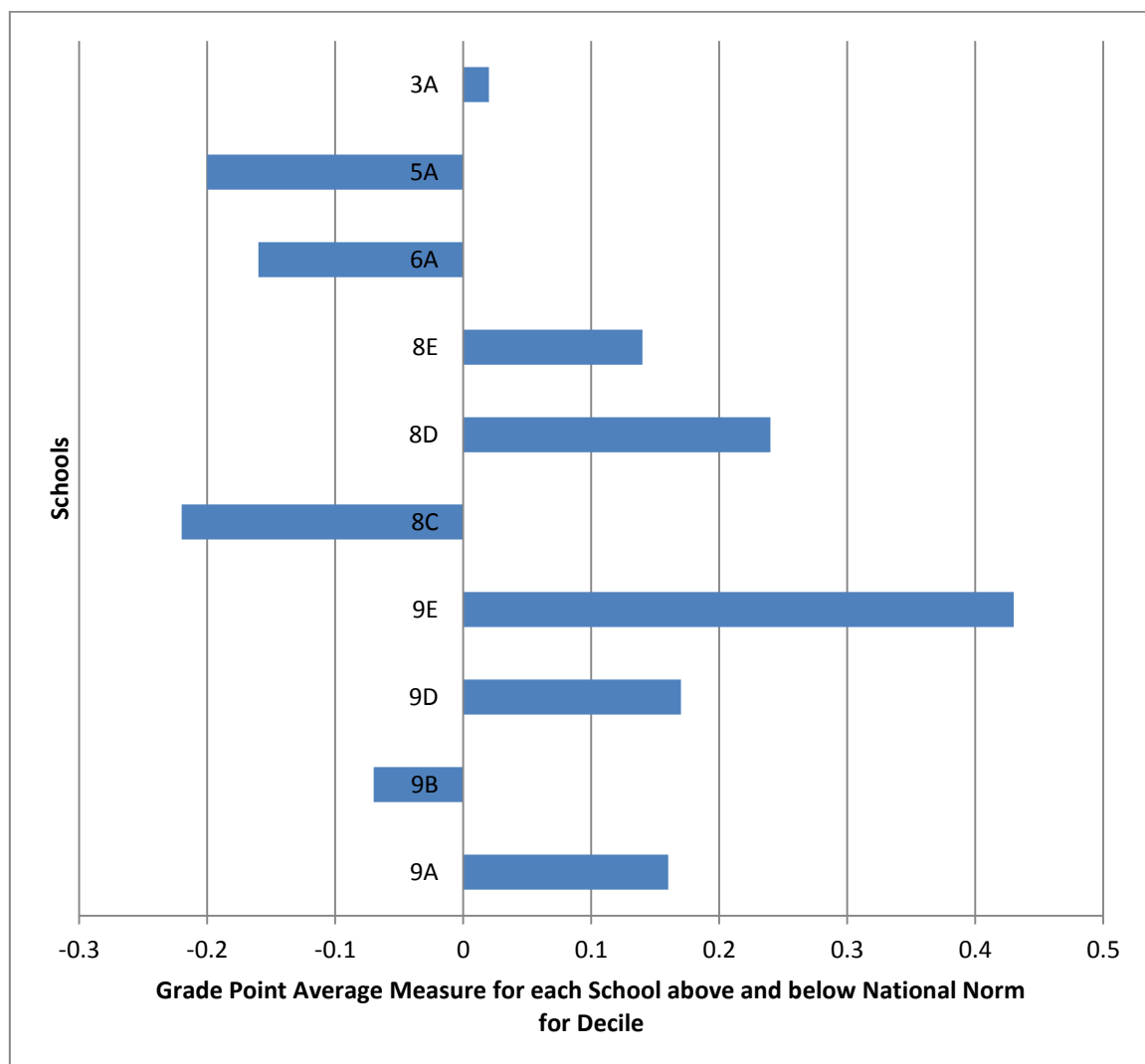

Figure 7.2. Difference between school average Level 1 Grade Point Average in English, mathematics and science across three years as compared to the national norm for that decile

Within-school variation in Phase 2 schools

Cohen's *d* was used to calculate the size of the difference between department results to determine the extent of within-school variation. Figure 7.3 shows the effect size for the Phase 2 schools in 2008, 2009 and 2010. This calculation shows the amount of within-school variation in the performance at Level 1 (15-year-olds) of the English, mathematics and science departments in the three years these data have been collected. Using the GPA, the maximum difference between the highest and lowest performing departments was calculated for each school. A calculation was then made to determine the standard deviation for each subject using the national results for all of the deciles. The effect size for each school was achieved by calculating the difference between the highest and lowest performing departments for each school and then applying Cohen's formula to divide this number by the appropriate average standard deviation calculated from all the national decile results. For example, in school 9A the highest performing subject in Level 1 in 2010 is mathematics with a GPA of 2.45 and science with a GPA of 1.83, which produces a difference of 0.62. The mean standard deviation for the difference between mathematics and science when calculated from the national results is 0.32. When 0.62 is divided by 0.32, the effect size is 1.95. The national effect size for decile 9 schools is 0.91 so school 9A has an effect size double that of the national sample for that decile. The effect-size calculation can only be made using the Level 1 results because the cohort of students is the same at this level. In NCEA Level 2 and 3, the subjects are likely to be optional; therefore, students have selected subjects where their motivation and aptitude is greatest.
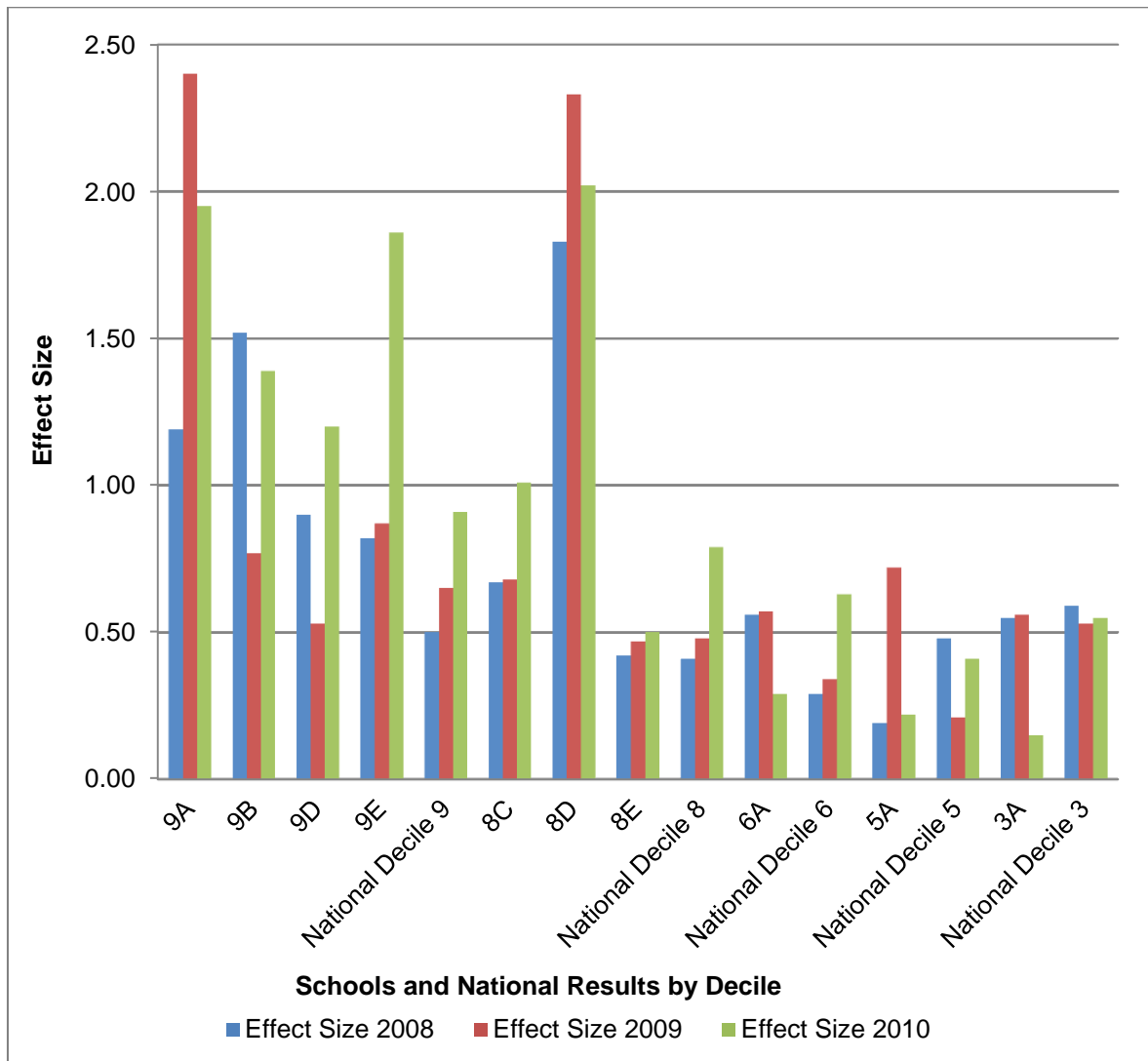
Figure 7.3. Phase 2 school effect sizes in 2008, 2009, 2010

The effect-size results are relevant in terms of comparing departments and schools. The lower decile schools have less within-school variation and the national results also have less fluctuation, as would be expected given the number of schools reflected in the national results. The lower decile schools also have less within-school variation at national level compared to the decile 8 and 9 schools. Many of the schools appear to have strong 'spikes' in variation in particular years. Six schools consistently have an effect size significantly greater than 0.5.

Department performance in Phase 2 schools

The same measures used to describe whole-school performance have been used to describe relative department performance. Figure 7.4 shows one measure of how a department performed *over time* by averaging three years of Level 1 results. Averaging the results over three years rather than using the results for a single year levels out the variation between years that can occur in a department or school due to a particular

cohort, policy or leadership change (Thomas et al., 1997b). Figure 7.5 shows department academic performance *in depth* by collecting and averaging student achievement data from NCEA Levels 1, 2 and 3 in 2010. In six schools, the pattern of department variation using the two measures is almost identical. For example, when the results represented in Figures 17 and 18, schools 9A, 9D, 9E, 8C, 8D and 8E are compared the pattern of highest, lowest and middle performing departments are the same for both measures. The variation between departments in the schools is more pronounced when comparing the department performance in the single 2010 year (Figure 7.5). In four of the schools (9B, 6A, 5A and 3A), the highest and lowest performing department changes when comparing the two measures. Some departments appear relatively more successful in one measure compared to the other. The differences in the department results are less pronounced for the mean Level 1 performance over three years, compared to the results when all three levels of NCEA are included. For this measure (Figure 7.5) all the decile 9 departments are performing at or above the national norm for their decile. School 8C performs below national norms for decile as does the English department in school 8D. Schools 6A, 5A and 3A are all performing below national norms for their decile in both measures.
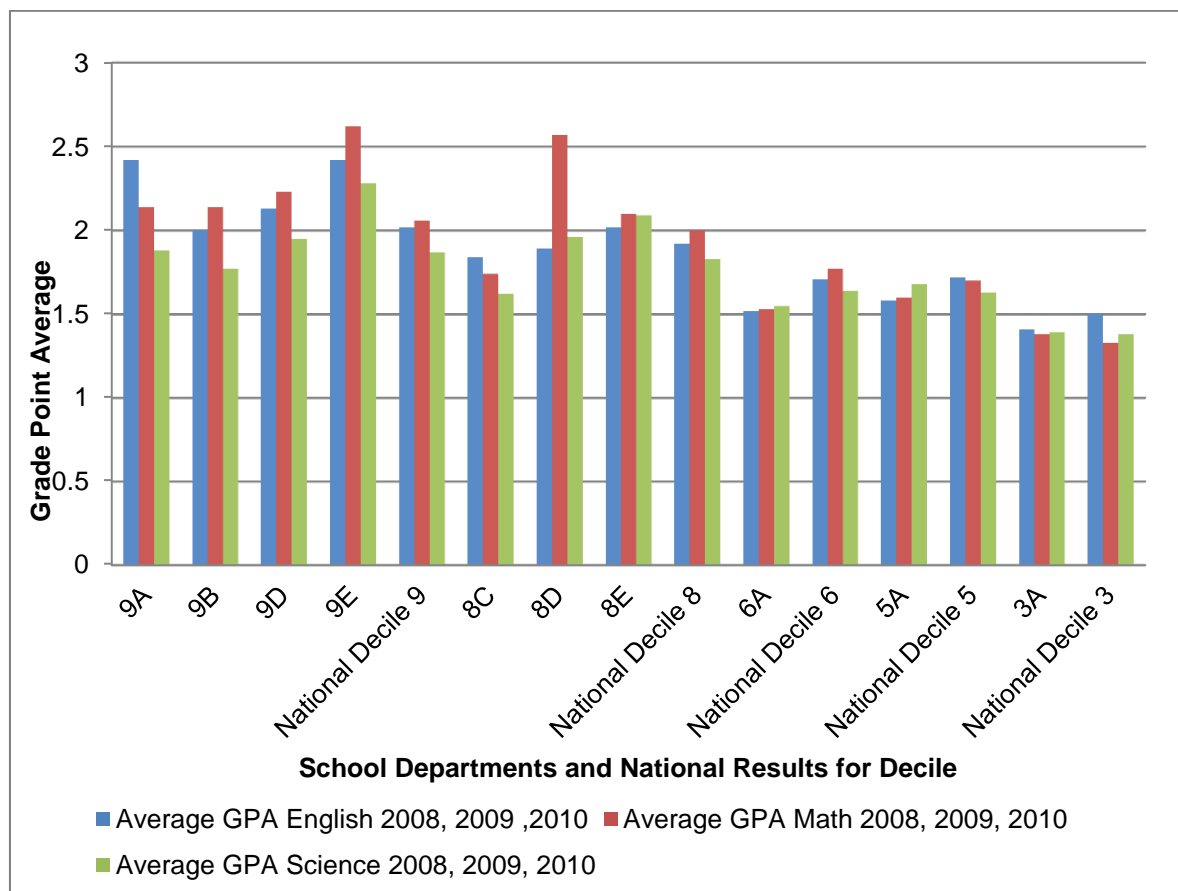

Figure 7.4. Averaged Level 1 (15-year-old) performance in Phase 2 schools compared to national norms in 2008,2009, 2010
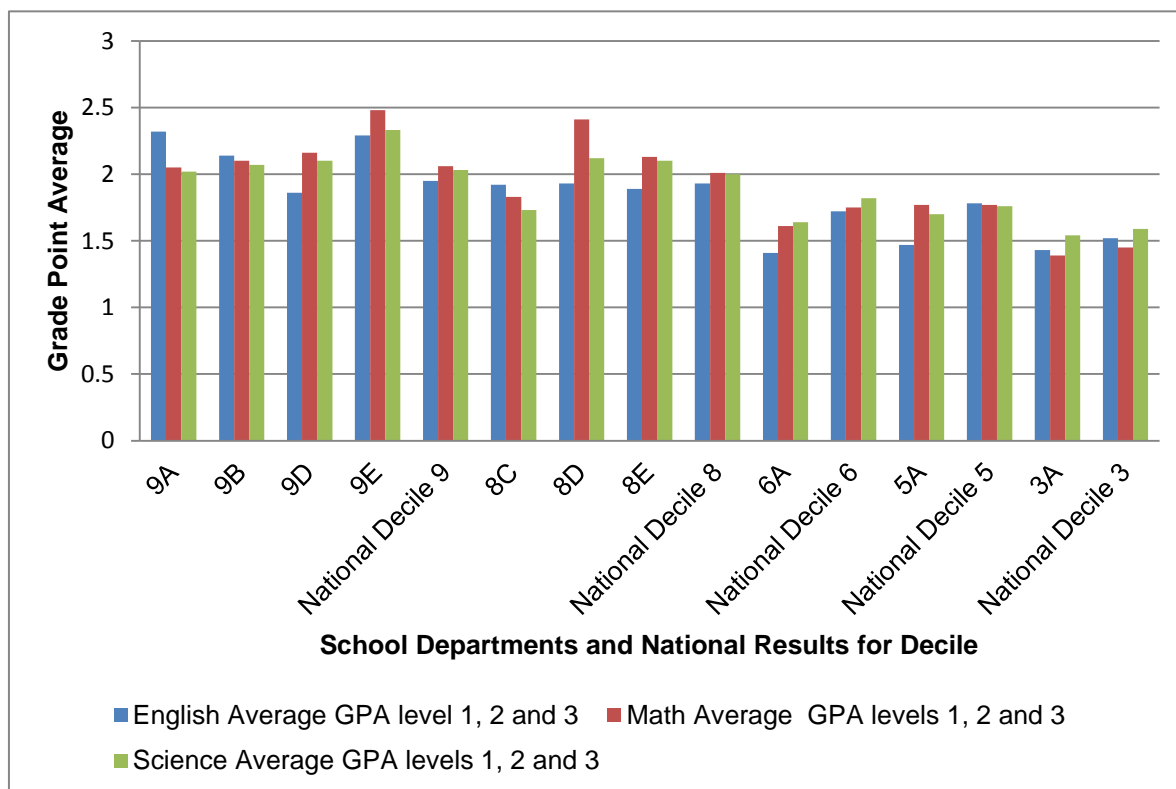
Figure 7.5. Department performance across Level 1, 2 and 3 in each of the Phase 2 schools in 2010

Another measure of department performance is the percentage of Merit and Excellence grades students achieve. When considering performance using this measure (Figure 7.6) the highest and lowest performing departments can be identified in terms of students gaining the best results as opposed to 'just passing.' The measures are useful in comparing departments within schools, across schools and against national norms for decile. It is possible to identify where whole-school performance 'masks' underperformance in some departments. For example, Figure 7.2 shows school 9D performs above the norm for decile and in Figure 7.6 the English department in school 9D underperforms compared to the national norms for English. Conversely, department performance above expectation can also be hidden by whole-school results. While Figure 7.2 shows that schools 6A, 5A and 3A all underperform in terms of whole-school performance compared to other schools of similar socioeconomic background, Figure 7.6 shows that the 6A mathematics and science departments and 5A mathematics department are performing above expectation for decile in this measure.
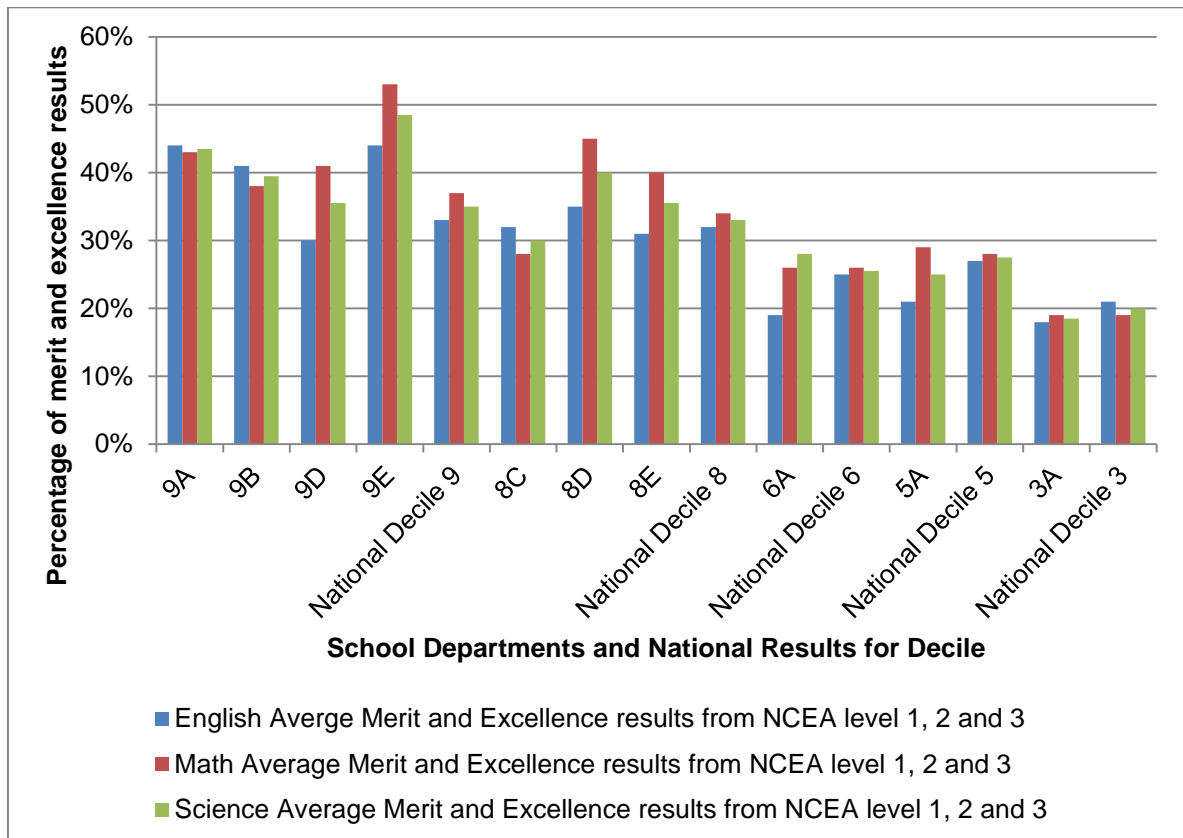
Figure 7.6. Percentage of Merit and Excellence results averaged for each department in Phase 2 schools in 2010

Identifying high- and low-performing departments in terms of student academic outcomes

In order to categorise departments in the sample as either underperforming or high performing, a range of measures were assembled in order to detect patterns, academic progress of students and to compare departments to each other and national norms. The table in Appendix F shows department performance described by a range of measures. These measures draw on a range of data that can be compared to national norms. In the first column, the average percentage of Merit and Excellence results at Levels 1, 2 and 3 in 2010 have been shown for each department with the national norm in brackets. The second column shows the average GPA for Levels 1, 2 and 3 in 2010 and the national norm is shown adjacent in brackets for comparative purposes. The third column shows the Level 1 GPA in 2010 compared to the national norm. The fourth column is designed to show progress or fluctuations in results for each department across three years by showing the Level 1 GPA for each department consecutively. The fifth column shows the percentage of Merit and Excellence at Level 3 in 2009 and 2010 respectively compared to national norms and the last column shows the number of scholarships attained by students in each department in 2009 and 2010.

The departments that are ineffective or effective in terms of raising student academic outcomes in relation to national norms are highlighted. The departments highlighted have either performed consistently above or below on most of the measures and those with no highlight can be regarded to be performing at expectation. Appendix F shows that eight departments in the sample are performing well above the national norm for their decile in a range of comparative measures. In school 9E, all three departments are performing above the national norm. There are seven departments that are underperforming in four or more of the measures, although two of these departments (5A science and 6A science) have gained good scholarship results. The two departments that are underperforming in school 8C are also improving as shown in the GPA progress at Level 1 in 2008, 2009 and 2010. The most effective departments have performed above the norms in all the comparative measures and in some cases have many students who have achieved scholarship in the subject.

