MINISTRY OF EDUCATION
Te Tāhuhu o te Mātauranga

New Zealand

Shifting Balances

The Impact of Level 1 NCEA on Teaching of Mathematics and Science

Report to the Ministry of Education

Rosemary Hipkins and Alex Neill

RESEARCH DIVISION
Wāhanga Mahi Rangahau

ISBN 0-478-13542-4

Web Copy ISBN 0-78-13539-4

Research reports are also available on the Ministry's website: www.minedu.govt.nz under the Research heading.

Opinions expressed in this report are those of the authors and do not necessarily coincide with those of the Ministry of Education

Shifting Balances

The Impact of Level 1 NCEA on the Teaching of Mathematics and Science

Report to Ministry of Education

Rosemary Hipkins and Alex Neill

December 2003

NEW ZEALAND COUNCIL FOR EDUCATIONAL RESEARCH

TE RÜNANGA O AOTEAROA MÖ TE RANGAHAU I TE MÄTAURANGA

WELLINGTON

2003

Acknowledgments

First and foremost we thank the mathematics and science teachers who made time in their busy working days to generously share their professional expertise and NCEA implementation experiences. We are grateful to their principals who gave permission for us to work in the study schools. We also thank participants in the two regional training days we attended who gave us time and welcome advice as we shaped the project.

The "best practice" findings from the *Science in Schools* research of Associate Professor Russell Tytler and his team at Deakin University in Australia were particularly helpful as we designed our research methodology.

Our research was supported by many people on the NZCER staff. We thank Cathy Wylie for her wise counsel and overview of the project. She and Karyn Dunn led the *NZCER National Survey*, from which we present some preliminary NCEA-related findings. Edith Hodgen and Hiliary Ferrral provided statistical advice and support, Christine Williams and Suzanne Hay provided research support, Kristina Louis was our library sleuth, and Bev Webber provided editorial support.

This project was funded by the Ministry of Education.

Table of Contents

List of Tables

List of Figures

Appendices

Executive Summary

Setting the scene

This research was carried out in the year following the first year of implementation of the NCEA at level 1. The curriculum areas of mathematics and science were chosen in the context of early indications that teachers of these subjects generally held more reservations about the NCEA than did teachers in other subject areas. The research sought to establish whether positive changes in mathematics and science teaching related to the NCEA implementation could be identified and documented, notwithstanding the concerns being expressed by a number of teachers of these subject areas.

Introduction to the project

This small-scale project investigated changes in teaching and learning in 18 case study schools, nine in mathematics and nine in science, as the NCEA implementation beds in at Year 11. Schools were representative of a range of school types — single sex and co-educational, high decile and low, city, town and rural, large and small. The teachers were mostly highly experienced, and were nominated for participation because they were seen to be effectively implementing the NCEA in their schools. Some were more supportive of the NCEA than others. (The recently completed *NZCER National Survey* found that mathematics and science teachers are somewhat less likely to see the NCEA making changes to learning than teachers of other curriculum areas.) All the study teachers were determined to make the NCEA work for their students.

Teachers in each school completed a self-reflection sheet designed to capture changes in their practice — whether positive or negative. Each teacher's response patterns identified their priorities and pre/post-NCEA practices for 19 identified aspects of "best practice" that were adapted from the findings of the Australian Science in Schools (SIS) project. Open-ended interview questions were used to describe aspects of the implementation context in each school, and to probe teachers' beliefs about student learning and assessment for qualifications in the NCEA regime.

This research provides an in-depth analysis of the dynamics of change in the study teachers' mathematics and science classrooms in response to the NCEA implementation. We have found a series of inter-related changes. While some of these seem to have readjusted existing balances in classroom practices, with little real change overall, others have intersected with different types of professional development initiatives underway in schools and positive changes have occurred.

The research questions

1. As a result of the introduction of the internally assessed achievement standards, are there identifiable changes in the content, structure, and balance within programmes for Maths and Sciences?
2. Are there identifiable changes in teaching and learning styles used within Maths and Science programmes that support the development of practical skills, or that allow teachers to address students' attitudes and values relevant to the subject area?
3. What case study/best practice lessons can be drawn from 2002/2003 practice in Maths and Science programmes for NCEA level 1?

The content and structure of mathematics and science programmes

All schools have modified the curriculum they offer in mathematics or science, and more such decisions are pending. The main reason for reshaping curriculum content is to reduce time pressures teachers perceive to have been exacerbated by the NCEA, particularly as they accommodate new internal assessment practices, and prepare their students for external examinations in an as yet unfamiliar format. For both mathematics and science, a number of schools have dropped, or are considering dropping, at least one internally assessed standard. Some schools have also selectively dropped externally assessed standards in both curriculum areas. Since there is no single pattern to these changes, the proportion of internally and externally assessed credits that students can gain in each subject varies between subjects and between-schools.

Internal assessments are typically carried out under strictly supervised conditions, rather than as part of the learning activities of the classroom. This practice has added substantially to teacher workloads.

In most schools, preparation for standards-based assessment is now beginning at Years 9 and 10. Conversion of existing assessments to this format has increased workloads in the short-term but teachers appreciate the chance to experiment in a "low-stakes" context. Some internal achievement standards are being assessed at Year 10 to ease curriculum coverage pressures at Year 11.

Learning programmes are typically organised around the discipline-specific divisions of the various achievement standards, which has led to perceptions that the curriculum is segmented. Some schools are beginning to offer innovative courses that combine science achievement standards in new ways — for example, to create “physical sciences”, “biological sciences”, or “environmental science” courses.

Classes are likely to be streamed, especially in mathematics, with students of differing ability levels offered courses that combine different combinations of achievement and unit standards. In these schools, students in science and mathematics classes with different types of learning needs may be experiencing the NCEA differently:

- “less able” students are likely to have a higher proportion of their course internally assessed and to be assessed with unit standards rather than achievement standards;
- “more able” students are more likely to be encouraged to try for merit or excellence level awards, especially for externally assessed standards; and
- students in some schools are well supported to try for reassessment but those in other schools get “one shot” at internally assessed achievement standards.

Changing balances in classroom programmes

The patterns of ranking of priorities and practices reported by the study teachers show a mix of changes pre- and post-NCEA implementation at level 1.

Both mathematics and science teachers say they are now spending more time