Subject comparison

Comparing the performance of subjects when the Level 2 and 3 data are available reveals different patterns for subject or department performance than the examination of Level 1 data. As students progress into higher levels and deeper learning of a subject the teaching and learning becomes more complex. Figure 7.7, 7.8 and 7.9 show subject performance in 2010 for all the Phase 2 schools across all three levels of the qualification. These graphs show the variation by subject for each level of the qualification and how this compares to the national norm for similar schools in that decile. These results reflect the high and low performance of departments shown in Appendix F. Figure 7.7 shows the results for all of the English departments in the sample. Seven of the 10 English departments achieved a higher average GPA for Level 1 NCEA. Six of the departments achieved their worst results in Level 2 and all but one of the departments experience a drop in average GPA at Level 2. A trend in all but one of the schools (5A) is for the English results to improve again at Level 3. The English departments in schools 6A, 5A and 3A do not achieve results above the norm for schools in the same decile at any level of the qualification and 9D performed below norm for decile at Level 2 and 3. The English departments in schools 8D and 8E perform below national decile at Level 3.
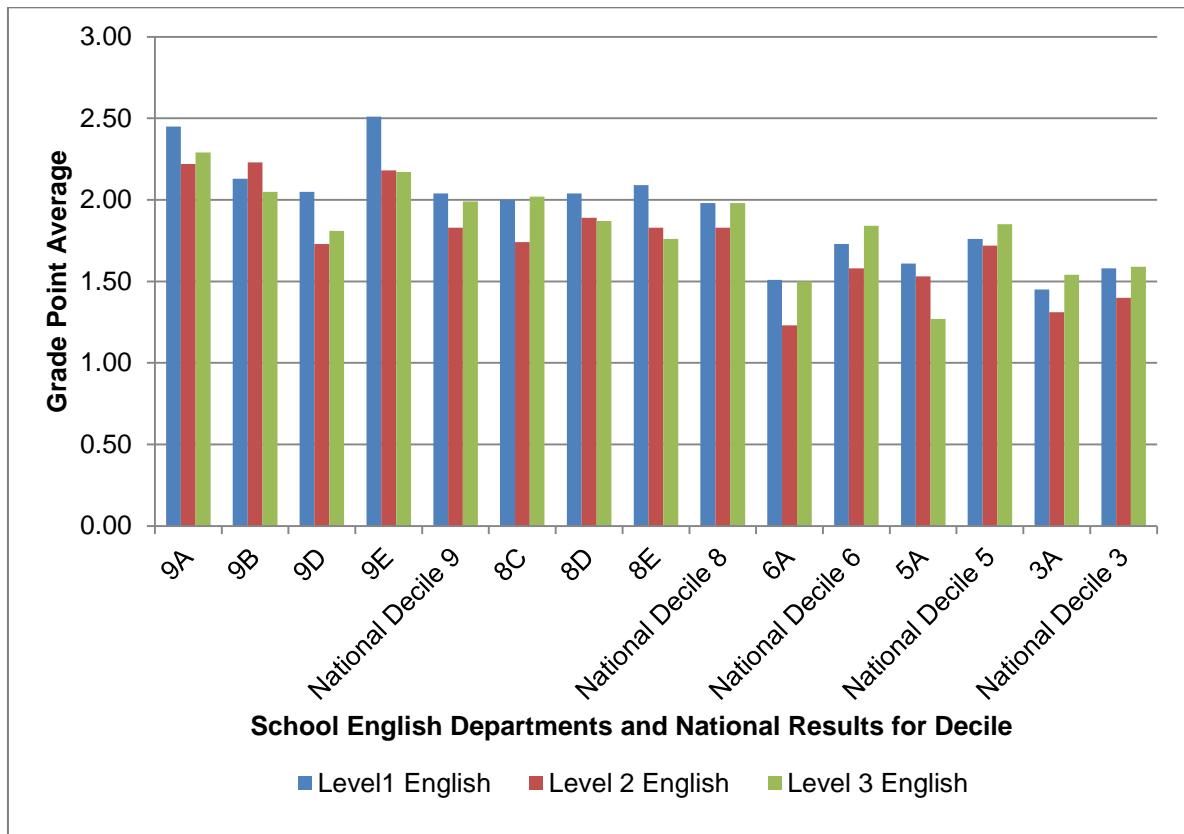
Figure 7.7. English department performance in Level 1, 2 and 3 in 2010 in Phase 2 schools

All of the higher decile mathematics departments (except 8E) experience the highest GPA at Level 1, but the pattern is different for the decile 3, 5 and 6 schools. Schools 8E and 5A achieve the highest results for Level 2 students and school 3A achieves the highest results at Level 3, above the national norm for decile 3 schools. The mathematics departments in schools 8C and 6A did not achieve results above the norm for their decile at any level. The decile 8 and 9 mathematics departments have a pattern of reducing GPA across Level 1, 2 and 3. Mathematics departments in schools 9A, 9B and 8C do not perform above norm for decile at Level 3. The highest performing mathematics departments in the sample are in schools 9E and 8D as they performed well above the norm at all three levels.
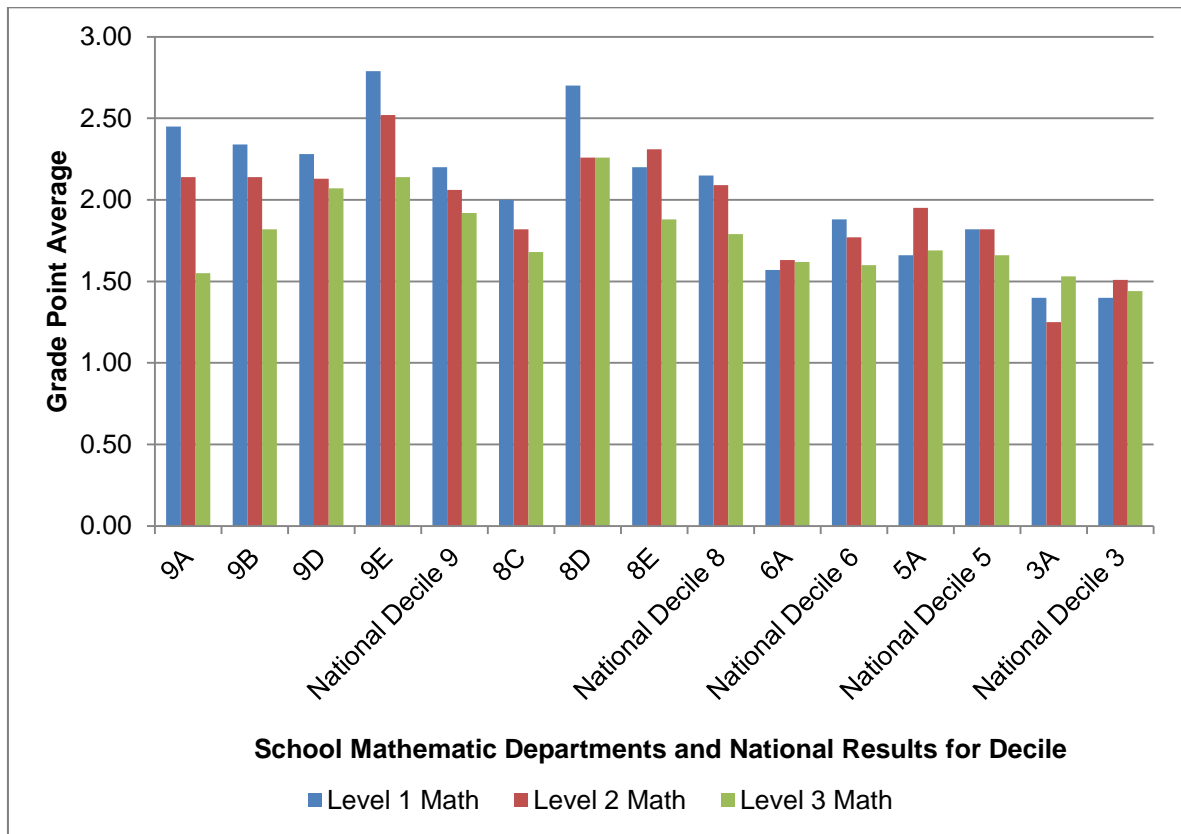
Figure 7.8. Mathematics department performance in Level 1, 2 and 3 in 2010 in Phase 2 schools

The patterns for achievement in the science departments are different to those in English and mathematics. The decile 9 schools have a general pattern of increasing academic achievement at Level 2 and 3, a pattern that is opposite to English and mathematics. The science department in school 3A also has this pattern. The science departments in schools 8C and 6A perform below norm for the decile. Three departments in schools 9E, 8D and 8E consistently perform above the norm for their decile at all levels of NCEA.
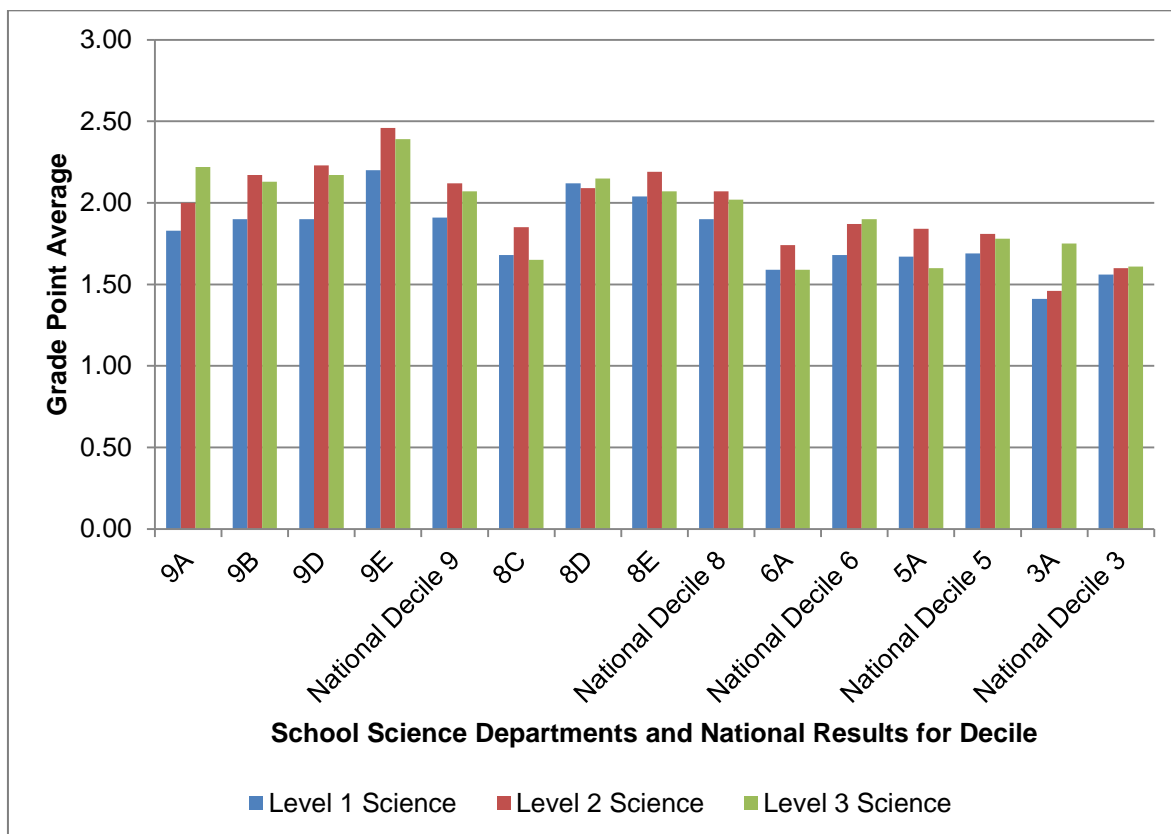
Figure 7.9. Science department performance in Level 1, 2 and 3 in 2010 in Phase 2 schools

Discussion

This section summarises the findings from the analysis of the academic results from the 10 schools in this phase of the research. These findings reflect some of the results of the analysis from the larger sample of the Phase 1 schools but when the Level 2, 3 and scholarship results are included in the analysis and results can be accumulated over time, a more in-depth profile of departmental achievement can be established.

The interaction between school and department effects are an important consideration for this project and have been able to be investigated in further depth in this phase. In schools where all departments are high performers compared to national norms the reasons for this performance could be attributed to a school level effect rather than a department effect. These school effects could include factors that have not been controlled for in this project such as student ethnicity, parent involvement and whole-school leadership. An example of possible whole-school effects is school 9E which was the second highest performing school in the Phase 1 sample of 41 schools. All departments in 9E are performing consistently well in terms of student achievement across three years. This is in contrast to a school such as 8D where the high-performing mathematics and science departments are outperforming the English department in all measures; however, the English department in 8D performs generally at or above the

national norms in most measures apart from the percentage of Merit and Excellence results. This would suggest that students are gaining NCEA level credits but not as many as might be expected at the higher levels of Merit and Excellence. This within-school variation by subject could support Luyten's (1994) claim that teacher effects are the most plausible explanation. For example, in school 8D the mathematics and science teachers appear to be supporting students to gain Merit and Excellence grades to a greater degree than those teachers in the English department. The departments where there are consistently higher results could contain a greater number of higher performing teachers. Good and Brophy (2003) summarise from observational research about effective teaching that teachers whose students achieve academically high grades accept responsibility for teaching their students and if students do not understand a concept they are willing to reteach in a different way because they believe that all students can learn. Whatever the reason for the variance, analysis of academic data provides a rationale for further investigation and indication of sites where high and low performance exists with the same student cohort.

Considerable detail about the extent of department performance or underperformance is described in Appendix F. Departments that appear to be more or less effective than others are highlighted. Using a range of academic measures to detail department performance has enabled a determination of the most effective and least effective departments in the sample. Thomas et al. (1997b) suggest that a stringent procedure for categorising schools in terms of effectiveness is appropriate for identifying and separating schools with the most stable pattern of results over a 3-year period. It is clear that overall NCEA performance provides only a partial measure of effectiveness of some schools, although it remains useful as a summary measure of overall achievement (Thomas et al., 1997b). Appendix F shows that in every department there are changes in student outcomes results year by year. Fourteen (46%) departments perform as expected compared to national norms (not highlighted). Eight (27%) are classified as high performing (red highlighted) and eight (27%) as underperforming (blue highlighted). In the Thomas et al. (1997b) study of 77 schools, only a minority of schools performed both consistently (across subjects) and with stability (over time). School 9E in this sample is the only school that could be described in those terms.

The analysis of results for these 10 schools shows considerable variance between subjects in some schools and variable performance of subjects in different years. This is similar to the findings of Luyton's (1994) study of the stability and variance of subjects across years in over 350 Dutch secondary schools. His work revealed that there are differences in subjects within schools, which are fairly stable themselves, and the

department differences appear to be more important than the general school differences. Luyton claimed that a school may be effective with respect to a certain subject in one particular year but much less effective the next year. In three schools in the Phase 2 sample (3A, 9B, 8C) the performance across departments appears to be relatively consistent, whether it be high, middle or low compared to national norms. In seven schools, the within-school variance is more pronounced such as in 9D, 8D and 6A. Luyton suggests that differences between subjects within schools, appear to be more important than the general school differences so researchers should turn their attention to the functioning of departments within secondary schools. In addition, he states that if researchers want to establish which school and/or classroom characteristics are related to academic outcomes, individual pupil characteristics such as intelligence, previous achievement and family background need to be controlled as they explain a considerable amount of variance in academic performance. To some extent, this issue has been managed when considering the results in this project because academic outcomes for groups of students have been compared to those from schools of similar decile to take into account the socioeconomic background of students.

Thomas et al. (1997b) agree that when three years of data is used rather than the results of a single cohort, school context such as location or size is less significant than individual pupil factors such as socioeconomic and cultural aspects. In addition to decile information, Table 5.2 (page 95) provides some contextual information for all of the Phase 2 schools and provides further detail about the numbers of students in the school for whom English is a second language and the percentage of students from groups statistically at risk of underachieving, such as students of Māori and Pacific Island origin. These student background factors are critical in the discussion of the relative success of departments such as in the English department of school 3A where over 50% of students in classrooms would either have English as a second language and/or be of Māori or Pacific Island descent. Thomas et al. suggest that in English, pupil factors are more significant as the socioeconomic context was shown to have a greater influence on academic outcomes than in other subjects. These factors do not in themselves designate students as academic failures but contextualise the teaching and learning environment for a department and the teachers within it, when comparing their academic results to other similar schools.

Two of the underperforming departments in the Phase 2 sample require special mention in terms of consideration of the school context. The demographic data in Table 5.2 (p. 95) shows school 3A is situated in a deprived urban area of Auckland and the

demographic data shows that 54% of the students in the school are either Māori/Pasifika/refugee or recent migrant students and/or have English as a second language. In this setting, the English department staff will have had to overcome a range of obstacles to ensure their students are performing at or above national norms for their decile. The 3A English department underperforms in all of the student academic outcome measures by not reaching the same results as the national norms, although the deficit is not significant. This case differs to the English department in school 9D where the context is in complete contrast to school 3A (refer Table 5.2, p. 95). School 9D is a single-sex girls' school in an affluent urban area with just 3% of students have English as a second language. It could be argued that there are less barriers for achievement for the students in this school and that the English department results could be expected to be above national norms rather than below them. These examples support the statement of Thomas et al. (1997b) that there is little value in trying to make judgements about effectiveness from school's raw examination data that are often published in the media. They suggest that effectiveness is best seen as a feature that is outcome and time specific. Therefore, identifying results that are significantly different from those expected, taking account of student intake over more than one year in terms of overall performance and in specific subjects is critical (Thomas et al., 1997b, p. 194).

The results of this analysis of the Phase 2 schools show that there is a considerable diversity in academic student outcomes in schools and departments in this group of volunteer schools. There is one school (9E) with three very high-performing departments and therefore very high overall whole-school performance. There are also some high-performing schools (e.g., 9D) with one department where the results are consistently not at the same level as the other departments. There are some departments (8D mathematics) outperforming others in their decile range and the lowest decile school (3A) with considerably disadvantaged pupils generally performing at and above national norms. The next chapter of this thesis provides an analysis of the middle leadership practices in each of these departments in conjunction with student academic outcomes so that relationships between leadership practices and student outcomes can be tested.

Chapter Eight: Analysing the Relationship Between Academic Student Outcomes and Leadership Attributes Results

This chapter outlines the methods used to test the relationship between middle leadership practices and student achievement in the Phase 2 schools. Chapters Six and Seven described the methods used to test the hypothesis that middle leadership practices known to be effective have a positive impact on student academic outcomes. A range of analysis strategies have been employed to enable reliable analysis of the student outcomes data so that it can be compared with the results from the middle leadership questionnaire.

Results

In the first phase of this project the focus on NCEA Level 1 data enabled a comparison of results within and across schools for the same cohort of students. The selection of 10 schools for in-depth analysis of results for the second phase of the project allowed the collection of the Level 2 and 3 data for the 30 departments to be added to SPSS for analysis purposes. The results of this analysis led to a focus in this section on the relationship between NCEA Level 2 and 3 student academic results and the middle leadership scores at department and school level because the initial test (Table 8.1) shows no significant relationship between middle leadership scores and academic results for Level 1 (15-year-olds). The results show a much stronger relationship between middle leadership practices and teaching and learning at the higher levels of the qualification and the percentage of Merit and Excellence grades achieved by students.

Initially, data from the questionnaire responses and the student academic results were combined in SPSS to perform analyses. Table 8.1 shows the relationship between the mean results for each of the middle leadership factors for the whole sample and the student academic results using GPA for NCEA Levels 1, 2 and 3, and the percentage of Merit and Excellence results in 2010. The scores for middle leadership across the sample did not correlate significantly with the student academic outcome data at Level 1. This pattern is the same for the factor *Collegial working environment,* which had no significant relationship with the academic results at qualification level. The factor *Focus on student academic results* was statistically significantly correlated with the results for students at Level 3 ($r = .37$, $p < .05$). There is a statistically significant relationship between the factor *Goals and expectations* and student academic results at NCEA

Level 2 ($r = .39$, $p < .05$) and at NCEA Level 3 ($r = .38$, $p < .05$). There was a statistically significant relationship between the factor *Positive learning environment for students and teachers* at both NCEA Level 2 student achievement ($r = .38$, $p < .05$) and the percentage of Merit and Excellence results in the department in 2010 ($r = .38$, $p < .05$). There was also a statistically significant relationship between factor *Management of resources* and NCEA Level 2 student achievement ($r = .40$, $p < .03$).

Table 8.1

Correlations Between Middle Leadership Practices and Student Academic Outcomes for the Whole Sample of Phase 2 Schools

Middle leadership factors	Correlations	GPA Level 1 2010	GPA Level 2 2010	GPA Level 3 2010	Percentage of Merit and Excellence results 2010
Collegial working environment ($n = 30$)	Pearson Correlation	.13	.27	.12	.21
Focus on student academic results ($n = 30$)	Pearson Correlation	.17	.20	.36[*]	.23
Management of resources ($n = 30$)	Pearson Correlation	.15	.40[*]	.34	.34
Goals and expectations ($n = 30$)	Pearson Correlation	.27	.38[*]	.37[*]	.35
Positive learning environment for students and teachers ($n = 30$)	Pearson Correlation	.33	.38[*]	.31	.38[*]

Notes.

Number of departments = 30

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

In order to establish patterns and relationshilineraetween the middle leadership scores and the academic results, summary measures were required. Using a set of criteria for positive and negative middle leadership practices detailed in the footnote of Table 8.2, departments were classified on a 4-point scale based on the mean scores for each department from the middle leadership questionnaires. The average mean for the sample was 5.5, so departments that achieved a score higher than the average mean were attributed a plus sign and those above a score of 6 were attributed a double plus sign. The departments with a score less than 5.5 were attributed a minus sign and those below 5.0, a double minus sign. The student academic outcome data was then given a positive or negative sign in terms of whether the average results for 2010 for Level 2 and 3 were above or below the national norm for the school decile in that subject. Level 2 and 3 results have been selected for comparison because these were the results that significantly correlated to three of the factors as seen in Table 8.2.

Table 8.3 shows that nine departments (30%) have consistent positive relationships between middle leadership practices and student academic outcomes, less than a third of the sample. Three departments (10%) have consistently negative results because the department leadership has scored less than average across the sample and the academic results at Level 2 and Level 3 NCEA are less than the national norm for decile. Nine departments have scored positively in middle leadership practices but have negative outcomes in terms of meeting the national norm for student outcomes at Levels 2 and 3. Nine departments have scored negatively in terms of middle leadership practices but positively in terms of student academic outcomes. The results of this comparison show that in over half the sample (60%) the results are contradictory and in 40% of the schools there is a relationship between the middle leadership and academic outcome scores. Of the nine departments with a positive leadership score, six had just one minus score against an academic outcome measure and three departments had two minus scores. Three departments achieved below average leadership results but are achieving above the norms for their decile in terms of student outcomes. There were mixed results for the six departments with the very highest average leadership score. Four of those six also gained good student academic outcomes, one has mixed results and the department working in the most challenging circumstances (3A) does not reach national norms for decile for student academic outcomes at either Level 2 or 3. The department members have scored their department leadership as one of the highest in the sample and Appendix F shows that the English department in the school 3A is just 3% lower than the national norm for decile for the percentage of Merit and Excellence results for the department in 2010 and just 2% below norm for decile for Level 3 in 2010. These measures provide an overall summary of the trend but closer scrutiny needs to be paid to the detail of the data and the context for each school in order to understand the extent of the relationship between the two measures.

Eight of the 30 departments (26.6%) have positive scores for all of the summary measures. Three of the schools (9A, 9B and 9E) have two departments each with positive summary measures for both middle leadership and student academic outcomes. Two of the three schools are the same as the three top schools (9B, 8D, 9E) ranked in Table 8.2 for middle leadership. Three schools (8D, 9A, 9B) have received a full set of positive measures for middle leadership shown in Table 8.2 with 9B having two departments with double positives. This result supports the ranking in Table 6.3 (p. 100), where school 9B is the highest ranking school for middle leadership, 8D ranks second highest, while 9A ranks fifth highest. When the student academic outcomes are also used for ranking or sorting purposes the results have some similarities and some differences.

Table 8.2

Middle Leadership Summary Measures Compared to Student Academic Outcomes at Level 2 and 3 Organised by School

School code	Department	Middle leadership score for department*	Department performance above or below national norm for decile at Level 2 in 2010 (+ sign shows above the norm and – sign below the norm)	Department performance above or below national norm for decile at Level 3 in 2010 (+ sign shows above the norm and – sign below the norm)
3A	English	++	-	-
3A	Mathematics	-	-	+
3A	Science	+	-	+
5A	English	-	-	-
5A	Mathematics	-	+	+
5A	Science	-	+	-
6A	English	--	-	-
6A	Mathematics	-	-	+
6A	Science	+	-	-
8C	English	--	-	+
8C	Mathematics	+	-	-
8C	Science	+	-	-
8D	English	++	+	-
8D	Mathematics	++	+	+
8D	Science	+	+	-
8E	English	+	+	-
8E	Mathematics	-	+	+
8E	Science	+	+	+
9A	English	+	+	+
9A	Mathematics	+	+	-
9A	Science	+	+	+
9B	English	++	+	+
9B	Mathematics	+	+	-
9B	Science	++	+	+
9D	English	--	-	-
9D	Mathematics	-	+	+
9D	Science	+	+	+
9E	English	++	+	+
9E	Mathematics	-	+	+
9E	Science	+	+	+

Notes.

++ = average middle leadership score over 6.0

+ = average middle leadership score over the average score of 5.5 and below 6.0

- = average middle leadership score under the average 5.5 but above 5.0

-- = average middle leadership score less than 5.0

Table 8.3 describes subject performance across the 10 schools. The mathematics department in school 8D was the highest ranking mathematics department in the sample for middle leadership and is the only mathematics department to get a full set of positive summary measures. English departments had the largest diversity of results

using these measures. Three departments scored positive indicators in all measures and three departments received all negative measures. No mathematics or science departments received a full set of negative scores using these summary measures. Five science departments have a full set of positive measures.

Table 8.3

Middle Leadership Summary Measures Compared to Student Academic Outcomes at Level 2 and 3 Organised by Subject

School code	Subject	Middle leadership score for department*	Department performance above or below national norm for decile at Level 2 in 2010 (+ sign shows above the norm and – sign below the norm)	Department performance above or below national norm for decile at Level 3 in 2010 (+ sign shows above the norm and – sign below the norm)
3A	English	++	-	-
5A	English	-	-	-
6A	English	--	-	-
8C	English	--	-	+
8D	English	++	+	-
8E	English	+	+	-
9A	English	+	+	+
9B	English	++	+	+
9D	English	--	-	-
9E	English	++	+	+
3A	Mathematics	-	-	+
5A	Mathematics	-	+	+
6A	Mathematics	-	-	+
8C	Mathematics	+	-	-
8D	Mathematics	++	+	+
8E	Mathematics	-	+	+
9A	Mathematics	+	+	-
9B	Mathematics	+	+	-
9D	Mathematics	-	+	+
9E	Mathematics	-	+	+
3A	Science	+	-	+
5A	Science	-	+	-
6A	Science	+	-	-
8C	Science	+	-	-
8D	Science	+	+	-
8E	Science	+	+	+
9A	Science	+	+	+
9B	Science	++	+	+
9D	Science	+	+	+
9E	Science	+	+	+

Notes.

++ = average middle leadership score over 6.0

+ = average middle leadership score over the average score of 5.5 and below 6.0

- = average middle leadership score under the average 5.5 but above 5.0

-- = average middle leadership score less than 5.0

The extent to which effective middle leadership indicates strong academic outcomes at department and whole-school level can be seen in Figure 8.1 and 8.2. Figure 8.1 shows a bar for each of the 30 departments. This bar represents a combination of the middle leadership score for the department at the bottom (blue), with a red section at the top of the bar that represents a student academic outcome score for that department. This score represents an average GPA from Level 2 and 3 NCEA results for that subject in 2009 and 2010. In Figure 8.2 the departments for each school are placed adjacent to each other so the extent of the within-school variation in middle leadership and student outcomes can be identified. Six departments from four different schools (9B, 8D, 9E, 9A) gain a score over 8 with the lowest department in the sample scoring 5.04. Four schools have departments ranked in the bottom seven with a total score of less than 7. A direct comparison of departments and schools using student outcome measures is complicated by student intake differences because lower decile schools would not expect the student outcomes to match that of the general population or those in the higher decile schools. Figure 8.1 and 8.2 show that when middle leadership scores and academic outcomes data are combined, the bottom-ranked departments are not always in the lowest decile schools. The third bottom-ranked department (9D) is a decile 9 school and the English department in school 3A is ranked in the top half of the sample.
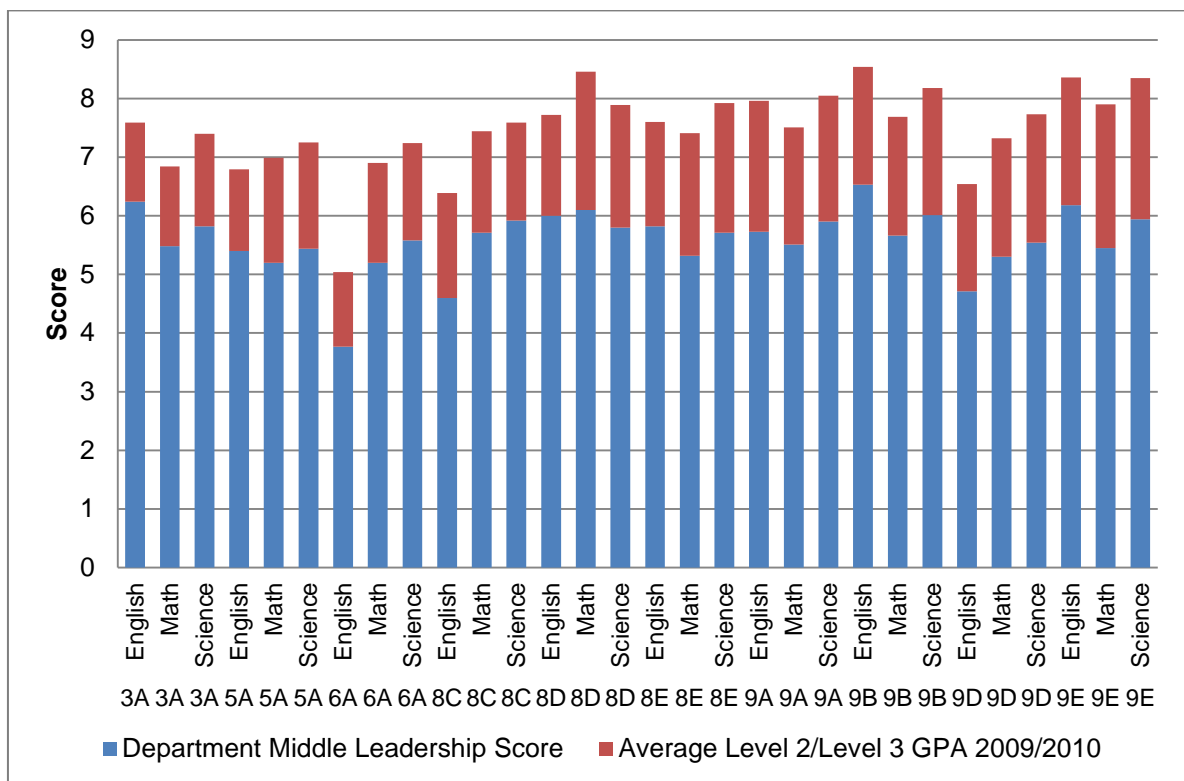

Figure 8.1. Department middle leadership combined with student academic outcomes for Level 2 and 3

This combined score derived from the department mean middle leadership questionnaire and the averaged Level 2 and 3 GPA is useful to compare within-school variation on both middle leadership and student academic outcome measures simultaneously. Some schools have considerable variation between departments and others have overall high performance in both measures. The clearest within-school variation can be seen in school 6A where the English department is underperforming in both measures in relation to the rest of the school but also the whole sample. School 8D has interesting within-school variation. The mathematics department is ranked second highest in the sample for middle leadership and is the highest ranking mathematics department. The student academic outcome results in this department are considerably higher than those of the English department. The middle leadership scores for the three departments in school 8D are similar but when the student outcome results are added the within-school variation is much more pronounced.

The extent of the across-school variation can be seen in Figure 8.2. This graph uses the same measures as Figure 8.1 but at whole-school level so that middle leadership and student outcome results for the English, mathematics and science departments in the school have been combined and averaged to produce a whole-school score. The blue section of the bar at the bottom is the whole-school middle leadership score and the red section of the bar at the top represents student outcomes. In this graph, three schools (9E, 9B, 9D) have an overall score over 8. These are the three top-ranked schools for the middle leadership factors (Table 6.3, p. 100) and when the student academic outcome scores are added, the ranking of these three schools is not changed. School 6A remains the lowest performing when these measures are combined and despite 3A having a lower GPA score than the rest of the sample, which would be expected due to its lower decile status, the higher middle leadership score has ranked this school sixth out of the 10 schools.

The summary measures for the sample shown in Table 8.2 and 8.3 show department effectiveness compared to academic norms for decile but the results shown in Figure 8.1 and 8.2 do not take student socioeconomic background into account by controlling for decile. Overall, the results represent a snapshot at one time in 30 departments in 10 schools. Although the student academic outcome data was been collected over a 3-year period it is important to recognise that change in school and department effectiveness is likely over time periods of more than 1 or 2 years due to changes in staff, in pupil intakes and in ethos. Schools may also change rapidly as a result of a professional development intervention or a new principal (Sammons, 1996). The student academic outcomes data are however more stable than the middle leadership

questionnaire results. The overall 75% response rate to the questionnaire (Table 5.1, p. 93) and the changes in staff leadership responsibilities that occur in schools means the results give an indication of within-school variability at one point in time over a cross section of schools.
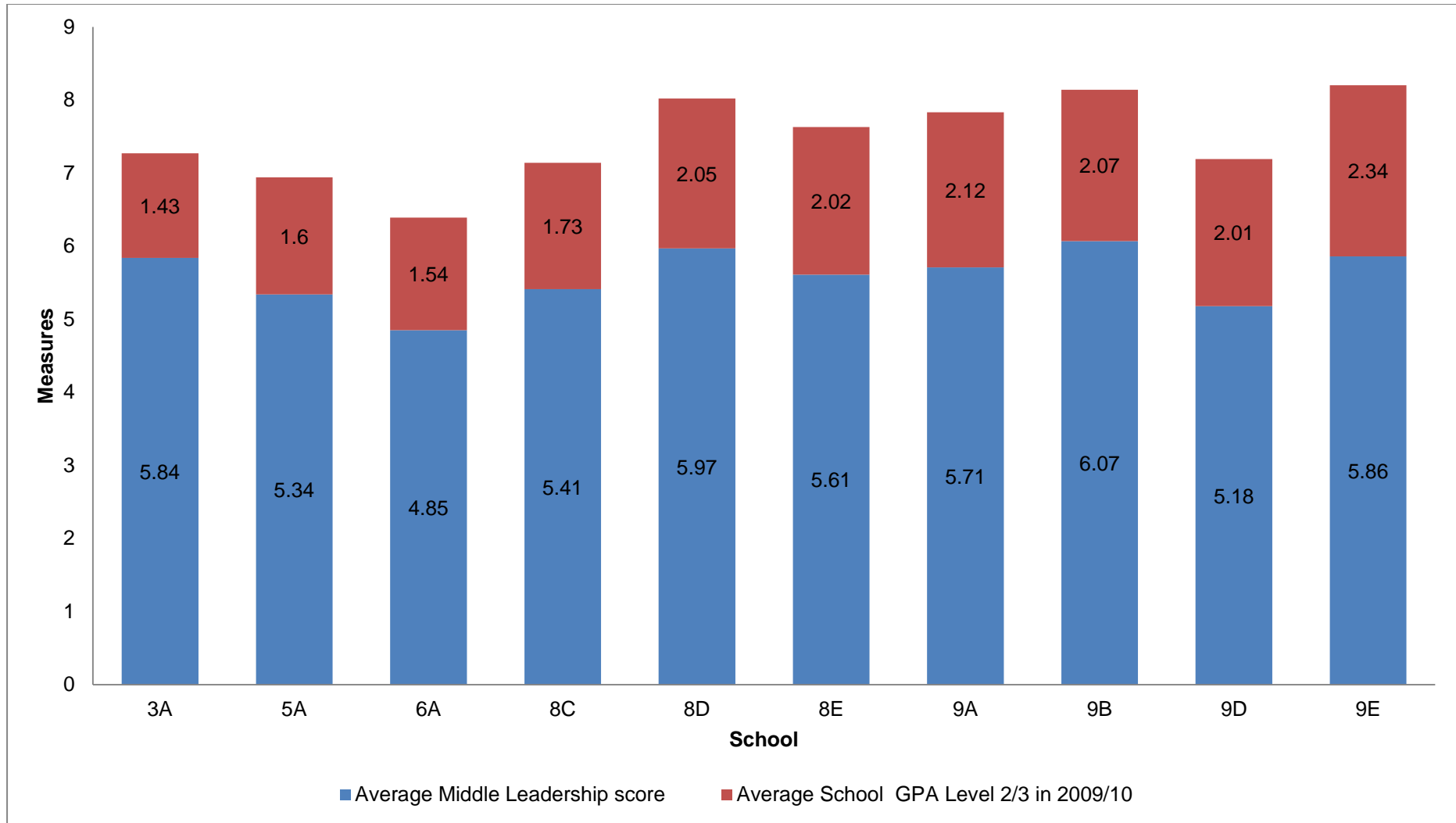
Figure 8.2. School middle leadership scores combined with average GPA across English, mathematics and science Level 2 and 3 in 2009/10

Discussion

The results from comparing the middle leadership scores and the student outcomes data produce some key ideas for discussion. The summary measures in Table 8.3 suggest a pattern of decile 8 and 9 schools having higher performing departments when effectiveness of middle leadership and student academic outcomes are summarised and compared. The specific leadership factors that had significant correlations with student academic achievement at NCEA Level 2 and 3 have helped develop a deeper understanding of the relationship and are tested in a simultaneous regression later in this chapter. The variable middle leadership and student academic outcomes within schools and across subject are also discussed in this section.

The results indicate that English, mathematics and science middle leaders and teachers in the higher decile schools in the sample are more likely to report having effectively led departments. The higher decile schools in this Phase 2 sample also show a pattern associated with student academic outcomes at NCEA Level 2 and 3 that are above the national norm for decile. Of the 21 decile 8 and 9 departments shown in Table 8.3 (p. 131), just five of these departments have a middle leadership score under the average for the sample. Three of those five departments are mathematics departments where despite the negative reporting of middle leadership, the student academic outcomes are still above the national norm. There are just two English departments in the high decile schools where the middle leadership scores are well below average and the student academic outcomes are also below the norms. Figure 8.1 shows that the highest performing departments when middle leadership and GPA data are combined are either decile 8 or 9. Silins and Mulford (2002) argue that schools add value to student achievement against a background of family support and socio-economic status so it could be claimed that these departments are advantaged by the students who are enrolled at the school. However, the argument that higher decile schools have more students with background factors that support strong academic achievement is mitigated in this study because the department results outlined in Table 8.3 (p. 131) are measured in relation to norms for decile and there are differential results for departments within one school. Harris and Muijs (2005) argue that teacher effectiveness is intrinsically related to the overall effectiveness of the department and the way it is led. Indications from the results of this study are that higher decile schools attract more effective teachers and leaders to some departments. However, the higher decile schools are not immune to having departments where the middle leadership is reported to be ineffective.

The analysis detailed in Table 8.1 (p. 128) showed that there were four middle leadership factors that had a significant correlation with student academic outcomes at Level 2 and 3 of the NCEA qualification and the percentage of Merits and Excellence grades gained. These were the factors, *Focus on student academic results, Management of resource, Goals and expectations and Positive learning environment for students and teachers.* These factors are supported by the literature that describes effectively led departments as those that have a central focus on teaching and learning and a student-centred approach (Bolam & Turner, 2003). When middle leaders are concentrating their efforts on departmental leadership that will directly impact positively on students, the academic results in the critical last two years of schooling are improved.

In contrast, the factor *Collegial working environment* had no significant relationship with the academic results at qualification level. This finding is important as it adds a different perspective to the existing literature. Bird and Little (1986) describe norms of collegiality as those that are realised in staff discussion of teaching practices, mutual observation, cooperative work in planning and selecting or designing classroom teaching resources and joint learning. Timperley and Robinson (1998) discuss collegiality in terms of the extent to which the conditions for effective problem-solving can occur in a school or department setting. These definitions are not well aligned with the items that grouped into the collegiality factor for this study, which are very focussed on how staff feel about their working environment, the culture of working together as a team and staff feeling recognised. However, participants in the current study were asked to rate the extent to which middle leaders ensured robust discussions about successes and challenges occurred and also the level of effectiveness that conflict and feedback about professional matters was managed. These middle leadership practices could be argued as being closely related to promoting effective problem-solving strategies as described by Timperley and Robinson (1998). The practices that Bird and Little (1986) describe are embedded in the questionnaire items of this study that grouped into the factor *Positive learning environment for students and teachers*, which does have a positive relationship with student academic outcomes. Timperley and Robinson (1998) assert that staff need to access or have the expertise to reflect critically to problem solve effectively. Future follow-up work would need to decipher exactly which aspects of collegiality impact positively on student academic outcomes and test the questionnaire items that have a negative relationship to student results. Hargreaves and Dawe (1990) argue that teachers and leaders are often encouraged to participate in a kind of *contrived collaboration*, which involves teachers grouped together in a way that is administratively designed to smooth the path of externally imposed innovation. They

argue that bureaucratically driven systems of collaboration are often hindered by the hierarchical relationships that undermine the trust necessary to engage in a culture that is truly reflective and critical. Hargreaves (1994) argues that *comfortable collaboration* among teachers occurs where staff might be happy and content but students make little progress. Hargreaves (1999) does go on to highlight the research on collegiality that shows that when people work in collaborative cultures, goal-centred and inquiry-oriented discussions about improving teaching and learning occur that have the potential to subsequently improve student outcomes. This statement is supported by the results of this project because when the teachers and middle leaders focus their collegial activity on students, rather than on themselves, there is a direct link to positive academic outcomes. This is evidenced in the statistically significant relationship between the factor *Positive learning environment for students and teachers* and improved academic achievement at NCEA Level 2 and the percentage of Merit and Excellence results achieved by students. The positive correlation for student academic outcomes at NCEA Level 3 and middle leadership that promotes a focus on student academic results also supports the existing literature on effective middle leadership in secondary schools as a critical force in improving the teaching and learning within the department (Bolam & Turner, 2003).

When considering the summary measures from a subject perspective where leadership and student academic outcomes are combined (Table 8.3) there are disparate results that support the literature describing subject department differences. Subject matter is often intertwined with membership in a department and collegiality and professional culture are often tied to departments (McLaughlin & Talbert, 2007; Siskin, 1994). Siskin's (1995) study in 16 schools considering the departmental structure as a context for high school leadership showed systematic and persistent differences at department level in department members' responses to questions around items she describes as being related to collegiality. In Siskin's (1995) study, the within-school differences were as strong as across schools and were revealed in survey responses, observation data and interviews. Siskin states that "Teachers in these different departments inhabited strikingly different contexts and professional climates. The 'schools' they saw themselves working in reflect substantially different sites" (p. 610).

The summary measures for the subjects English, mathematics and science within this project are variable. The English departments had the largest diversity of results with three departments having all positive summary measures and three English departments with all negative summary measures. Although only one mathematics department had a full set of positive summary measures, no mathematics or science

departments had a full set of negative measures for middle leadership and student academic outcomes below the norm. Six of the 10 mathematics departments reported middle leadership attributes below the average for the sample. Half (5) of the science departments in these Phase 2 schools had a full set of positive measures. This variation in reported effectiveness of leadership within a department is supported by the literature that describes the differences of colleagues who work in different curriculum areas. The differences reported in this project could indicate that the perceptions of leadership practices are variable for teachers in different disciplines. For example, with respect to the comparison of goals between English and mathematics teachers, Grossman and Stodolsky (1994) found they differed in the goals they held for students and instructional practices. English teachers scored significantly higher on personal growth goals (self-esteem) and human relations goals. English teachers in their study strongly endorsed instructional approaches such as providing opportunities for personal expression and personalised learning while mathematics teachers were significantly more supportive of grouping students in class based on prior achievement. Siskin (1991) found mathematics teachers were more systematic in their tracking and assessment of student learning. Siskin (1995) and Grossman and Stodolsky (1994) strongly assert that secondary school departments are powerful subject subcultures and that teachers identify very strongly with the collective thinking of their discipline. Siskin (1995) suggests that specialised ways of thinking and doing would be evident in all aspects of department decision-making, teaching and ideology. She states that middle leaders gather and dispense resources, make and mediate policies and bring in differing disciplinary perspectives that have far-reaching consequences. The evidence in this study suggests that subject differences occur in all school types and have no relationship to the socioeconomic status of students being taught. Mathematics teachers and leaders report middle leadership as less effective but in some departments this does not translate to student academic outcomes that are below the norm.

Emerging patterns between key middle leadership factors and student academic outcomes

In order to determine whether leadership practices predicted Level 2 and Level 3 NCEA results, multiple regressions were run. In each of these regressions, decile was controlled, so that it was not a confounding variable in the equation. For the first hierarchical linear regression, the first equation included decile predicting achievement (percentage of Merit and Excellence results) while in the second equation decile was entered first and then the mean score for teachers and middle management for each of

the middle leadership factors was entered in the next step. This test was used to determine to what extent the *Collegial working environment*, *Focus on student academic results*, *Management of resource, Goals and expectations* and a *Positive learning environment for students and teachers* predicted achievement in schools when decile was controlled, first at Level 2 NCEA (GPA *M* = 1.95, *SD* .34) and then for Level 3 NCEA (GPA *M* = 1.87, *SD* .29).

Means and standard deviations for the middle leadership factors are presented in Table 8.4. This table reports the R^2 and R^2 change while the text includes the adjusted R^2. *Collegial working environment*, a *Focus on student academic results*, *Management of resource, Goals and expectations* and a *Positive learning environment for students and teachers* predicted achievement at Level 2 NCEA. For the first equation that included only decile predicting achievement, the ANOVA was significant, $F_{(1, 28)} = 46.20$, $p < .001$, accounting for 62% of the variance in achievement. When middle leadership practices were added into the equation, the result was still significant, $F_{(5, 23)} = 9.03$, p $< .001$, adjusted $R^2 = .84$, thus accounting for 84% of the variance in the NCEA Level 2 academic results, a large effect. The Beta weights in Table 8.5 shows that both *Collegial working environment* ($p = .02$) and a *Focus on student academic results* ($p = .001$) negatively predicted achievement. *Management of resources* was a positive predictor of achievement ($p = .001$) along with a *Positive learning environment for students and teachers (p = .06)* and *Goals and expectations* ($p = .003$).

Further, leadership practices also predicted Level 3 NCEA controlling for decile. When decile alone was entered into the equation, the ANOVA was significant, $F_{(1, 28)} = 23.69$, $p < .001$, accounting for 46% of the variance in achievement. When middle leadership practices were added into the equation, the result was still significant, $F_{(5, 23)} = 3.56$, $p = .02$, adjusted $R^2 = .62$, thus accounting for 62% of the variance in the NCEA Level 3 academic results, a large effect. The Beta weights show that a *Collegial work environment* negatively predicted student achievement ($p = .003$) while *Management of resources* was a positive predictor of student academic achievement ($p = .01$) as was a *Positive learning environment for students and teachers* ($p = .05$) (see Table 8.5).

Table 8.4

Mean and Standard Deviation for Regression

Factors	*M*	*SD*
	1.9357	.34493
Collegial	5.53	.607
Results	5.47	.518
Resources	5.48	.922

	6.9	.597			
Goals	6.9	.597			
Learning	5.77	.438			

Table 8.5

Simultaneous Regressions for Teacher Leadership Practices Predicting Achievement at Levels 2 and 3 NCEA.

Variable	B	SEB	β	R^2	ΔR^2
Predicting **Level 2** NCEA achievement					
Step 1 (*n* = 29)					
Decile	.14	.02	.79	.62	.62
Constant	.925	.154			
Predicting Level 2 NCEA achievement				.87	.25
Step 2 (*n* = 29)					
Decile	.12	.01			
Collegial	-.27	.10	-1.13**		
Results	-.43	.98	-.04		
Resources	.23	.05	.72*		
Goals	.29	.09	.35		
Learning	.22	.10	.52†		
Constant	.564	.410			
Predicting **Level 3** NCEA achievement					
Step 1 (*n* = 29)					
Decile	.09	.02	.67	.46	.46
Constant	1.145	.155			
Predicting Level 3 NCEA achievement				.69	.24
Step 2 (*n* = 29)					
Decile	.08	.01	.57		
Collegial	-.43	.13	-.90		
Results	-.01	.12	-.02*		
Resources	.20	.07	.64		
Goals	.07	.11	.16		
Learning	.29	.14	.44		
Constant	.458	.535			

Notes.

The coefficients in the β column are standardised while those in the *B* and *SEB* columns are unstandardised.

† $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

The multiple regression for Level 2 NCEA student achievement shows that decile of school can account for 62% of the variance in student achievement and when middle leadership practices are included in a simultaneous regression, 84% of the variance can be accounted for. Decile of school was a less important predictor for Level 3 NCEA student academic achievements (46%), but as with Level 2 when middle leadership practices were added the ability to predict student achievement rose to 62%. Therefore a combination of school decile and middle leadership practices were a good overall predictor of student academic achievement at Levels 2 and 3 NCEA.

The regression analysis shows that middle leadership factors relate both positively and negatively to student achievement at Levels 2 and 3 NCEA, at all decile schools in the Phase 2 study. The Level 1 NCEA results were not used in this analysis as they had already been shown to have no relationship with middle leadership practices (Table 8.3). The scale *Collegial working environment* had a high Cronbach's alpha reliability (Appendix D) of 0.95, however as a factor it has a negative relationship with student achievement at Levels 2 and 3 NCEA. The items within this factor related to collegiality are focussed on the social and professional working environment of teachers. Teachers and middle leaders in departments with high collegiality scores report that their department is a collegial place to work, where staff support each other and function well as a team. A department *Focus on student academic results* was also a negative predictor for academic results for Levels 2 and 3 NCEA. This factor embodies practices such as teachers setting academic goals based on students' achievement data and having regular routines around assessment and feedback. The regression analysis shows that these two factors of department leadership characteristics have a negative relationship to student achievement in the senior academic years where the stakes are high for students wanting to gain entrance to university or further training and employment.

In contrast, the factor *Positive learning environment for students and teachers* had a positive relationship with student academic achievement at both NCEA Levels 2 and 3. This factor included items where teachers and leaders rated the extent to which students experienced the classroom environment as supportive of their learning and received high-quality feedback that helped improve their work. The factor *Goals and expectations* was a positive predictor of student academic outcomes at Level 2 NCEA. This factor included items related to clear and measurable department goals that linked to school goals and teachers understanding how their work in the classroom contributed to the department and school goals. This factor also included an item describing middle leaders that make time to observe department teachers at work in their classrooms. The factor *Management of resources* was also a positive predictor of student academic achievement at Levels 2 and 3 NCEA. This factor measures how equitably and strategically resources are allocated in a department, how they are targeted to support student learning and teaching goals and the extent to which middle leaders are involved in the recruitment of high-quality staff.

Discussion

The results from the regressions show that when school decile or student socioeconomic status is controlled an additional predictor of student academic results in

English, mathematics and science at Level 2 and 3 NCEA are the leadership practices in a department. The factors that have a negative impact on student academic results are those that relate most directly to the collegiality and social structures of the colleagues within the department. All participants in this study would experience a level of collegiality because they all teach the same subject in the same department and have a substantial set of work activities in common within the same school context (Bidwell & Yasumoto, 1999). To some extent the results from this analysis run counter to some of the existing theory. Collegiality as a term can have multiple interpretations but Campbell and Southworth (1992) define collegiality as teachers working together and sharing power through consensus decision-making processes within a cohesive school culture. Bidwell and Yasumoto (1999) describe networks of collegial relationships that are pathways for communication, influence and sanctioning. They assert that these collegial relationships allow individual and collective solutions of instructional problems in a school, allowing teachers to consult and work with colleagues to adapt and enforce occupational norms. Effective departments have been described in the literature as highly collegial and collaborative work places for teachers (D. Hill, 1995). Increased collegiality is considered desirable because it improves teacher professionalism and accountability (Firestone, 1996; Hargreaves, 1991). Harris et al. (1995) describe a striking feature of the effective departments they studied was the emphasis on collegiate styles of management. The researchers found departments effective in improving student academic outcomes were marked by the constant interchange of professional information at both a formal and an informal level. They described the leaders of these departments as those that were skilled at managing interpersonal relationships, exhibited trust in their colleagues and encouraged teachers to take on responsibilities that would provide them an opportunity to lead the whole department in a particular area.

New Zealand researchers have also tested the link between a collegial school environment and outcomes for students in a secondary school context. Timperley and Robinson (1998) state that the assumption that collegiality produces better outcomes is increasingly recognised as problematic. They support Hargreaves' (1984) assertion that when teachers work together, the result may be the pooling of ignorance rather than of expertise. Timperley and Robinson's (1998) study of problem-solving in a secondary school showed that collegiality was enhanced only when quality information is available and coherent understandings could be formulated within the group through the surfacing and testing of assumptions. They observed a form of collegiality that enabled staff to reach consensual decisions that they privately believed had a low probability of implementation (Timperley & Robinson, 1998, p. 625).The teachers working in Lipman's

(1997) study to develop solutions to the problem of underachievement for African American students spent increasing amounts of time discussing the deficits of their students, rather than examining their own practice and how they may be contributing to the problem. Collegial processes are limited when teachers fail to recognise the limitations of their knowledge (Timperley & Robinson, 1998) and seek the expertise they require to challenge assumptions and ways of working (Harris & Muijs, 2005). Bidwell and Yasumoto (1999) would support this claim as they state that department members acquire a body of knowledge of empirical information about teaching techniques that work relatively well or relatively badly under specific circumstances (p. 236). Data they gathered about the collegial focus of staff from 13 public high schools in Chicago showed that considerable polarising tendencies in departments arose especially in relation to issues such as the access to scarce resources or assignment to desirable curricular levels (p. 249). The following quote from a principal in a Harris and Muijs (2005) case study focussed on teacher leadership summarises the key idea:

> There's a difference, isn't there, between collegiality and conviviality. I've worked in places where people are terribly matey, and we buy each other cream buns and go for drinks, but having a climate where people can be critical of each other, hold each other to account is different. (p. 110)

In summary, researchers (Hargreaves, 1984, 1991; Hargreaves & Dawe, 1990; Lipman, 1997; Timperley & Robinson, 1998) argue that collegiality is a positive influence on student outcomes and school performance when it involves increased professional discussion, critical reflection and problem-solving. The opposing evidence from the Harris et al. (1995) study would suggest that there are some features of a collegial working environment worthy of specific and further investigation. The items that were grouped within the collegiality factor in this study all related to participants' views on the effectiveness of department leaders to ensure the department functioned well. Participants rated items that described a collegial and supportive environment, including a culture of professionalism, feedback and advice. These items also included aspects of a 'team' approach with shared opportunities for leadership (Appendix D). The results from this project show that the collegial culture within a department can be functioning well from the perception of department members and will no doubt mean that the working environment for the adults is positive but having these things in place does not guarantee a positive impact on student academic outcomes. Further clarification and definitions of the specific features of collegiality that have direct relationships with student academic outcomes could be probed in larger scale qualitative studies.

A further result from this regression that appears to counter much of the existing theory is the factor related to a *Focus on student academic results.* This factor was a negative predictor for academic results for Levels 2 and 3 NCEA and yet there is considerable literature that would argue middle leadership practices that are included in this factor are consistently linked with improved academic outcomes for students. Areas of practice such as teachers setting academic goals for students based on data and having an academic results focus is negatively correlated in this study with student outcomes at NCEA Level 2 and 3. This finding is interesting when considering the results of Harris et al. (1995) who in a study of effective departments in secondary schools found that in contexts where departments were effective, senior management teams scrutinised the assessment and test results of students at department level. Middle leaders knew that they were being held accountable for the results in their subjects, but in accelerating departments this was not viewed as a threat but a justifiable and necessary way of improving department performance. Departments that were effective in the Harris et al. (1995) study modularised assessment and found it useful in diagnosing pupil strengths as well as areas that required development. Their study asked questions that were different from those used in this thesis, particularly in the area of assessment related to homework. They found that effective departments had clear routines for the setting and marking of homework, and this homework often involved or had the potential to involve parents. Harris et al. found that a distinctive feature of effective departments was the care and attention paid to assessment, including detailed and up-to-date record keeping. In the current New Zealand secondary school context, having efficient and effective record keeping is critical for department and whole-school performance and systems and processes are regularly scrutinised by government agencies. Further probing of the specific aspects of practices related to assessment and an academic results focus that can be found in effective departments needs further scrutiny and examination.

The factors that have a positive impact on student academic success at NCEA Levels 2 and 3 are those that are most directly related to student and staff professional learning and the targeted resources available with teacher expertise that creates a positive learning environment for both students and teachers. The factor *Goals and expectations* was a positive predictor of student academic outcomes at Level 2 NCEA. In the Harris et al. (1995) study, all of the effective departments were marked by a clear sense of vision that largely emanated from and was propagated by the middle leaders. The vision for the department embraced the nature of the subject and how it should be organised for teaching purposes. Department meetings had a clear purpose, were linked to the goals and were sometimes used for professional development. In a 2006

case study research project in three large Auckland secondary schools, Feist (2008) found that faculty heads' interactions with staff developed out of a shared professional ethic with the key focus on teaching and learning. The evidence from the data in her study showed that collegial relationships developed from robust discussions and generating shared understandings about classroom practice.

The factor *Management of resources* was a positive predictor of student academic achievement at Levels 2 and 3 NCEA and this result also supports the existing literature. Effective middle leaders have been shown to organise key resource elements that directly impact on teaching and learning in a positive way. In the study completed by Harris et al. (1995), all of the effective departments had detailed and agreed schemes of work that were consistent with the mandated curriculum, easily accessible, detailed and that provided clear guidance to all staff. The departments in the study were not particularly well resourced but the resources they had were well targeted to purchase items that students could access in the classroom to enhance and improve their learning. The optimum allocation of human resources was also exemplified in this study (Harris et al., 1995) by the use of support staff and technicians who enabled teachers to focus on the core business of teaching their students as their priority. Well-established systems and processes for resource management, dissemination and sharing made teaching in these departments more successful and catered for students' individual learning styles and met their learning needs. With good resource allocation, agreed schemes of work, record keeping and systematic reviews department members had the scope to develop their individual capacities and strengths. In a comparative qualitative research project that examined links between middle leadership and learning in four schools in England and New Zealand (Fitzgerald & Gunter, 2006), it was found that the investment of resources that directly impact on the leadership of learning and the creation of a learning community cannot be overestimated. Fitzgerald and Gunter (2006) asserted that when leaders are prioritising resources a significant factor in their decision-making must be how to best advantage student learning opportunities. In the Harris et al. (1995) study, there was distinctly low staff turnover in the effective departments, so it would seem when the material and human resources are managed well students can expect a consistent and stable work force within a department, which is a feature of effective schooling.

The factor *Positive learning environment for students and teachers* had a significantly positive relationship with student academic achievement at both NCEA Levels 2 and 3. Central to this concept is the creation of a learning environment that encourages teachers to make a difference to student learning and understanding both inside and

outside the classroom context. There are many references in the existing schooling effectiveness literature that support the strong relationship with this factor and academic success. The results of this study support the ongoing investigation of the way middle leadership practice at department level can influence the quality of the teaching and learning. The current literature highlights some of the characteristics of middle leadership that appear to have a positive impact on the teaching and learning. Bolam and Turner (2003) stress that middle leaders need to model good teaching practice themselves in order to communicate a vision about the subject and how it should be taught. In the case study work of Harris and Muijs (2005), they describe a shared vision for the department as a key component of successful teacher leadership. They directly report a middle leader's comment, "We are all very aware of where the school is going, we are consulted on the strategic plan, and we know we are responsible in our particular area to help the school meet its vision" (Harris & Muijs, 2005, p. 110).

The ability of middle leaders to communicate about and influence the quality of the teaching is exemplified in the study by Harris et al. (1995). In their study, they observed middle leaders encouraging teachers to develop clear routines and practices within lessons that included structured lessons and regular feedback with a strong pupil-centred ethos (p. 297). In the departments they studied they found middle leaders and teachers had a high level of commitment to developing their own professional skills in order to activate engaging classroom environments.

A further key feature of department leadership that supports a positive learning environment for students is a focus on professional reflective inquiry into classroom teaching practice. Hargreaves (1999) argues that collaborative cultures that are focussed on goal-centred inquiry orientated discussions can lead to the improvement of teaching and learning. This is supported in Hill's (1995) case study where he describes reflective practice about teaching and learning as the effective aspect that gives the department "turbo power." Bird and Little (1986) describe staff discussions of teaching practices, mutual observation, cooperative work in planning and developing curriculum materials and teaching content as "norms of collegiality." Hill (1995) describes teachers that actively contribute to professional learning communities at school and district level as those that are able to broaden their focus and actively contribute with more strength and knowledge at the department level. A study by Fitzgerald and Gunter (2006) revealed that where middle leaders were regarded as professionals and encouraged to engage in critical reflection and take on extra responsibilities, they flourished. This positively associated factor that focuses on student learning underpins the practices that relate directly to classroom practice and the learning and development of both

students and teachers. These practices are more important in making a difference to student outcomes than those factors that are more closely linked to teachers' professional and personal happiness and well-being in the work place.

The regression analyses completed as part of this thesis in linking the NCEA results at Level 2 and 3 to middle leadership practices revealed that not all aspects of leader and teacher practice can be positively associated with student academic outcomes. These findings suggest that there are areas for further investigation and some results substantiate the existing middle leadership and within-school variation theory. In summary, the results from this project that are consistent with the existing relevant research are that when middle leaders ensure a positive learning environment, develop goals and expectations and manage resources to the benefit of students there is a positive correlation with student academic outcomes for 16- and 17-year-olds. Counter to this is that when middle leaders focus their efforts on developing a culture of collegiality, this by itself not impact positively on student academic outcomes at the senior secondary school as the 'collegial' factor was a negative predictor of student outcomes, as was a focus on student academic results. These negatively associated factors appear to exist in departments with poorer academic outcomes. These factors are described in some of the existing literature as positive but in previous studies have not been directly linked to academic data as they are for this project. The results show, that particularly in mathematics the reported absence of the collegial middle leadership characteristics will not have a negative impact on academic outcomes at the senior level of the secondary school. The literature would suggest these characteristics could have an interdependent relationship with some of the key features of middle leadership that do impact positively. It is worth noting that three of the items in the questionnaire that grouped within the collegial factor relating to teamwork and staff supporting each other were added to the questionnaire after it was trialled. This was due to the participant feedback that there were not enough questions that asked about the collegial working environment in the department. Further work is needed to eliminate some of these items from the questionnaire and strengthen them in terms of the concept of the department as a professional learning community. The results revealed that there was a mismatch between the beliefs that were held by the practitioners who trialled the questionnaire, the existing theory on collegial relationships that make a difference to teaching and learning and the findings from this project. A further qualitative study would be required to gather evidence that could develop further knowledge about the findings of this regression analysis.

This chapter set out the multiple analysis strategies employed in order to explore the quantitative relationships between the middle leadership and student achievement data from the three departments across the 10 schools. The findings have been discussed in relation to the extent to which they substantiate or contradict the existing middle leadership literature. The next chapter will discuss the key findings from all phases of the study and the implications for further research and professional development.

Chapter Nine: Conclusion

This chapter outlines the major findings of the study, details its limitations, and describes the further research and professional learning and development opportunities afforded by this project.

Major findings

The research question that underpinned the initial stages of this project was about the extent of variation in student academic outcomes between the English, mathematics and science departments within and across a sample of 41 Auckland secondary schools. By using data from the same cohort of 15-year-old students across subjects and within one year group, the differential performance of departments could be established. The sub questions related to the patterns of student academic achievement within a school and across three years to determine whether the results were consistently variable or stable by year. The data collected also provided an opportunity to consider the performance of subjects across schools that serve students with diverse socioeconomic and cultural backgrounds. The use of school decile to control for student socioeconomic background has been a key feature of the project and has played a large part in the interpretation of the findings. Variation of academic results within and across similar schools was confirmed in the first phase and the subsequent hypotheses that the role of middle leadership plays a part in the differential performance of students across subjects and departments was then tested. When the middle leadership practices in a subset of schools were investigated, it was found that some aspects of the department leadership role were rated as more effective than others and that differences in student academic results in the senior secondary school correlated with middle leadership practices.

Phase 1 findings

The analysis of student academic outcomes in English, mathematics and science, using whole-school measures revealed that generally schools that draw on students from the highest socioeconomic backgrounds have better results than those that serve the most socially disadvantaged students. When overall school achievement is calculated across Level 1 English, math and science departments, higher decile schools (9 and 10) always rank in the top five. The bottom five ranks are occupied by decile 1 and 2 schools. At NCEA Level 1, there is a statistically significant difference between the achievement of the lowest decile (1, 2, 3 and 4) and highest decile (8, 9 and 10)

schools. There was no statistically significant differences between the high- or low-decile schools and the mid-decile schools when comparing the results for 15-year-olds.

This study has produced some interesting findings regarding predictors of school and subject performance in relation to the overall socioeconomic status of the students in the school. School decile appears to make a difference to academic results in some subjects more than others. Subject departments in mid-decile schools for 15-year-olds had results that were comparable to departments in other mid-decile schools whereas variation between subjects was more pronounced when comparing the higher and lower decile schools. Students attending a decile 5, 6 or 7 school have a statistically significant chance of getting similar results in English, mathematics and science to those of their peers in higher or lower decile schools. The results show that the statistically significant differences in mathematics and science department results in schools with different deciles are larger than they are in English. There was no statistically significant difference in NCEA Level 1 academic results in English unless a student attended a decile 8, 9 or 10 school whereas in mathematics and science students would have a greater chance of success if they attended a decile 6, 7, 8, 9 or 10 school.

There is a considerable range of performance of schools within one decile. This is consistent with the Sammons et al. (1997) study of inner city London schools where schools were shown to be differentially effective for students from low-income families. In this study, schools in the sample perform well above and below national norms for their decile. Neither high- nor low-decile schools were immune from lower performance levels (Bendikson et al., 2011, p. 295) and schools regardless of their socioeconomic status or decile can impact positively on student academic outcomes. The results for the low-decile schools are concerning because many of these schools have high proportions of attendees who come from socially disadvantaged backgrounds. Previous research in these low-decile schools show that Māori and Pasifika students are overrepresented in the statistics for students performing below expectations (McNaughton et al., 2011) and Pacific students are only half as likely as their non-Pacific peers to gain University Entrance (UE) (Statistics New Zealand & Ministry of Pacific Island Affairs, 2010, p. 39). Although education outcomes for these minority groups are improving steadily, Māori and Pacific students are performing well below European and Asian students in their crucial last year at high school.[11] University of Auckland researchers advocate that more Māori and Pacific students need to stay at school longer in order to study appropriate Level 3 subjects (Madjar, McKinley,

---

[11] In 2009, 46% of the Pacific students and 53.4% of Māori in Year 13 gained NCEA Level 3, in comparison with 75.4% of European and 74.4% Asian.

Deynzer, & Van der Merwe, 2010). School leaders and teachers have a responsibility to ensure that academic pathways are made available for these students so that they can successfully transition to further degree level study.

Despite student academic results for schools and subjects being available in raw form on the NZQA website (NZQA, 2011c), there is little use made of these data by independent researchers. The main users appear to be the media. An analysis of the academic student results of 83 Auckland schools in a mainstream magazine in 2011 (Wilson & McGregor, 2011) produced tables that clearly show attendance at a lower decile school in Auckland is not necessarily a barrier to going on to university — although the writers acknowledged that there were some low-decile schools where students had a significantly lower chance of achieving that goal. There does appear to be whole-school effects in some schools where there is overall high performance across departments across three years but other schools exhibit inconsistency in results across departments and across years. Like secondary schools in London (Sammons et al., 1997), whole-school measures in New Zealand schools mask significant departmental differences. Since whole-school GPA does conceal department underachievement within a school it is important to calculate in judging school effectiveness, the percentage of Merit and Excellence results within a department. Phase 2 of the project revealed that NCEA results at Level 2 and 3 were also needed, because in some schools Level 1 academic performance is comparable to national norms but academic performance in departments drifts downwards at the higher levels of the qualification.

Phase 2 findings

Phase 2 of this project involved investigating the middle leadership practices at department level and analysing the student academic outcomes data for NCEA Level 1, 2 and 3 for 30 departments within 10 schools. Combining the results from these two investigations provided an opportunity to establish the extent of the relationship between the effectiveness of reported middle leadership and student academic achievement.

When an investigation of the level of effectiveness of middle leadership was conducted it was found there was considerable agreement between teachers and middle leaders in how they rated the effectiveness of the middle leadership practices in the department in which they all worked. This was true for each of the 30 departments, whether the ratings were low or high. There was a tendency for some schools to have reported overall effective (8D) or consistently poor (5A) middle leadership across every

department but there was also within-school variation. Higher decile schools appear to have teachers and leaders that report more effective middle leaders with some exceptions. One English department within a school (3A) with socially disadvantaged students scored higher in effective middle leadership practices than many departments with socially advantaged students. This was an encouraging finding and would support the literature that argues that teachers and middle leaders can be encouraged and supported to lead in any school context and become collaborators in the reform and improvement process (Harris & Muijs, 2005). The trend appeared to be that if the English department in a school was a poor performer in relation to middle leadership practices, the mathematics and science departments were also poor performers.

The results showed that the factor *Management of resources* is a highly contested area as there was considerable reported differentiation within schools and across subjects in terms of how resources were managed to best advantage student learning. The results of the analysis from the middle leadership questionnaire revealed a statistically significant difference for the subjects mathematics and science and English and science and this factor was also statistically different between all 10 schools. The items that sit within this factor such as prioritising the budget to meet the teaching goals and the equity of access to resources are well within the control of the middle leaders in a secondary school and yet there seems to be considerable variation between participants' responses within a school and across subjects throughout the sample, in their perception of how these resources are being managed. This factor had a positive relationship with student academic results at NCEA Level 2, so determining further evidence of the specific practices of effective middle leadership in this area would be worthwhile. Further, school decile makes little difference to the way teachers and middle leaders score this factor. It could reasonably be expected that teachers in lower decile schools would report less satisfaction with the allocation of resources in their department where there would be many needs and demands. The literature clearly points to the resource demands of subjects such as science and English being very different (Grossman et al., 2004) but in this project the responses within the same subject were also variable. The way resources are managed within a department has shown to have a significant relationship with student academic results at the senior level of secondary schooling so investigation into the specific middle leadership practices that positively impact on student learning, particularly in schools where resources are comparatively stretched, requires further investigation.

The department least likely to score the highest in terms of middle leadership practices within a school was mathematics. Just one school (8D) could claim that the

mathematics department was reported to be the most effectively led in the school. The respondents from mathematics departments reported that they experience less collegiality than science teachers. The maths respondents in this sample experienced a less collegial department environment, teamwork, fewer opportunities for leadership, or robust opportunities for professional discussion. As a subject group mathematics teachers and leaders gave lower scores on items related to a safe, supportive and well-organised working environment, compared to their science colleagues who responded. The reported lower scores for middle leadership attributes appear in most cases to *not* be impacting on the academic results for mathematics departments. There were six mathematics departments where middle leadership was reported by department members as being less effective than the average compared to the rest of the sample, but student academic outcomes for students in those departments were above the national norm at NCEA Level 2 and/or Level 3 in 2010.

The relationship between middle leadership and student academic outcomes

The third research question was focussed on analysing the relationship between student academic outcomes and middle leadership practices in order to establish which practices predicted student academic results. The scores for middle leadership across the sample did not correlate significantly with the student academic outcome data at NCEA Level 1 but at NCEA Level 2 and 3 there were statistically significant relationships. The results suggested a much stronger relationship between the factors *Goals and expectations*, *Management of resources* and *Positive learning environment for students and teachers* at the higher levels of the qualification. This is important because NCEA Level 3 is a critical level of schooling in New Zealand because students need to achieve a required number of credits from an approved list of subjects for University Entrance (UE).[12] Teachers require sound pedagogical content knowledge to deliver the curriculum requirements at this level and middle leaders in a school or department have a responsibility to ensure the course offerings are meeting the UE requirements so students have a chance at success. It is worth noting that generally 48% of school leavers in New Zealand achieve the University Entrance qualification and in the top decile schools that figure rises to approximately 80% (Wilson & McGregor, 2011). New Zealand educators, like those in many OECD countries aim to improve the proportion of students from challenging circumstances, gaining access to university in order to improve their life chances.

---

[12] University entrance requires the completion of a minimum of 42 credits at NCEA Level 3 or higher, including a minimum of 14 credits at Level 3 or higher in each of two subjects from the approved list of subjects, and a further 14 credits at Level 3 or higher from no more than two domains on the National Qualifications framework or other approved subjects. It also requires at least 14 credits at Level 1 or higher in mathematics and at least 8 credits at Level 2 or higher in English (Madjar et al., 2010).

When summary measures were used in this project to compare middle leadership scores and student academic outcomes at NCEA Level 2 and 3, the results were mixed. Forty percent of the departments in the sample had leadership scores that aligned to the student academic results, 30% had consistently positive relationships and 10% had consistently negative relationships. Nine (30%) departments scored positively in middle leadership practices but had negative outcomes in terms of meeting the national norm for student outcomes. Nine (30%) departments scored negatively in terms of middle leadership practices but positively in terms of student academic outcomes. The simultaneous regression provided more useful detail about the leadership factors that could be positively and negatively associated with student academic results. This is because it showed that school decile accounted for 62% of the variance in student academic outcomes at Level 2 and when school decile was controlled middle leadership could account for a further 22% of the student academic outcomes. Decile of school was a less important predictor for Level 3 NCEA student academic achievements (46%), but as with Level 2 when middle leadership practices were added the ability to improve student achievement rose to 62%. Therefore, a combination of school decile and middle leadership practices were strong overall predictors of student academic achievement at Level 2 and 3 NCEA. This is a critical finding of the study because it demonstrates that middle leaders and the teachers in their departments can make a difference to student academic outcomes despite the socioeconomic status or ethnicity of students. When decile is controlled for the departments with good student academic outcomes had higher middle leadership scores for the factors *Management of resources, Goals and expectations,* and *Positive learning environment for students and teachers.*

The factors related to *Collegiality* and *Focus on student academic results* were a negative predictor of student academic outcomes. This result is important as it supports the existing research that explains that collegial relationships are important only when they are improvement-focussed and pupils are making progress (Hargreaves, 1999). Department middle leaders need to ensure that collegial activity and effort is professionally focussed and goal centred, with a clear focus on teaching and learning so that students gain the best outcomes from the energy and effort of the teachers. This does not assume that departments that are collegial and focus on results will not be successful but that in order to be successful in terms of student academic outcomes, other leadership factors that are focussed on students learning needs require consistent attention and energy. The middle leaders in the departments with poorer academic results should focus their efforts on making a difference to the aspects that make the largest impact directly on students. Collegial relationships will always be positive in an

155

adult work environment but they need to be based on a shared professional ethic with the key focus on teaching and learning (Feist, 2008). These findings support the recommendations of Harris and Muijs (2005) who suggested a new form of professionalism as teachers are increasingly perceived of as leaders. They promote the concept of shifting from teaching at the centre to learning at the centre and from individualism to professional community, and from technical and managed work to inquiry and leadership (Harris & Muijs, 2005, p. 139). The results of the regression analysis in this research project revealed the specific middle leadership practices that impact both positively and negatively on student achievement in the senior academic years and the management and leadership practices that need to change and develop so that students gain greater advantages.

Contribution to knowledge

Calls continue to be made for research that clearly identifies the teacher and leadership practices that will make a difference to student academic outcomes. This project goes some way to identify that there is a significant link between middle leadership and student academic outcomes. This in turn, contributes to the schooling improvement research and develops further the findings of Sammons et al. (1997) that schools, even in disadvantaged areas, can impact on student academic outcomes. The results shed some light on the key practices at department level that are a positive or negative predictor of student academic outcomes. The significance of this study is supported by the validity of the data and tools used and the numbers of participants included in the study in Phases 2 and 3. In 2007, Bennet et al. published a review of the empirical research published in the English language that reported on the work of middle leaders in secondary schools, mostly in the United Kingdom. The research team reviewed 101 studies completed between 1988 and 2005. Although there have been some larger studies of this nature completed (Sammons et al., 1997), many of the studies included in this review consisted of small qualitative case studies. It appears that there was a strong research interest building in this field in the late 1990s but that within the last 5–10 years there have been no further significant studies that have built on the work of researchers such as Sammons et al. (1997), Harris (1999) and Hofman et al. (2001). This project produces further evidence that the impact of middle leadership is critical in the secondary school system and provides further ground to test and examine the practices that make the most difference to student outcomes at department level.

This study further develops methods and systems for analysing student academic results in a standards-based assessment system at senior secondary level and provides some procedures for categorising and analysing large sets of data. The Phase

One findings particularly contribute to the New Zealand evidence on within-school variation, including the evidence provided by the New Zealand results for 15-year-olds participating in PISA studies (OECD, 2010a, 2010b). By using the department results for English, mathematics and science across three years, a more detailed understanding of the extent of the variability within and across identified schools in urban Auckland can be examined. This research adds to the national picture provided by PISA by verifying and contextualising some of the results. PISA provides a national picture of how student ethnicity and socioeconomic status are linked to student academic outcomes. This study provided the results for individual departments and schools where these 15-year-olds are actually learning and revealed information about the middle leadership practices that were impacting on their academic outcomes.

The middle leadership questionnaire developed for this project has been shown to have a high level of statistical reliability. It has been developed from key theories and is not New Zealand-context specific so could be replicated and used for research and professional learning and development purposes in many OECD countries. It could and should be adapted based on these results, with some of the collegiality items being removed and focussed on problem-solving and developing professional learning communities (Feist, 2008; Harris & Muijs, 2005) as opposed to teams.

Limitations of the study

The most striking limitation of this work is its reliance on quantitative data. Although there is strength in its objectivity, the design methodology has not allowed for further in-depth investigation at department level to observe, test and probe some of the findings that have been reported at both school and department level.

A second limitation is that the data that have formed the basis for this study is that which is publicly available through the NZQA website. These data do not reveal the numbers of students who were absent from class, left school before NCEA exams, or were excluded. If that school-based data were available, the way that department success was judged may be different. By using data that are publicly available, this researcher was unable to investigate which 15-year-olds who were enrolled at a particular school were not represented in the data.

A third limitation is that the sample of schools that participated in Phase 2 were volunteers and it was the principal of the school who made the decision about whether the English, mathematics and science middle leaders and teachers would be invited to participate. Middle leaders and teachers had a choice to attend the meeting to hear

about the project and complete the questionnaire and this had some impact on the response rates (75%). Although the sample can be used for exploratory purposes, it cannot be generalised to any population of schools. The schools and departments that volunteered to be involved in the research for Phase 2 are not a true sample of the original Phase 1 schools. There are no decile 10, 4, 2 or 1 schools in the Phase 2 sample. If a greater range of schools had participated in Phase 2 the general findings regarding middle leadership would have been applicable to New Zealand high schools from all deciles. The departments in this study were volunteers and staff were informed at the time of the administration of the questionnaire that results would be matched with student academic outcomes. Therefore, a judgement of their effectiveness in terms of student academic outcomes was not made prior to their selection. In-depth case study work would be best done in specific contexts where department leadership had been identified as effective or not effective, therefore using a purposive sampling technique. This would test the common features of effective departments and identify the multiplicity of issues and widespread inconsistency often apparent in the ineffective schools and departments (Sammons et al., 1997). The weaknesses and limitations identified in this section could be overcome with a follow-up case study approach where data could be verified and observations of middle leadership practices could be carried out, particularly in departments that had been identified as either very effective or ineffective.

Implications for further research

The findings from the two phases of the study provide considerable information that warrants further investigation. Phase 1 of this study is an analysis of the effective and not so effective departments in terms of student academic outcomes for 41 urban schools, while Phase 2 provides information about the variable middle leadership capability in 10 of those schools. There are three key areas that require further research as a result of the findings. The first is an inquiry into the negative relationship between student academic outcomes and the factor *Collegial Working Environment*. The second, key area is a deeper scrutiny of the overall positive relationship between academic outcomes and leaders effectively managing resources. The third, is an examination of the differing leadership practices in the curriculum areas, with the positive influence of subject English.

The middle leadership practices that were grouped into the 'collegiality' factor in this study were negatively related to student academic outcomes at NCEA Level 2 and 3. There is contradictory evidence in the literature regarding the importance of collegial relationships in subject departments (Bird & Little, 1986; Hargreaves, 1999) and the

way that collegiality is defined in the literature differs. In studies where collegiality has been defined as a strong professional learning community approach, the outcomes for students are positive, whereas when the collegial practices are focused mostly on the positive relationships between the adults there are mixed results in terms of outcomes for students. Philips (1997) found that in schools where teachers are more collaborative on a social level and not as concerned with academic learning, mathematics scores were lower. Research supports the notion that effective professional learning communities or in this context 'the subject department' are successful when they are deeply rooted in the academic and social learning goals of the school. These communities are not merely focussed on creating pleasant work environments (Goldring, Porter, Murphy, Elliott, & Cravens, 2007) but are professional communities where concerns about effective teaching and learning are a persistent focus. In this current study the items within the collegiality factor generally assessed whether the adults in the department had good social relationships and operated as a team. This might explain the negative relationship to student academic outcomes.

In contrast with the collegiality factor, the factor *Positive Learning Environment for Students and Teachers* included items that focused on classroom practice and quality teaching which align more strongly with aspects of collegial professionalism described in the literature (Hargreaves, 1999). A clear definition of the types of middle leadership practices that are considered 'collegial' needs to be established, combining and eliminating some of the items from the factors *Collegial Working Environment* and *Positive Learning Environment for Students and Teachers* would be an important first step in developing research tools for a further study. Establishing a clear definition and ensuring the items describe middle leadership practices considered to be effective would be important in any further study. Trialling these items in school departments and revising them accordingly to ensure that the practices described are those that link positively to student outcomes would also be critical. Further investigation that involved observing those practices that appear to strengthen the teaching and learning in a department would assist middle leaders to identify and prioritise the work required to lead and manage their departments effectively. A study of this type would provide insights about a range of teachers and leaders in a school that describe the types of middle leadership attitudes and practices that motivate others, are conducive to a highly professional environment and efficiently mobilise the staff in the department (Dinham & Rowe, 2007). Observing how the 'collegial' practices of an effective professional community are deployed in different curriculum areas is worth investigating in light of the subject differences revealed in this current study These practices could be

observed, exemplified and tested in department focus groups, through observations and video/audio evidence.

The factor that had the greatest degree of positive impact on student academic outcomes for all subjects was the *Management of Resources.* There were just three questionnaire items included in this factor and they assessed whether staff had equitable access to department resources and how those resources were allocated to building a culture focused on student learning. Teachers and middle leaders were also asked about the extent to which the budget decisions aligned with priority teaching goals. There is evidence to suggest that middle leaders in schools with exceptional student outcomes are prepared to invest money and time in empowering others (Dinham & Rowe, 2007). An investigation into the ways effective leaders carry out this responsibility at department level and in subject specific contexts would add further knowledge about exactly how the effective deployment of resources is done at department level when there are multiple demands and expectations from teachers and students. Middle leaders in secondary schools often have direct responsibility for advocating and deploying department resources. They also experience considerable time pressures, so investigating the way effective middle leaders use their time to prioritise and manage resources to best support student learning and focus on priority teaching goals, would be helpful information that could enhance performance across the sector.

Further research is required to enquire into how middle leadership is practiced in different 'subject' departments. A case study approach to investigate and acknowledge the curriculum and pedagogical contexts for English, mathematics and science middle leaders and teachers who work within varying resource demands and expectations, would enhance understanding and knowledge of subject requirements and practices. A comparison of the similarities and differences between departments would provide an understanding of the extent to which subjects or curriculum expertise dictates leadership attributes and the leadership requirements, particularly around the management of resources.

The current project suggests that when the leadership of the English department in a school is strong there is a positive relationship to the mathematics and science department leadership practices and student academic results. Researchers need to investigate, understand and test the power of effective middle leadership in subject English to raise student achievement across the whole school. The reasons for this positive relationship could be that effective leader(s) of a large English department in a school influence the culture of leadership practice across other departments and

therefore contribute to whole-school effects. In addition, if English departments are well led and students are being effectively taught it is conceivable that student's skills and knowledge in reading and writing improves and they are able to better meet the literacy demands of other subjects and their overall academic outcomes are improved accordingly. Understanding how this positive influence might occur would require investigation into the leadership attributes, relationships and networking abilities of effective and less effective middle leaders of English.

Implications for policy and professional learning and leadership development

In this study, all the participants were secondary school teachers who were interested in knowing and understanding more about effective middle leadership practices. As the researcher administered the questionnaire in person, many participants commented on the questionnaire and indicated their desire to have a copy for their own or departmental reflection. Middle leaders were very interested in the results of the study and keen to compare their department performance with those in other similar schools. The importance of middle leaders getting direct feedback from the teachers they lead on criteria they know to be effective cannot be underestimated because this provides them with clear guidance on goals for improvement. Middle leaders are also interested in comparing their department performance to other subjects and other schools. The questionnaire can be best used as a tool to assist the development of departments as learning communities and to monitor their improvement.

When principals were invited to participate they were offered a staff professional development session that summarised the findings from the research. This will be initiated with the 10 schools in Phase 2 in 2012. The general findings will be of interest to many secondary school middle leaders as they provide a good indication of which practices are likely to be more effective for student outcomes and how collegiality might need to be conceptualised if it is to have any impact on improved professional practice.

Professional learning and development needs to focus on middle leaders as having a key role in promoting and managing all curriculum and assessment development and change in the secondary education system. In order to embed these changes in teaching and assessment practices and build sustainable shifts in schools, it will be essential to engage middle leaders in the specific leadership practices that enhance their capability and capacity as curriculum and pedagogical leaders. School leaders need to be supported to understand that enhanced middle leadership capability will contribute to school-wide coherent curriculum review and effective pedagogical

practice. This research also suggests that school leaders should require departments to function as effective professional learning communities focussed on the effectiveness of teaching and learning, collaborative problem-solving and reflective practice.

Professional standards and accountability for middle leaders

The Ministry of Education's Statement of Intent (2009a, p. 11) states as a priority: "*Every young person has the skills and qualifications to contribute to their and New Zealand's future."* The Government has set a goal of all young people achieving NCEA Level 2 or an equivalent qualification, as students' educational success is critical to New Zealand's economic and social success. The Ministry has also made a commitment to implement changes and improvements to NCEA. There are approximately 18,000 teachers in secondary schools in New Zealand, so this project shows that an investment by government in supporting the increased middle leadership practices of the department leaders has a statistically significant chance of improving the student academic outcomes at Level 2 and 3 NCEA. Targeted resourcing that encourages accountability and enhanced professional practice of middle leaders is an investment that will improve outcomes for students.

A solid evidence based framework for middle leader professional standards and staff development needs to be established and developed. Middle leadership practices that are strongly centred on expectations of consistently high quality teaching in the department and on giving priority to their own and their staff's professional development needs are critical for student academic success. Middle leaders who recruit high quality teachers and then encourage them to implement and share new ideas have also been shown in this project to be important. The findings suggest that in departments where there is an expectation for middle leaders and teachers to be highly accountable and to be seeking to improve their professional knowledge, the impact for students is consistently positive. In schools where middle leaders are prepared to set high standards and ensure all staff are accountable they are likely to reap the rewards in terms of student outcomes.

There is a strong argument for strengthening professional accountability using rigorous performance standards across the education system. The factor *Goals and Expectations* in this project had a positive relationship with student academic outcomes. Middle leaders who ensured that department goals and plans related to those of the school and were measurable and observable suggests that this type of accountability within schools and departments is important. Accountability stems from performance mechanisms from both within the school and externally across the whole system

(Adams & Kirst, 1999). Recent research in America has reported a positive relationship between the strength of a states' accountability system and student achievement (Carnoy & Loeb, 2002; Hanushek & Raymond, 2004). Internal accountability systems and local expectations are important for improving the capability of teachers and leaders (Goldring et al., 2007). Evidence suggests that schools that have rigorous academic standards, high-quality instruction and a focus on academic success are more likely to meet the equity challenge (Goldring et al., 2007). School systems where goals and expectations are set need to develop within them a mechanism for appropriately assessing and developing leaders that is driven by central government. Adequate training and support at the regional and school level is required well before principalship and in the secondary school context, when a teacher is emerging as a candidate for middle leadership.

School leaders must understand the requirements of the external accountability system in their region or jurisdiction and ensure that they integrate those requirements with their own school goals, appraisal systems and provision of professional learning opportunities. This can be achieved by sharing information and holding their staff accountable for implementing strategies and setting achievement goals and targets for themselves as professionals and also as a department (Goldring et al., 2007). It is critical for all school leaders to play an integral role in focusing their staff and students on professional performance standards that are likely to contribute to student success. These criteria need to be referenced in department meetings, performance reviews, classroom observations, and in discussion of curriculum and instructional strategies (Goldring et al., 2007).

In New Zealand schools middle leaders can use the Registered Teacher Criteria (New Zealand Teachers Council, 2012) to build their knowledge and confidence in appraising teachers. They have the responsibility to support the teachers in their departments to work toward and maintain full registration. Middle leaders are embedded in the day to day running of the department and therefore have the ability to strengthen the culture of self-responsibility, accountability and improvement within the departments that they lead. They can ensure that their department has strong evaluative processes that underpin sound appraisal and that the relationship between appraisal and professional learning and development for all teachers is clear and direct. This study indicates that the ability to engage in a range of appraisal conversations including those needed to address any gap between a teacher's current practice and agreed elements of practice to enhance student outcomes is critical for department and whole-school effectiveness.

Middle leaders who are currently performing their role or aspiring to step up to leadership require support and development to understand the middle leadership practices on which they should focus for improved student academic outcomes. The current teacher registration standards could be extended to provide accountabilities for middle leaders so that effective practice is clearly mandated. Performance appraisal linked to professional standards for middle leaders could be developed so that there are clear guidelines for middle leadership development that include a 'leading learning' culture within a department. School leaders need to provide time to ensure middle leaders are adequately trained for their role so that they are focussed on the practices that will enhance student learning. Providing resources so that their own professional learning and development needs are met as well as those of their staff is also critical. The resourcing of the development of professional standards for middle leaders that are linked to the current professional standards for teachers is a strong recommendation from this project.

Conclusion

The complex results of this study are a further reminder that categorisation or labelling schools does not describe the differential effectiveness that occurs within the department and classrooms and across years. Sammons et al. (1997) stated that attempts to describe schools as *good* or *bad* were genuinely unhelpful. They argued that effectiveness should only be judged on the basis of three years of performance. Yet parents and communities require this information and given that their taxes are paying for state schools, it could be argued that they are entitled to understand how schools perform at department or subject level and in relation to each other. The ongoing dilemma and political debates on the value of league tables balanced with the desire of communities to understand how effective schools are in educating young people in a world that is increasingly diverse and competitive, continues to be of great interest to educators and the general public. Sammons et al. state that the focus on the individual school is also relevant because while students and parents may have had some choice in school, they will seldom have been in a position to choose a particular set of teachers. They assert that raw league tables cannot tell parents and prospective students how a school performs, nor whether a school is differentially effective for some groups of students such as boys or particular ethnicities. The research indicates that judgments about schools are fairly complex (Sammons, Thomas, & Mortimore, 1996) and steps need to be provided to produce community-friendly and independent information about schools (Wilson & McGregor, 2011). Supporting communities to engage with the teachers and leaders at their local school by taking an interest in the

teaching and learning of their teenagers will help students to achieve to their full potential. Providing information about relative school performance in a range of forms and languages that are accessible will empower parents and support them to engage with their local school. Increased cooperation and meaningful relationships between parents and secondary schools and departments have been shown to advantage both parties when there is an integration of home and school in supporting students' learning (Busher, Harris, & Wise, 2000).

This study showed how effective middle leaders make a contribution to department performance in the senior secondary school. Subject leaders as opposed to school leaders are uniquely positioned as they have the pedagogical content knowledge required to impact positively on the quality of teaching and learning in a department, particularly at the senior level. Previous studies have shown that middle leaders also need to set clear goals and expectations, manage resources effectively and develop a culture of professional learning and inquiry that impact positively on student academic outcomes (Busher et al., 2000; Dinham, 2007; Sammons et al., 1997). Interventions in teaching and learning are essential, but on their own will not result in sustainable levels of improvement unless the critical school leaders are committed to embedding change (Busher et al., 2000). As the increased workload of middle leaders becomes a barrier for improvement, the focus needs to be on developing understanding of the leadership behaviours and practices that put students in the centre (Ministry of Education New Zealand, 2007, p. 8) so that their aspirations and abilities, no matter what their socioeconomic status, is always the teaching priority. Government officials need to focus on strengthening the accountability mechanisms for middle leadership performance both centrally and within schools. The development and support of middle leaders must be encouraged and resourced at all levels of the education system, because the rewards in terms of student academic achievement in the senior secondary school have been shown in this project to impact positively regardless of the decile of the school in which students are studying.

Existing New Zealand research indicates that if we are to increase the proportion of minority and socially disadvantaged students gaining UE, students and their teachers need to receive clear messages about how best to prepare their students for the demands of tertiary study (Earle, 2007). Students need to receive quality teaching so that they can achieve the credits necessary alongside clear and detailed guidance to ensure that the school courses they take are appropriate for their long-term goals. This guidance needs to occur in the early high school years (Madjar et al., 2010). Department middle leaders have a key responsibility in this area as this study shows

that effectively employed leadership practices will contribute to students' chances of academic success in a very direct way.

This conclusion takes us back to the introduction of this thesis, the moral imperative for the project. The research imperative to inform ongoing school improvement to support educators to understand the evidence of what makes a difference and enhance outcomes for students, continues to take a high priority at both government and school level. The recent Briefing to the Incoming Minister (Ministry of Education New Zealand, 2011, p. 3) in New Zealand states:

> There is evidence of gains in our education system over the last ten years particularly in achievement and participation. Those improvements have done little to address the fundamental problem that the system works well for some learners and poorly for others.

Best practice needs to become common practice and we need to move away from the idea that the problems lie in a fairly small number of schools in poor areas, because they do not. Strengthening middle leadership and the teaching and learning at department level will help more young people achieve the academic outcomes they will need to enable them to lead successful and fulfilling lives.

Appendices

Appendix A: Middle leader questionnaire

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

Middle Leadership in Secondary Schools Survey

This survey asks about leadership in the context of your department.
Please answer the questions considering the leadership of all the people who hold formal leadership positions in your department such as the HOF, HOD, assistant HOD, and curriculum/year level leaders.

1. In which department do you currently have a leadership role? (*circle one*)

English   Maths   Science

2. What is your current position title? (*circle one*)

HOD   HOF   Assistant HOD   Other (*please state_____*)

Read the 34 statements on the following pages about leadership practice and use the rating scale below to indicate the one response that comes <u>closest</u> to describing your opinion for each item. **Indicate how strongly you agree with these statements about the leadership practices in your department. Note: the leadership practices relate to your own leadership practices and those of other leaders in your department.**

➢ Strongly Disagree
➢ Mostly Disagree
➢ Slightly Disagree
➢ Slightly Agree
➢ Moderately Agree
➢ Mostly Agree
➢ Strongly Agree

1

	The leaders of this department ensure that:	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Moderately Agree	Mostly Agree	Strongly Agree
Goals	1. School goals are translated into clear department goals.	☐	☐	☐	☐	☐	☐	☐
	2. Teachers have performance appraisal goals and plans that relate to department and school goals.	☐	☐	☐	☐	☐	☐	☐
	3. Teachers understand how their classroom work contributes to the department and school goals.	☐	☐	☐	☐	☐	☐	☐
	4. The department has clear and measurable academic goals.	☐	☐	☐	☐	☐	☐	☐
Managing resources	5. The department budget is aligned to priority teaching goals.	☐	☐	☐	☐	☐	☐	☐
	6. Staff have equitable access to department resources.	☐	☐	☐	☐	☐	☐	☐
	7. Resources are allocated to building a culture focused on student learning.	☐	☐	☐	☐	☐	☐	☐
	8. Meeting time is allocated to support the achievement of priority teaching and learning goals.	☐	☐	☐	☐	☐	☐	☐
Ensuring an orderly and supportive environment	9. The department provides a safe, supportive, and well organised environment for teaching and learning.	☐	☐	☐	☐	☐	☐	☐
	10. Staff handle student discipline problems with fairness and equity.	☐	☐	☐	☐	☐	☐	☐
	11. Problems that staff raise about the work environment are addressed in a timely and effective way.	☐	☐	☐	☐	☐	☐	☐
	12. Conflict within the department is quickly identified and resolved.	☐	☐	☐	☐	☐	☐	☐
	13. Students experience the classroom environment as supportive of their learning.	☐	☐	☐	☐	☐	☐	☐

2

169

	The leaders of this department ensure that:	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Moderately Agree	Mostly Agree	Strongly Agree
Effective teaching and student outcomes	14. There is an expectation of consistent high quality teaching.	☐	☐	☐	☐	☐	☐	☐
	15. They are involved in the recruitment of high quality staff for the department.	☐	☐	☐	☐	☐	☐	☐
	16. They take responsibility for the learning of ALL students in this department.	☐	☐	☐	☐	☐	☐	☐
	17. Students who are struggling to achieve are taught by the most effective teachers.	☐	☐	☐	☐	☐	☐	☐
	18. They make time to observe department teachers at work in their classrooms.	☐	☐	☐	☐	☐	☐	☐
	19. All students have the opportunity to attempt NCEA standards that will provide access to university.	☐	☐	☐	☐	☐	☐	☐
Assessment and data	20. Assessment data are used to plan teaching.	☐	☐	☐	☐	☐	☐	☐
	21. Teachers set academic goals that are based on analysis of their students' achievement data.	☐	☐	☐	☐	☐	☐	☐
	22. They lead discussion about students' results with staff.	☐	☐	☐	☐	☐	☐	☐
	23. Record keeping processes are efficient and effective	☐	☐	☐	☐	☐	☐	☐
	24. Student's receive high quality, timely feedback that helps them to improve their work	☐	☐	☐	☐	☐	☐	☐
Culture of reflection and inquiry	25. Teachers are provided with effective and timely advice about any concerns they have about their teaching	☐	☐	☐	☐	☐	☐	☐
	26. They take care of their own professional development needs and share their new learning with colleagues.	☐	☐	☐	☐	☐	☐	☐
	27. There are opportunities for teachers to give feedback to departmental leaders and vice versa.	☐	☐	☐	☐	☐	☐	☐
	28. Professional development ideas are implemented and used in the classroom.	☐	☐	☐	☐	☐	☐	☐

3

	The leaders of this department ensure that:	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Moderately Agree	Mostly Agree	Strongly Agree
	29. Student feedback is regularly collected, analysed and acted upon.	☐	☐	☐	☐	☐	☐	☐
Team approach	30. The staff work together as a team and support each other.	☐	☐	☐	☐	☐	☐	☐
	31. Leadership roles are shared among the department's teaching staff according to individual strengths and interests.	☐	☐	☐	☐	☐	☐	☐
	32. The department is a collegial place to work where staff support each other.	☐	☐	☐	☐	☐	☐	☐
	33. There are robust discussions about successes and challenges.	☐	☐	☐	☐	☐	☐	☐
	34. Staff talents and achievements are recognised and celebrated.	☐	☐	☐	☐	☐	☐	☐

4

171

	The leaders of this department ensure that:	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Moderately Agree	Mostly Agree	Strongly Agree
	29. Student feedback is regularly collected, analysed and acted upon.	☐	☐	☐	☐	☐	☐	☐
Team approach	30. The staff work together as a team and support each other.	☐	☐	☐	☐	☐	☐	☐
	31. Leadership roles are shared among the department's teaching staff according to individual strengths and interests.	☐	☐	☐	☐	☐	☐	☐
	32. The department is a collegial place to work where staff support each other.	☐	☐	☐	☐	☐	☐	☐
	33. There are robust discussions about successes and challenges.	☐	☐	☐	☐	☐	☐	☐
	34. Staff talents and achievements are recognised and celebrated.	☐	☐	☐	☐	☐	☐	☐

4

172

PLEASE COMPLETE THE FOLLOWING DEMOGRAPHIC QUESTIONS

1. Gender (please tick): Male ☐   Female ☐

2. **Which ethnic group do you belong to**? *(please tick)*

New Zealand European	☐
Maori	☐
Samoan	☐
Cook Island Maori	☐
Tongan	☐
Niuean	☐
Chinese	☐
Indian	☐

3. **Leadership Experience in a Secondary School**
Please indicate your leadership experience in any secondary school in any country

Tick	Leadership Role	Total years	Experience In NZ	Experience in another country (please state)
☐	Assistant HOD			
☐	HOD			
☐	Subject Head (*specify e.g., Chemistry*)			
☐	Head of Faculty			
☐	Curriculum Leader			
☐	Year/Form Dean			
☐	Specialist Classroom Teacher			
☐	Senior Management Team (i.e., AP, DP, SM)			
☐	Other (*specify*)			

4. **How long have you been in your CURRENT leadership role in THIS school?**
Please circle one of the following:

Less than 3 years   3-10 years   11-20 years   More than 20 years

5. **Qualifications**
What qualifications do you hold? For each qualification you hold, indicate the level, year gained, and any area of speciality.

Tick	Qualification	Year gained	Area/Specialty (*e.g. Chemistry, MEd Admin, etc*)
☐	Undergraduate Diploma/Certificate		
☐	Bachelor Degree		
☐	Post graduate Diploma/Certificate		
☐	Masters		
☐	Doctorate		

Thank you for your time. Your help is greatly appreciated.

Please return survey into the box anonymously

Appendix B: Teacher questionnaire

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
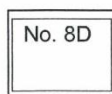
Teachers Perceptions of Department Leadership

This survey asks about leadership in the context of your department.
Please answer the questions considering the leadership of all the people who hold formal leadership positions in your department such as the HOF, HOD, assistant HOD and curriculum/year level leaders.

1. In which department are you currently mostly teaching? (*choose one*)

English   Maths   Science

Read the 34 statements on the following pages about leadership practice and use the rating scale below to indicate the one response that comes <u>closest</u> to describing your opinion for each item.

Indicate how strongly you agree with these statements about the leadership practices you experience in your department.

  ➤ Strongly Disagree
  ➤ Mostly Disagree
  ➤ Slightly Disagree
  ➤ Slightly Agree
  ➤ Moderately Agree
  ➤ Mostly Agree
  ➤ Strongly Agree

1

	The leaders of this department ensure that:	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Moderately Agree	Mostly Agree	Strongly Agree
Effective teaching and student outcomes	14. There is an expectation of consistent high quality teaching.	☐	☐	☐	☐	☐	☐	☐
	15. They are involved in the recruitment of high quality staff for the department.	☐	☐	☐	☐	☐	☐	☐
	16. They take responsibility for the learning of ALL students in this department.	☐	☐	☐	☐	☐	☐	☐
	17. Students who are struggling to achieve are taught by the most effective teachers.	☐	☐	☐	☐	☐	☐	☐
	18. They make time to observe department teachers at work in their classrooms.	☐	☐	☐	☐	☐	☐	☐
	19. All students have the opportunity to attempt NCEA standards that will provide access to university.	☐	☐	☐	☐	☐	☐	☐
Assessment and data	20. Assessment data are used to plan teaching.	☐	☐	☐	☐	☐	☐	☐
	21. Teachers set academic goals that are based on analysis of their students' achievement data.	☐	☐	☐	☐	☐	☐	☐
	22. They lead discussion about students' results with staff.	☐	☐	☐	☐	☐	☐	☐
	23. Record keeping processes are efficient and effective	☐	☐	☐	☐	☐	☐	☐
	24. Student's receive high quality, timely feedback that helps them to improve their work	☐	☐	☐	☐	☐	☐	☐
Culture of reflection and inquiry	25. Teachers are provided with effective and timely advice about any concerns they have about their teaching	☐	☐	☐	☐	☐	☐	☐
	26. They take care of their own professional development needs and share their new learning with colleagues.	☐	☐	☐	☐	☐	☐	☐
	27. There are opportunities for teachers to give feedback to departmental leaders and vice versa.	☐	☐	☐	☐	☐	☐	☐
	28. Professional development ideas are implemented and used in the classroom.	☐	☐	☐	☐	☐	☐	☐

3

	The leaders of this department ensure that:	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Moderately Agree	Mostly Agree	Strongly Agree
	29. Student feedback is regularly collected, analysed and acted upon.	☐	☐	☐	☐	☐	☐	☐
Team approach	30. The staff work together as a team and support each other.	☐	☐	☐	☐	☐	☐	☐
	31. Leadership roles are shared among the department's teaching staff according to individual strengths and interests.	☐	☐	☐	☐	☐	☐	☐
	32. The department is a collegial place to work where staff support each other.	☐	☐	☐	☐	☐	☐	☐
	33. There are robust discussions about successes and challenges.	☐	☐	☐	☐	☐	☐	☐
	34. Staff talents and achievements are recognized and celebrated.	☐	☐	☐	☐	☐	☐	☐

4

PLEASE COMPLETE THE FOLLOWING DEMOGRAPHIC QUESTIONS

1. Teaching and Leadership Experience in a Secondary School
Please indicate your teaching and leadership experience in any secondary school in any country.

Tick	Leadership role	Years experience	Country
□	Teaching		
□	Assistant HOD		
□	HOD		
□	Subject Head (*specify e.g., Chemistry*)		
□	Head of Faculty		
□	Curriculum Leader		
□	Year/Form Dean		
□	Senior Management Team (i.e., AP, DP, SM)		
□	Other (*specify*)		

2. How long have you been teaching in this department in THIS school?
Please circle one of the following: < 3 years   3-10 years   11-20 years   > 20 years

3. Qualifications
What qualifications do you hold? For each qualification you hold, indicate the level, year gained, and any area of speciality.

Tick	Qualification	Year Gained	Area/Speciality (*e.g. Chemistry, MEd Admin, etc*)
□	Undergraduate Diploma/Certificate		
□	Bachelor Degree		
□	Post graduate Diploma/Certificate		
□	Masters		
□	Doctorate		

4. Have you had previous leadership experience in a secondary school?
Please circle: YES/NO
If you answered YES, please provide details:_____

Thank you for your time. Your help is greatly appreciated.

Please return survey into the box anonymously

5

Appendix C: Ethics material including Participant Information Sheets and Consent Form for schools

School of Teaching Learning and Development
Te Kura Whakawanake Akoranga

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tamaki Makaurau
INCORPORATING THE AUCKLAND COLLEGE OF EDUCATION

Epsom Campus
Gate 3, 74 Epsom Avenue
Auckland, New Zealand
Telephone 64 9 623 8899
Facsimile 64 9 623 8898

The University of Auckland
Private Bag   92601
Symonds St, Auckland

Participant Information Sheet
School Principal
EdD Research Project: Middle Leadership in Secondary Schools
Researcher: Camilla Highfield

Researcher Introduction:
My name is Camilla Highfield, and I am a doctoral student at the University of Auckland. I am approaching you, to request your permission to ask leaders and teachers in your English, maths and science departments to participate in a research project.
This research will examine the relationship between secondary schools' academic outcomes and middle leadership practices. The study is being supervised by Professor Viviane Robinson and Dr Christine Rubie-Davies, from the University of Auckland.

Project Description:
My academic interest has developed out of my work involving identifying good leadership practices at department level in secondary schools. The aim of this research is to ascertain what effective leadership practices are carried out by department leaders in English, maths and science departments in secondary schools, and how they relate to student outcomes. The ultimate aim is to develop tools and strategies to help all middle leaders in secondary schools so they know what leadership practices are effective and why.

Project Procedures:
I am requesting each middle leader and teachers that report to them in maths, English and science departments to complete a leadership practices questionnaire. I anticipate this taking no more than 30 minutes in a staff meeting. Heads of Department will only be contacted if you have given permission for your staff to become involved in the project. The field work for this study is being carried out between August and December 2010.

If you are willing for staff in your school to participate all you would need to do is notify me and I will liaise directly with the HOD's of the English, maths and science departments. I will then set up an appropriate time at either a department or staff meeting when convenient to explain and hand out the

anonymous questionnaires. No personal information such as names needs to be recorded on the questionnaire. The school and department is identified by a code on the questionnaire and the middle leaders and teachers are required to complete similar questions but from their own perspective and deposit them into a box. Teachers and middle leaders will be asked to complete the questionnaire in private and return it (completed or left blank) to a box that I will bring along on the day of my visit and collect 24 hours later. This ensures that no individual completing the anonymous questionnaire can be identified.

Data storage/retention/destruction and future use
The questionnaires filled in by teachers and middle leaders in your school will be kept in secure storage for six years at the University of Auckland and will inform the writing of my thesis and other academic papers. After six years the questionnaires will be shredded and any electronic files of the data will be wiped from the password-protected computer used to store them.

Right to withdraw from participation
All teachers and middle leaders in the department will be asked to put a questionnaire into the box within 24 hours of the visit. The questionnaire could be completed or left blank if staff do not want to participate. Participants have the right to withdraw from participation at any time up until submitting the anonymous questionnaire into the box. Once the questionnaire is submitted anonymously it is not possible to withdraw that participant's data from the research as individual participants cannot be identified. Staff may decide not to participate. Should any staff mention their involvement in this project, I seek your assurance that participation or non-participation in this research project will not affect the staff members' relationship with the school in any way.

Anonymity and Confidentiality
Individual anonymity can be guaranteed in this project. Your school and department will have a code only known to me as the researcher. Individuals will not be identified at any stage of the project. There are sufficient schools involved in this project that in any written publications about the work the schools and departments will not be able to be identified.

Advantages for your school in being involved in this project
The results from the whole study across all schools could be shared with you initially but could also be presented in a staff meeting for professional development purposes after July 2011.

Contact Details
If you have any questions or want any more information at this stage, My contact details are:
Camilla Highfield
Faculty of Education
The University of Auckland
T: 09 623 8929 ext. 48929
M: 0276 882 881
F: 09 623 8952
E: c.highfield@auckland.ac.nz

My Supervisors Details are:

Professor Viviane Robinson
The University of Auckland
Faculty of Education
Phone: 09 3737599 ext. 87372
Email vmj.robinson@auckland.ac.nz

Dr Christine Rubie-Davies
The University of Auckland
Faculty of Education
Ph 096238899 ext 8274
Email:c.rubie@auckland.ac.nz

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human participants Ethics Committee, The University of Auckland, office of the Vice- Chancellor, Private Bag 92019, Auckland 1142. Ph 093737599 ext 83711

Thank you for giving this matter your consideration. I will phone you to discuss this matter and make an appointment to meet with you and discuss the details if you would find that helpful.

Kind regards

Camilla Highfield

THE UNIVERSITY OF AUCKLAND
FACULTY OF EDUCATION
Te Kura Akoranga o Tamaki Makaurau
INCORPORATING THE AUCKLAND COLLEGE OF EDUCATION

Epsom Campus
Gate 3, 74 Epsom Avenue
Auckland, New Zealand
Telephone 64 9 623 8899
Facsimile 64 9 623 8898

The University of Auckland
Private Bag   92601
Symonds St, Auckland

CONSENT FORM
School Principal

EdD Research Project: Middle Leadership in Secondary Schools
Researcher: Camilla Highfield

I have read the Participant Information Sheet, have understood the nature of the research and why I have been asked to give permission for some of my staff to participate in this research, should they choose. I have had the opportunity to ask questions and have them answered to my satisfaction.

I agree for HOD'S in English, maths and science and their staff in my school to be approached by the researcher to be asked to take part in this research.

I understand that participation will take up to 30 minutes of their meeting time.

I give my assurance that participation or non-participation by any of my staff will not affect their relationship with the school in any way.

I understand that staff are free to withdraw from participation at any stage, prior to the questionnaires being placed in the 'anonymous' box after the questionnaires are administered.

I understand that no one will have access to the individual data and only the researcher will have access to the school and department data.

I understand that I can request a professional development session with my staff that details the findings from the whole project once it is completed, from July 2011.

I understand that data will be kept for 6 years, after which they will be destroyed.

Name_____

Signature_____   **Date**_____

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANT ETHICS COMMITTEE ON18th
AUGUST, 2010 FOR 3 YEARS REFERENCE NUMBER 2010/303

Middle Leadership in Secondary Schools   4
Participant Information Sheet - Principal

Appendix D: Reliability indices Factor analysis

Five factors and items from questionnaire related to effective departmental leadership in secondary schools

Items	Factor Loadings					Communality
	1	2	3	4	5	
Factor 1: Collegial Working Environment α = .95						
32 The department is a collegial place to work where staff support each other.	**.91**	.35	–	.43	.45	.85
30 The staff work together as a team and support each other.	**.88**	.40	.31	.43	.45	.79
33 There are robust discussions about successes and challenges.	**.86**	.49	–	.43	.45	.76
12 Conflict within the department is quickly identified and resolved.	**.81**	.30	.37	.37	.47	.68
31 Leadership roles are shared among the department's teaching staff according to individual strengths and interests.	**.81**	.49	.34	.46	.46	.68
34 Staff talents and achievements are recognised and celebrated.	**.77**	.48	–	.43	.44	.61
9 The department provides a safe, supportive, and well organised environment for teaching and learning.	**.79**	.37	.47	.46	.54	.67
11 Problems that staff raise about the work environment are addressed in a timely and effective way.	**.73**	.47	.44	.49	.58	.64
25 Teachers are provided with effective and timely advice about any concerns they have about their teaching.	**.72**	.49	–	.52	.60	.60
16 They take responsibility for the learning of ALL students in this department.	**.67**	.53	–	.48	.58	.57
Factor 2:Focus on Student Academic Results α = .87						
21 Teachers set academic goals that are based on analysis of their students' achievement data.	.39	**.88**		.42	.36	.78
22 They lead discussion about students' results with staff.	.52	**.87**	–	.50	.44	.78
20 Assessment data are used to plan teaching.	.39	**.80**	–	.48	.39	.66
29 Student feedback is regularly collected, analysed and acted upon.	.47	**.70**	–	.47	.40	.53
8 Meeting time is allocated to support the achievement of priority teaching and learning goals.	.60	**.64**	.46	.51	.40	.60
23 Record keeping processes are efficient and effective.	.53	**.64**	–	.56	.54	.55
Factor 3:Management of Resources α = .80						
6 Staff have equitable access to department resources.	.53	.33	**.80**	.44	.43	.75
7 Resources are allocated to building a culture focussed on student learning.	.60	.43	**.75**	.51	.54	.77
5 The department budget is aligned to priority teaching goals.	.56	.40	**.70**	.48	.48	.67
Factor 4:Goals and Expectations α = .85						
2 Teachers have performance appraisal goals and plans that relate to department and school goals.	.40	.37	–	**.87**	.35	.76
1 School goals are translated into clear department goals.	.37	.41	–	**.85**	.36	.75
3 Teachers understand how their classroom work contributes to the department and school goals.	.48	.50	–	**.85**	.48	.74
4 The department has clear and measurable academic goals.	.37	.60	.33	**.70**	–	.61
18 They make time to observe department teachers at work in their classrooms.	.57	.56	–	**.63**	.62	.52
Factor 5:Positive Learning Environment for Students and Teachers α = .87						
13 Students experience the classroom environment as supportive of their learning	.45	.37	.32	.45	**.83**	.71
19 All students have the opportunity to attempt NCEA standards that will provide access to university.	.41	.43	.30	.34	**.66**	.49
24 Student's receive high quality, timely feedback that helps them to improve their work.	.51	.61	–	.50	**.74**	.67
28 Professional development ideas are implemented and used in the classroom.	.65	.51	–	.50	**.73**	.65
10 Staff handle student discipline problems with fairness and equity.	.56	–	.41	.47	**.64**	.55
26 They take care of their own professional development needs and share their new learning with colleagues.	.63	.46	.36	.44	**.65**	.57
14 There is an expectation of consistent high quality teaching.	.58	.53	–	.63	**.63**	.61
15 They are involved in the recruitment of high quality staff for the department.	.50	–	–	.43	**.55**	.50
Total Variance	16.339	2.185	1.391	1.135	1.109	
Percentage of Variance	48.055	6.425	4.091	3.337	3.262	

Note. N = 330 Boldface indicates highest factor loadings.

Appendix E: Means and standard deviations for each factor for each department

School Code	Department	n	Collegial Working Environment	Goals and Expectations	Focus on Student Academic Results	Management of Resources	Positive Learning Environment for Students and Teachers
3A	English	12	6.36 (0.91)	6 (1.01)	6.18 (0.80)	6.29 (0.54)	6.37 (0.27)
3A	Mathematics	10	5.41 (1.9)	5.43 (1.89)	5.31 (1.72)	5.41 (2.26)	5.85 (1.03)
3A	Science	12	5.62 (1.17)	5.53 (1.00)	5.72 (0.80)	6.44 (0.34)	5.81 (0.79)
5A	English	21	5.66 (1.02)	5.78 (0.65)	5.03 (1.00)	4.74 (1.50)	5.79 (0.71)
5A	Mathematics	18	5.2 (1.48)	5.55 (1.29)	4.75 (1.6)	5.17 (1.67)	5.34 (1.35)
5A	Science	16	5.82 (0.55)	5.1 (0.61)	4.87 (1.01)	5.79 (0.78)	5.65 (0.72)
6A	English	10	3.94 (1.55)	3.57 (1.47)	4.1 (0.68)	2.72 (1.45)	4.56 (0.51)
6A	Mathematics	12	5.09 (1.29)	5.6 (0.88)	5.38 (1.33)	4.75 (1.36)	5.2 (1.35)
6A	Science	9	5.36 (1.12)	6 (0.46)	5.56 (0.88)	5.33 (1.41)	5.66 (0.86)
8C	English	8	3.95 (1.18)	5.62 (0.63)	5.52 (0.50)	2.85 (1.54)	5.1 (0.66)
8C	Mathematics	6	5.7 (0.72)	6.4 (0.58)	5.7 (0.59)	5.2 (1.59)	5.59 (0.97)
8C	Science	5	6.01 (0.53)	5.66 (0.90)	5.88 (0.53)	6.11 (0.38)	5.95 (0.26)
8D	English	8	5.8 (0.53)	6.3 (0.46)	6.05 (0.63)	5.66 (1.22)	6.22 (0.49)
8D	Mathematics	9	6.07 (0.91)	6.4 (0.85)	5.95 (0.70)	5.92 (1.69)	6.17 (0.57)
8D	Science	9	6.26 (0.66)	5.8 (0.73)	5.16 (0.50)	5.72 (0.85)	6.1 (0.43)
8E	English	14	6.04 (0.75)	5.71 (1.20)	5.48 (1.54)	5.74 (0.49)	6.15 (0.57)
8E	Mathematics	14	5.31 (0.96)	5.16 (1.37)	4.95 (1.20)	5.48 (1.12)	5.71 (0.72)
8E	Science	17	5.55 (0.80)	5.8 (0.72)	5.41 (0.56)	5.8 (0.67)	6 (0.48)
9A	English	5	5.12 (2.56)	6.16 (0.59)	6.2 (0.62)	5.5 (2.1)	5.71 (0.71)
9A	Mathematics	5	5.75 (0.31)	5.66 (0.75)	5.33 (1.42)	5.11 (1.89)	5.7 (0.43)
9A	Science	5	5.45 (0.38)	6.3 (0.70)	6.15 (0.49)	6.33 (0.47)	5.31 (0.97)
9B	English	6	6.54 (0.38)	6.66 (0.41)	6.27 (0.58)	7.00 (0.00)	6.2 (0.50)
9B	Mathematics	9	5.77 (0.66)	5.68 (0.83)	5.64 (0.52)	5.38 (0.73)	5.83 (0.57)

School Code	Department	n	Collegial Working Environment	Goals and Expectations	Focus on Student Academic Results	Management of Resources	Positive Learning Environment for Students and Teachers
9B	Science	6	5.86 (0.33)	6.01 (0.65)	5.96 (0.69)	6.6 (0.27)	5.66 (0.31)
9D	English	11	4.8 (1.72)	4.42 (1.56)	4.73 (1.76)	4.61 (1.88)	5.01 (1.91)
9D	Mathematics	16	4.96 (1.62)	5.63 (0.79)	5.17 (1.11)	5.07 (1.55)	5.69 (0.84)
9D	Science	15	5.42 (1.15)	5.54 (0.99)	5.15 (1.06)	5.57 (1.45)	6.03 (0.53)
9E	English	14	6.09 (1.68)	6 (0.66)	5.95 (1.31)	6.3 (0.99)	6.6 (0.33)
9E	Mathematics	14	5.1 (1.89)	5.44 (1.07)	5.05 (1.49)	5.8 (1.17)	5.86 (1.01)
9E	Science	11	6.06 (0.76)	5.94 (0.75)	5.5 (0.90)	5.95 (0.91)	6.26 (0.55)

Appendix F: Department academic performance measures for Phase 2 schools

Department	Average 2010 %of M/E at Level, 1, 2, 3 above or below national norm for decile	Average 2010 GPA for L1,2,3 above or below national norm for decile	Level 1 GPA 2010 compared to national norm	Progress at Level 1 measured by GPA in 2008, 2009, 2010	Percentage of M/E at Level 3 in 2009 and 2010 compared to national norm for decile	Number of scholarships in 2009 and 2010
3A Eng	**18%(21%)**	**1.43 (1.52)**	1.45(1.58)	1.48,1.30,1.45	16% (22%) 22%(24%)	1/0
3A Mathematics	19%(19%)	1.39 (1.45)	1.40(1.40)	1.29,1.45,1.40	20%(17%) 26%(18%)	1/1
3A Science	25%(25%)	1.54(1.59)	1.41(1.56)	1.31,1.45,1.41	26%(26%) 33%(27%)	0/5
5A Eng	**21%(27%)**	**1.47 (1.78)**	1.61 (1.76)	1.61,1.53,1.61	19%(27%) 18%(31%)	2/2
5A Mathematics	29%(28%)	1.77(1.77)	1.66 (1.82)	1.66,1.50,1.66	29%(23%) 26%(23%)	2/1
5A Science	**28%(30%)**	**1.70(1.76)**	1.67(1.69)	1.66,1.73,1.67	34%(30%) 26%(30%)	9/9
6A Eng	**19%(25%)**	**1.41(1.72)**	1.51(1.73)	1.38,1.67,1.51	18%(28%) 19%(29%)	1/0
6A Mathematics	26%(26%)	1.61(1.75)	1.57(1.88)	1.48,1.56,1.57	27%(23%) 25%(17%)	1/3
6A Science	**28%(31%)**	**1.64(1.82)**	1.59(1.68)	1.55,1.52,1.59	27%(31%) 24%(33%)	4/6
8C Eng	32%(32%)	1.92 (1.93)	2.00(1.98)	1.59,1.93,2.00	30%(32%) 39%(35%)	0/1
8C Mathematics	**28%(34%)**	**1.83 (2.01)**	2.00(2.15)	1.37,1.85,2.00	34%(28%) 21%(22%)	0/0
8C Science	**30%(36%)**	**1.73(2.0)**	1.68(1.90)	1.43,1.75,1.68	27%(36%) 28%(37%)	0/1
8D Eng	35%(32%)	1.93 (1.93)	2.04(1.98)	1.93,1.72,2.04	27%(32%) 33%(35%)	1/2
8D Mathematics	**45%(34%)**	**2.41(2.01)**	2.70(2.15)	2.54,2.48,2.70	41%(28%) 26%(22%)	0/1
8D Science	**41%(36%)**	**2.12 (2.00)**	2.12(1.90)	2.03,1.74,2.12	42%(36%) 42%(37%)	2/4
8E Eng	31%(32%)	1.89(1.93)	2.09(1.98)	1.96,2.03,2.09	34%(32%) 28%(35%)	0/2
8E Mathematics	**40% (34%)**	**2.13 (2.01)**	2.20 (2.15)	1.97,2.13,2.20	34%(28%) 28%(22%)	2/4
8E Science	38%(36%)	2.10(2.00)	2.04(1.90)	2.08,2.15,2.04	42%(36%) 39%(37%)	3/7
9A Eng	**44%(33%)**	**2.32 (1.95)**	2.45(2.04)	2.41,2.40,2.45	47%(34%) 41%(34%)	1/1
9A Mathematics	43%(37%)	2.05 (2.06)	2.45(2.20)	2.01,1.97,2.45	30%(28%) 36%(30%)	0/2
9A Science	37%(37%)	2.02(2.03)	1.83(1.91)	2.07,1.76,1.83	40%(36%) 43%(38%)	0/5
9B Eng	**41%(33%)**	**2.14(1.95)**	2.13(2.04)	1.96,1.93,2.13	37%(34%) 37%(34%)	3/5
9B Mathematics	38%(37%)	2.10 (2.06)	2.34(2.20)	1.95,2.15,2.34	30%(28%) 27%(30%)	2/1
9BScience	37% (37%)	2.07 (2.03)	1.90(1.91)	1.52,1.90,1.90	35%(36%) 41%(38%)	0/8
9D Eng	**30%(33%)**	**1.86 (1.95)**	2.05 (2.04)	2.20,2.15,2.05	35%(34%) 30%(34%)	3/4
9D Mathematics	41%(37%)	2.16(2.06)	2.28(2.20)	2.20,2.22,2.28	34%(28%) 36%(30%)	3/0
9D Science	39%(37%)	2.10 (2.03)	1.90 (1.91)	1.90,2.05,1.90	36%(36%) 43%(38%)	5/9
9E Eng	**44%(33%)**	**2.29(1.95)**	2.51(2.04)	2.25,2.51,2.51	44%(34%) 40%(34%)	16/10
9E Mathematics	**53%(37%)**	**2.48(2.06)**	2.79 (2.20)	2.52,2.55,2.79	47%(28%) 37%(30%)	4/3
9E Science	**50% (37%)**	**2.33(2.03)**	2.20(1.91)	2.37,2.27,2.23	44%(36%) 48%(38%)	24/14

Notes.

Bracketed Numbers are the national norm for decile

🟥	Underperforming departments in student academic outcome measures
🟦	Overperforming departments in student academic outcome measures
⬜	Results of student academic outcomes are similar to norm for decile

References

Adams, J. E., & Kirst, M. W. (1999). New demands and concepts for educational accountability: Striving for results in an era of accountability. In J. Murphy & K. Louis (Eds.), *Handbook of research on educational administration* (2nd ed., pp. 463-489). San Francisco: Jossey-Bass.

Agnew, S. (2010). Accounting for the NCEA: Has the transition to standards-based assessment achieved its objectives? *Australasian Accounting Business and Finance Journal, 4*(4), 87–102.

Alton-Lee, A., & Rowe, K., J. (2007). *Policy Brief: Impact of teachers and schools on variance in student outcomes.* Retrieved from http://educationcounts.edcentre.govt.nz/goto/BES.

Ball, S., & Lacey, C. (1980). Subject disciplines as the opportunity for group action: A measured critique of subject sub-cultures. In P. Woods (Ed.), *Teacher strategies* (pp. 149–177). London, England: Croom Helm.

Bendikson, L., Hattie, J., & Robinson, V. M. J. (2011). Identifying the comparative academic performance of secondary schools. *Journal of Educational Administration, 49*(4), 433–449. doi: 10.1108/09578231111146498

Bennett, N. (1995). *Managing professional teachers: Middle management in primary and secondary schools.* London, England: Paul Chapman.

Bennett, N., Woods, P., Wise, C., & Newton, W. (2007). Understandings of middle leadership in secondary schools: A review of empirical research. *School Leadership & Management, 27*(5), 453–470. doi: 10.1080/13632430701606137

Bidwell, C. E., & Yasumoto, J. Y. (1999). The collegial focus: Teaching fields, collegial relationships, and instructional practice in American high schools. *Sociology of Education, 72*(4), 234–256. doi: 10.2307/2673155

Bird, T., & Little, J. W. (1986). How schools organise the teaching occupation. *The Elementary School Journal, 86*, 493–512. doi: 10.1086/461464

Blase, J. (1995). *Democratic principals in action: Eight pioneers.* Thousand Oaks, CA: Corwin Press.

Blase, J., & Anderson, G. (1995). *The micropolitics of educational leadership: From control to empowerment.* London, England: Cassell.

Bolam, R., & Turner, C. (2003). Heads of secondary school subject departments and the improvement of teaching and learning. In M. Wallace & L. Poulson (Eds.), *Educational leadership and management* (pp. 133–148). London, England: Sage.

Bosker, R. J., Kremers, J. J., & Lugthart. (1990). School instruction effects on mathematics achievement. *School Effectiveness & School Improvement, 1*(4), 233–248. doi: 10.1080/0924345900010401

Brown, G. T. L. (2004). Measuring attitude with positively packed self-report ratings: Comparison of agreement and frequency scales. *Psychological Reports, 94*, 1015–1024. doi: 10.2466/pr0.94.3.1015-1024

Brown, M., & Rutherford, D. (1998). Changing roles and raising standards: New challenges for heads of department. *School Leadership & Management, 18*(1), 75–88. doi: 10.1080/13632439869781

Brown, M., & Rutherford, D. (1999). A re-appraisal of the role of the head of department in UK secondary schools. *Journal of Educational Administration, 37*(3), 229–242. doi: 10.1108/09578239910275472

Brown, M., Rutherford, D., & Boyle, B. (2000). Leadership for school improvement: The role of the head of department in UK secondary schools. *School Effectiveness and School Improvement, 11*(2), 237–258. doi: 10.1076/0924-3453(200006)11

Busher, H. (1992). *The politics of working in secondary school: Some teacher perspectives on their schools as organisations.* (Unpublished doctoral dissertation), University of Leeds, England.

Busher, H., & Harris, A. (1999). Leadership of school subject areas: Tensions and dimensions of managing in the middle. *School Leadership & Management, 19*(3), 305–317. doi: 10.1080/13632439969069

Busher, H., Harris, A., & Wise, C. (2000). *Subject leadership and school improvement.* London, England: Paul Chapman.

Campbell, P., & Southworth, G. (1992). Rethinking collegiality: Teachers views. In N. Bennett, M. Crawford & C. Riches (Eds.), *Managing change in education: Individual and organizational perspectives* (pp. 61–79). London, England: Paul Chapman.

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis, 24*(4), 305-331.

Chapman, C., Mongon, D., Roxby, S., & Manns, S. (2009). *Training and Development Agency for Schools and National College of School Leadership in-school variation project.* Manchester, England: University of Manchester and Price Waterhouse Coopers.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Earlbaum.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45*(12), 1304–1312. doi: 10.1037/0003-066X.45.12.1304

Coleman, J. S., Campbell, E. Q., Hobson, C. F., McPartland, A. M., Mood, A. M., Weinfield, F. D., & York, R. L. (1966). *Equality of educational opportunity.* Washington, DC: Department of Health, Education and Welfare.

Cooper, H., & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis.* New York, NY: Russell Sage Foundation.

Creemers, B. P. M. (1994). *The effective classroom.* London, England: Cassell.

Darling-Hammond, L. (1995). Inequality and access to knowledge. In J. A. Banks (Ed.), *The handbook of research on multicultural education.* New York, NY: Macmillan.

Day, C., & Leithwood, K. (2008). The impact of school leadership on pupil outcomes. *School Leadership and Management, 28*(1), 1–4. doi: 10.1080/13632430701800045

Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K., Qing, G., . . . Kington, A. (2009). The impact of school leadership on pupil outcomes: Final report: University of Nottingham.

Dinham, S. (2007). The secondary Head of Department and the achievement of exceptional student outcomes. *Journal of Educational Administration, 45*(1), 62–79. doi: 10.1108/09578230710722458

Dinham, S., & Rowe, K. (2007). Teaching and learning in middle schooling: A review of the literature (Report to the New Zealand Ministry of Education). Melbourne, Australia.

Duke, D., & Leithwood, K. (Eds.). (1994). *Defining effective leadership for Connecticut's schools.* Hartford, CT: University of Connecticut.

Earle, D. (2007). *Te whai i nga taumata atakura: Supporting Maori achievement in bachelors degrees.* Wellington, New Zealand: Ministry of Education, New Zealand.

ERO. (2011a). Information for parents. Retrieved April 5, 2011, from http://www.ero.govt.nz/Review-Process/For-Parents

ERO. (2011b). School reports/About the school. Retrieved April 6, 2011, from http://www.ero.govt.nz/content/search

Feist, C. (2008). Positions and dispositions: Locating leaders within New Zealand secondary school faculties. *Journal of Educational Leadership, 23*(2), 60–73.

Firestone, W. A. (1996). Images of teaching and proposals for reform: A comparison of ideas from cognitive and organizational research. *Education Administration Quarterly, 32*, 209–235. doi: 10.1177/0013161X96032002003

Fiske, E. B., & Ladd, H. F. (2001). The uneven playing field of school choice: Evidence from New Zealand. *Journal of Policy Analysis and Management, 20*(1), 43–63. doi: 10.1002/1520-6688(200124)20

Fitz-Gibbon, C. T. (1991). Multilevel modelling in an indicator system. In S. Raudenbush, W & J. D. Willms (Eds.), *Pupils, classrooms and schools: International studies from a multilevel perspective* (pp. 67–83). San Diego, CA: Academic Press.

Fitz-Gibbon, C. T. (1992). School effects at A level: Genesis of an information system? In D. Reynolds & P. Cuttance (Eds.), *School effectiveness: Research, policy and practice* (pp. 96–120). London, England: Cassell.

Fitz-Gibbon, C. T., & Vincent, L. (1997). Difficulties regarding subject difficulties: Developing reasonable explanations for observable data. *Oxford Review of Education, 23*(3), 291–298. doi: 10.1080/0305498970230302

Fitzgerald, T., & Gunter, H. (2006). Leading learning: Middle leadership in schools in England and New Zealand. *Management in Education, 20*(3), 6–8. doi: 10.1177/0892020606200030201

Fletcher-Campbell, F. (2003). Promotion to middle management: Some practitioners' perceptions. *Educational Research, 45*(1), 1–15. doi: 10.1080/0013188032000086082

Fullan, M. (1999). *Change forces: The sequel*. London, England: Falmer.

Glass, G. V., & Hopkins, K. D. (1984). *Statistical methods in education and psychology* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Glickman, C., Gordon, S., & Ross-Gordon, J. (2001). *SuperVision and instructional leadership: A developmental approach*. Boston, MA: Allyn & Bacon.

Glover, D., Gleeson, D., Gough, G., & Johnson, M. (1998). The meaning of management: The development needs of middle managers in secondary schools. *Educational Management Administration and Leadership, 26*(3), 279–292. doi: 10.1177/0263211X98263006

Glover, D., Miller, D., Gambling, M., Gough, G., & Johnson, M. (1999). As others see us: Senior management and subject staff perceptions of the work effectiveness of subject leaders in secondary schools. *School Leadership & Management, 19*(3), 331–344. doi: 10.1080/13632439969087

Goldring, E., Porter, P. C., Murphy, J., Elliott, S. N., & Cravens, X. (2007). Assessing learning-centred leadership: Connections to research, professional standards and current practices: Learning Science Institute.

Goldstein, H. (1987). *Multi-level models in educational and social research*. London, England: Griffin.

Goldstein, H. (1995). *Multi-level statistical models*. London, England: Edward Arnold.

Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness & School Improvement, 8*(4), 369–395. doi: 10.1080/0924345970080401

Goldstein, H. (2001). Using pupil performance data for judging schools and teachers: Scope and limitations. *British Educational Research Journal, 27*(4), 433–442. doi: 10.1080/01411920120071443

Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., & Nuttall, D. (1993). A multilevel analysis of school examination results. *Oxford Review of Education, 19*(4), 425–433. doi: 10.1080/0305498930190401

Good, T. L., & Brophy, J. E. (2003). *Looking into classrooms* (9th ed.). Boston, MA: Allyn & Bacon.

Grossman, P. L., & Stodolsky, S. S. (1994). Considerations of content and the considerations of secondary school teaching. In L. Darling-Hammond (Ed.), *Review of research in education* (Vol. 20, pp. 179–221). Washington DC: American Educational Research Association.

Grossman, P. L., & Stodolsky, S. S. (1995). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher, 24*(8), 5–23. doi: 10.3102/0013189x024008005

Grossman, P. L., Stodolsky, S. S., & Knapp, M. S. (2004). Making subject matter part of the equation: The intersection of policy and content: Centre for the Study of Teaching and Policy.

Hallinger, P., & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness:1980–1995. *School Effectiveness and School Improvement, 9*(2), 157–191. doi: 10.1080/0924345980090203

Hanushek, E., & Raymond, M. E. (2004). Does school accountability lead to improved student performance?

Hargreaves, A. (1984). Experience counts, theory doesn't: How teachers talk about their work. *Sociology of Education, 57*, 244–254. doi: 10.2307/2112428

Hargreaves, A. (1991). Contrived collegiality: The micropolitics of teacher collaboration. In J. Blase (Ed.), *The politics of life in schools: Power, conflict and cooperation* (pp. 46–72). Newbury Park, CA: Sage.

Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in the postmodern age.* London, England: Cassell.

Hargreaves, A. (1999). Fielding errors? Deepening the debate about teacher collaboration and collegiality: Response to Fielding. *The Australian Educational Researcher, 26*(2), 45–53. doi: 10.1007/bf03219694

Hargreaves, A., & Dawe, R. (1990). Paths of professional development: Contrived collegiality, collaborative culture, and the case of peer coaching. *Teaching and Teacher Education, 6*(3), 227–241. doi: 10.1016/0742-051x(90)90015-w

Harris, A. (1997, September). *Differential departmental culture and strategies for change.* Paper presented at the British Educational Research Association, York, England.

Harris, A. (1998). Improving ineffective departments in secondary schools: Strategies for change and development. *Education Management and Administration, 26*(3), 269–278. doi: 10.1177/0263211X98263005

Harris, A. (1999). *Effective subject leadership in secondary schools.* London, England: David Fulton.

Harris, A. (2004). Distributed leadership and school improvement: Leading or misleading? *Educational Management Administration and Leadership, 32*(1), 11–24. doi: 10.1177/1741143204039297

Harris, A., & Chapman, C. (2002, January). *Democratic leadership for school improvement.* Paper presented at the International Congress for School Effectiveness and Improvement, Copenhagen, Denmark.

Harris, A., Jamieson, I. M., & Russ, J. (1995). A study of effective departments in secondary schools. *School Organisation, 15*(3), 283–300.

Harris, A., & Muijs, D. (2005). *Improving schools through teacher leadership.* New York, NY: Open University Press.

Hattie, J. (1992). *Influences on student learning.* Paper presented at the Inaugural Lecture, University of Auckland, New Zealand. http://74.125.155.132/scholar?q=cache:Xti5q03VZ-wJ:scholar.google.com/+hattie+influences+on+student+learning&hl=en&as_sdt=0,5

Hattie, J. (2002). Schools like mine: Cluster analysis of New Zealand schools. *Technical Report 14, Project AsTTle.* New Zealand: University of Auckland.

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

Heck, R. H. (1992). Principals' instructional leadership and school performance: implications for policy development. *Educational Evaluation and Policy Analysis, 14*(1), 21–34. doi: 10.1080/0924345910020204

Heck, R. H., Marcoulides, G. A., & Lang, P. (1991). Principal instructional leadership and school achievement: The application of discriminant techniques. *School Effectiveness and School Improvement, 2*(2), 115–135. doi: 10.1080/0924345910020204

Hill, D. (1995). The strong department: Building the department as learning community. In L. Siskin & J. W. Little (Eds.), *The subjects in question* (pp. 123–139). New York, NY: Teachers College Press.

Hill, P. (1998). Shaking the foundations: Research driven from school reform. *School Effectiveness and School Improvement, 9*(4), 419–436. doi: 10.1080/0924345980090404

Hill, P., Holmes-Smith, P., & Rowe, K. L. (1993). School and teacher effectiveness in Victoria. Key findings from Phase 1 of the Victorian Quality Schools' Project. Australia: University of Melbourne.

Hill, P., W, & Rowe, K., J. (1996). Multilevel modelling in school effectiveness research. *School Effectiveness & School Improvement, 7*(1), 1–34. doi: 10.1080/0924345960070101

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. Wellington, New Zealand: New Zealand Council for Educational Research.

Hofman, R. H., & Hofman, W. H. A. (2000). *Context effects and governance of schooling.* Paper presented at the ISA conference, Groningen, Netherlands.

Hofman, R. H., Hofman, W. H. A., & Guldemond, H. (2001). The effectiveness of cohesive schools. *International Journal of Leadership in Education, 4*(2), 115–135. doi: 10.1080/13603120117137

James, C., & Aubrey-Hopkins, J. (2003). The leadership authority of educational middle managers: The case of subject leaders in secondary schools in Wales. *International Studies in Educational Administration, 31*(1), 50–64.

Johnson, S. M. (1990). The primacy and potential of high school departments. In M. W. McLaughlin, J. E. Talbert & N. Bascia (Eds.), *The contexts of teaching in secondary schools: Teachers' realities* (pp. 167–184). New York, NY: Teachers College Press.

Kannapel, J., Aargard, 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, IL: University of Chicago.

Kilpatrick, V. E. (1905). Departmental teaching in elementary schools. *Educational Review, March*, 468–485.

Kline, P. (1994). *An easy guide to factor analysis.* London, England: Routledge.

Ladd, H. F., & Fiske, E. B. (2003). Does competition improve teaching and learning? Evidence from New Zealand. *Educational Evaluation and Policy Analysis, 25*(1), 97–112. doi: 10.3102/01623737025001095

Lam, T. C. M., & Klockers, A. J. (1982). Anchor point effects on the equivalence of questionnaire items. *Journal of Educational Measurement, 19*, 317–322. doi: 10.1111/j.1745-3984.1982.tb00137.x

Lamb, S., & Fullarton, S. (2000). *Factors affecting mathematics achievement in primary and secondary schools: Results from TIMSS*. Paper presented at the Twenty-Third Annual Conference of the Mathematics Education Research Group of Australia, Perth, Australia.

Latham, G., P, & Locke, E., A. (2006). Enhancing the benefits and overcoming the pitfalls of goal setting. *Organisational Dynamics, 35*(4), 332–340. doi: 10.1016/j.orgdyn.2006.08.008

Lee, V. E., Bryk, A. S., & Smith, J. B. (1993). The Organization of effective secondary schools. *Review of Research in Education, 19*, 171–267.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership and Management, 28*, 27–42. doi: 10.1080/13632430701800060

Levin, B. (2008). *How to change 5000 schools: A practical and positive approach for leading change at every level*. Cambridge, MA: Harvard Education Press.

Lipman, P. (1997). Restructuring in context: A case study of teacher participation and the dynamics of ideology, race, and power. *American Educational Research Journal, 34*(1), 3–37. doi: 10.3102/00028312034001003

Little, J. W. (1993). Professional community in comprehensive high schools: The two worlds of academic and vocational teachers. In J. W. Little & M. W. McLaughlin (Eds.), *Teachers' work: Individuals, colleagues and contexts* (pp. 137–163). New York, NY: Teachers College Press.

Luyten, H. (1994). Stability of school effects in Dutch secondary education: The impact of variance across subjects and years. *International Journal of Educational Research, 21*(2), 197–216. doi: 10.1016/0883-0355(94)90032-9

Madaus, G., & Clarke, M. (2001). The adverse impact of high-stakes testing on minority students: Evidence of 100 years of test data. In G. Orfield & M. Kornhaber (Eds.), *Raising standards or raising barriers? Inequality and high-stakes testing in public education*. Boston, MA: Century Foundation Press.

Madjar, I., McKinley, E., Deynzer, M., & Van der Merwe, A. (2010). Stumbling blocks or stepping stones? Students' experience of transition from low-mid decile schools to university. Auckland, New Zealand: University of Auckland.

Mann, K. V. (1998). Not another survey! Using questionnaires effectively in needs assessment. *Journal of Continuing Education in the Health Professions, 18*(3), 142–149. doi: 10.1002/chp.1340180303

Marsh, H. (1990). Self description questionnaire 1 - Manual *Self research centre instrument packages*. Oxford, England: University of Oxford.

McCauley, C. D., & Moxley, R. S. (1996). Developmental 360: How feedback can make managers more effective. *Career Development International, 1*(3), 15–19. doi: 10.1108/13620439610118555

McLaughlin, M. W., & Talbert, J., E. (2007). Building professional learning communities in high schools: Challenges and promising practices. In L. Stoll & L. Seashore, K (Eds.), *Professional learning communities: Divergence, depth and dilemmas* (pp. 151–165). Maidenhead, England: Open University Press.

McNaughton, S., Robinson, V. M. J., & Timperley, H. (2011). Building capacity in a self-managing school system. *Journal of Educational Administration, 49*(6), 720–738. doi: 10.1108/09578231111174848

Meagher-Lundberg, P. (2000). Comparison variables useful to teachers in analysing assessment results *Technical Report 1*: Project asTTLE, University of Auckland.

Metcalfe, C., & Russell, S. (1997). *The role of subject leaders in monitoring the work of teachers in secondary schools: The quest for consistency?* Paper presented at the British Educational Research Association, York, England.

Meyer, L., McClure, J., Walkey, F., McKenzie, L., & Weir, K. (2006). The impact of the NCEA on student motivation. Wellington, New Zealand: Victoria University.

Ministry of Education New Zealand. (2007). *The New Zealand curriculum.* Wellington, New Zealand: Learning Media.

Ministry of Education New Zealand. (2009a). *Ministry of Education Statement of Intent 2007–2012.* Wellington, New Zealand: New Zealand Government Retrieved from http://www.minedu.govt.nz/theMinistry/PublicationsAndResources/StatementOfInte nt/SOI2007/Appendix_1_Key_documents.aspx.

Ministry of Education New Zealand. (2009b). PISA 2006. Retrieved May 16, 2009, from http://www.educationcounts.govt.nz/themes/research/pisa_research/pisa_2006

Ministry of Education New Zealand. (2010a). School leavers with NCEA Level 1 or above. *Indicators.* Retrieved March 14, 2011, from http://www.educationcounts.govt.nz/indicators/education_and_learning_outcomes/ qualifications/28788

Ministry of Education New Zealand. (2010b). School leavers with NCEA Level 2 or above. *Indicators.* Retrieved March 14, 2011, from http://www.educationcounts.govt.nz/indicators/education_and_learning_outcomes/ qualifications/1781

Ministry of Education New Zealand. (2011). Briefing to the incoming Minister. Wellington, New Zealand: Ministry of Education.

Murphy, J. (1992). *The landscape of leadership preparation: Reframing the education of school administrators.* Newbury Park, CA: Corwin Press.

Murphy, J. (1994). Transformational change and the evolving role of the principal: Early empirical evidence. In J. Murphy & K. S. Louis (Eds.), *Reshaping the principalship: Insights from transformational reform efforts*: Newbury Park, CA: Corwin Press.

NCSL. (2006). Narrowing the gap: Reducing within-school variation in pupil outcomes. Nottingham, England: National College for School Leadership.

New Zealand Teachers Council. (2012). *Policy requirements and guidlines for registered New Zealand teachers.* Wellington: New Zealand Government Retrieved from http://www.teacherscouncil.govt.nz/required/Registration_Policy_2012.pdf.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257. doi: 10.3102/01623737026003237

nz.ratemyteachers.com. (2010). Retrieved April 5, 2011, from http://nz.ratemyteachers.com/

NZQA. (2009). NCEA: The details. Retrieved September 20, 2009, from http://www.nzqa.govt.nz/ncea/about/overview/ncea-thedetails.html

NZQA. (2011a). How NCEA works. Retrieved March 7, 2011, from http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/understanding-ncea/how-ncea-works/

NZQA. (2011b). Scholarship. Retrieved May 6, 2011, from http://www.nzqa.govt.nz/qualifications-standards/awards/scholarship/

NZQA. (2011c). Search standards and assessment. Retrieved March 28, 2011, from http://www.nzqa.govt.nz/ncea/assessment/search.do?query=English&view=achievements&level=01

NZQA. (2011d). Standards. Retrieved March 7, 2011, from http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/understanding-ncea/how-ncea-works/standards/

OECD. (2000). Knowledge and skills for life: First results from the OECD Programme for International Student Assessment (PISA) 2000.

OECD. (2003). Learning for tomorrow's world: First results from PISA 2003 (pp. 162).

OECD. (2010a). PISA 2009 results: Overcoming social background (Vol. 2): OECD.

OECD. (2010b). PISA 2009 results: What students know and can do: Student performance in reading, mathematics and science (Vol. 1).

Oppenheim, A. N. (1998). *Questionnaire design, interviewing and attitude measurement*: Continuum International Publishing Group.

Ortiz, A. (2000). *Including students with special needs in standards-based reform: Issues associated with the alignment of standards, curriculum and instruction*. Paper presented at the Third Mid-Continent Research for Education and Learning, Denver, Colorado.

Philips, D. (2003). Lessons from New Zealand's national qualifications framework. *Journal of Education and Work, 16*(3), 289–304. doi: 10.1080/1363908032000099458

Philips, M. (1997). What makes schools effective? A comparison of the relationships of communitarian climate and academic climate to mathematics achievement and attendance during middle school. *American Educational Research Journal, 34*, 633-662.

Porter, A. C., Murphy, J., Goldring, E., Elliott, S. N., Polikoff, M. S., & May, H. (2008). Vanderbilt assessment of leadership in education technical manual Version 1.0: Vanderbilt University, Peabody College.

Raudenbush, S., W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher, 34*(5), 35–21.

Raudenbush, S., W, & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioural Statistics, 20*(4), 307–335.

Rawlins, P., Brandon, J., Chapman, J., Leach, L., Neutze, G., Scott, A., & Zepke, N. (2005). Standards-based assessment in the senior secondary school: A review of the literature. *New Zealand Journal of Teachers Work, 2*(2), 107–115.

Reynolds, D. (2004, July). *Within school variation: Its extent and causes.* Paper presented at the National Within School Variation Conference.

Reynolds, D. (2007). *Schools learning from their best: The within school variation project.* Nottingham, England: National College of School Leadership.

Robinson, V. M. J. (2007). School leadership and student outcomes: Identifying what works and why. *ACEL Monograph Series, 41*, 27.

Robinson, V. M. J., Hohepa, M., & Lloyd, C. (2009). *School leadership and student outcomes: Identifying what works and why.* New Zealand: University of Auckland.

Robinson, V. M. J., & Timperley, H. (2007). The leadership of the improvement of teaching and learning: Lessons from initiatives with positive outcomes for students. *Australian Journal of Education, 51*(3), 247–262.

Rowan, B., Raudenbush, S. W., & Kang, S. J. (1991). Organizational design in high schools: A multilevel analysis. *American Journal of Education, 99*(2), 238–266. doi: 10.1086/443980

Rowe, K. L., & Hill, P. (1997). *Simultaneous estimation of multilevel structural equations to model students' educational progress.* Paper presented at the Tenth International Congress for School Effectiveness and Improvement, Memphis, Tennessee.

Sammons, P. (1996). Complexities in the judgement of school effectiveness. *Educational Research and Evaluation, 2*(2), 113–149. doi: 10.1080/1380361960020201

Sammons, P., Hillman, J., & Mortimore, P. (1994). Key characteristics of effective schools: A review of school effectiveness research: Institute of Education, University of London.

Sammons, P., & Luyten, H. (2009). Editorial article for special issue on alternative methods for assessing school effects and schooling effects. *School Effectiveness & School Improvement, 20*(2), 133–143. doi: 10.1080/09243450902879753

Sammons, P., Thomas, S., & Mortimore, P. (1996). Promoting school and departmental effectiveness. *Management in Education, 10*(1), 22–24. doi: 10.1177/089202069601000110

Sammons, P., Thomas, S., & Mortimore, P. (1997). *Forging links: Effective schools and effective departments.* London, England: Paul Chapman.

Sammons, P., Thomas, S., Mortimore, P., Walker, A., Cairns, R., & Bausor, J. (1998). Understanding differences in academic effectiveness: Practitioner's views. *School Effectiveness and School Improvement, 9*(3), 286–309. doi: 10.1080/0924345980090302

Scheerens, J. (1992). *Effective schooling: Research, theory and practice.* London, England: Cassell.

Schein, E. (1990). *Organisational culture and leadership: A dynamic view*. San Francisco, CA: Jossey Bass.

Shulruf, B., Hattie, J., & Tumen, S. (2008). The predictability of enrolment and first year university results from secondary school performance. *Studies in Higher Education, 33*(6), 685–698.

Silins, H., & Mulford, B. (2002). Leadership and school results. In K. Leithwood & P. Hallinger (Eds.), *Second international handbook of educational leadership and administration* (pp. 561–612). Norwell, MA: Kluwer Academic.

Siskin, L. (1991). Departments as different worlds: Subject subcultures in secondary schools. *Educational Administration Quarterly, 27*(2), 134–160. doi: 10.1177/0013161x91027002003

Siskin, L. (1994). *Realms of knowledge: Academic departments in secondary schools*. Washington, DC: Falmer Press.

Siskin, L. (1995). Subject divisions. In L. Siskin & J. W. Little (Eds.), *The subjects in question* (pp. 23–47). New York, NY: Teachers College Press.

Siskin, L., & Little, J. W. (1995). The subject department: Continuities and critiques. In L. Siskin & J. W. Little (Eds.), *The subjects in question* (pp. 1–22). New York, NY: Columbia University.

Smith, D. J., & Tomlinson, S. (1989). *The school effect: A study of multi-racial comprehensives*. London, England: Policy Studies Institute.

Spillane, J. P., Halverson, R., & Diamond, J., B. (2001). Investigating school leadership practice: A distributed perspective. *Research News and Comment,* 23–28.

Spillane, J. P., & Seashore, L. K. (2002). School improvement processes and practices: Professional learning for building instructional capacity *Yearbook of the National Society for the Study of Education, 101*(1), 83–104.

Statistics New Zealand, & Ministry of Pacific Island Affairs. (2010). *Education and pacific peoples in New Zealand*. Wellington, New Zealand: Statistics New Zealand and Ministry of Pacific Island Affairs.

Stodolsky, S. S., & Grossman, P. L. (1995). The impact of subject matter on curricular activity: An analysis of five academic subjects. *American Educational Research Journal, 32*(2), 227–249. doi: 10.3102/00028312032002227

Stoll, L., & Fink, D. (1996). *Changing our schools: Linking school effectiveness and school improvement*. Buckingham, England: Open University Press.

Strathdee, R. (2003). The qualifications framework in New Zealand: Reproducing existing inequalities or disrupting the positional conflict for credentials. *Journal of Education and Work, 16*(2), 147–164.

Stringfield, S. (1994). The analysis of large databases in school effectiveness research. In D. Reynolds, B. P. M. Creemers, E. C. Nesselrodt, S. Schaffer, S. Stringfield & C. Teddlie (Eds.), *Advances in school effectiveness research and practice* (pp. 52–72). Oxford, England: Pergamon Press.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology* (1st ed.). San Francisco, CA: Jossey-Bass Publishers.

Teddlie, C., & Reynolds, D. (2000). *International handbook of school effectiveness research*. London, England: Falmer.

Teddlie, C., & Stringfield, S. (1993). *Schools make a difference: Lessons learned from a 10 year study.* New York, NY: New York Teachers College Press.

Telford, M., & May, S. (2010). PISA 2009: Our 21st century learners at age 15. Wellington, New Zealand: Ministry of Education.

Thomas, S., Sammons, P., & Mortimore, P. (1995). *Stability and consistency in secondary school effects in student outcomes over 3 years.* Paper presented at the International Congress for School Effectiveness and Improvement, Leeuwarden, Netherlands.

Thomas, S., Sammons, P., Mortimore, P., & Smees, R. (1997a). Differential secondary school effectiveness: Comparing the performance of different pupil groups. *British Educational Research Journal, 23*(4), 451–469.

Thomas, S., Sammons, P., Mortimore, P., & Smees, R. (1997b). Stability and consistency in secondary schools' effects on students' GCSE outcomes over three years. *School Effectiveness & School Improvement, 8*, 169–197. doi: 10.1080/0924345970080201

Thrupp, M., & Willmott, R. (2003). *Education management in managerialist times: Beyond the textural apologists*. Maidenhead, England: Open University Press.

Timperley, H. (2008). Teacher professional learning and development. *International Academy of Education, 18*, 1–22.

Timperley, H., & Robinson, V. M. J. (1998). Collegiality in schools: Its nature and implications for problem solving. *Educational Administration Quarterly, 34*(Suppl. 1), 608–629. doi: 10.1177/0013161x980341003

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2008). Teacher professional learning and development: Best Evidence Synthesis [BES]. Wellington, New Zealand: Ministry of Education.

Tyack, D. B. (1974). *The one best system: A history of American urban education*. Cambridge, MA: Harvard University Press.

Tymms, P. (1992). The relative effectiveness of post-16 institutions in England (Including Assisted Places Scheme Schools). *British Educational Research Journal, 18*(2), 175–192.

Tymms, P. (1993). Accountability: Can it be fair? *Oxford Review of Education, 19*, 291–299.

Tymms, P., & Fitz-Gibbon, C. T. (1990). *The stability of school effectiveness indicators.* Paper presented at the BERA conference, London, England.

Tymms, P., & Fitz-Gibbon, C. T. (2001). Standards, achievement and educational performance: A cause for celebration? In R. Phillips & J. Furlong (Eds.), *Education,*

reform and the state: Twenty-five years of politics, policy and practice (pp. 156–173). London, England: Routledge/Falmer.

Weick, K. E. (1976). Educational organisations as loosley coupled systems. *Administrative Science Quarterly, 21*, 1–19.

Willms, J. D. (1992). *Monitoring school performance: A non-technical guide for educational administrators.* Lewes, England: Falmer Press.

Willms, J. D. (2000). Monitoring school performance for standards-based reform. *Evaluation and Research Education, 14*(3&4), 237–253.

Willms, J. D., & Cuttance, P. (1985). School effects in Scottish secondary schools. *British Journal of Sociology of Education, 6*(3), 289–306.

Wilson, S., & McGregor, C. (2011, June 27, 2011). The best schools in Auckland. *Metro Magazine, July-August,* 44–62.

Wise, C. (2001). The monitoring role of the academic middle manager in secondary schools. *Educational Management Administration & Leadership, 29*(3), 333–341. doi: 10.1177/0263211x010293007

Witziers, B., Bosker, R. J., & Kruger, M. L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly, 39*(3), 398–425.

Witziers, B., Sleegers, P., & Imants, J. (1999). Departments as teams: Functioning, variations and alternatives. *School Leadership & Management, 19*(3), 293–304. doi: 10.1080/13632439969050

Wylie, C. (2010, April). *Educational leadership practices survey: A tool for schools and policy makers.* Paper presented at the NZEALS Conference, Christchurch, New Zealand.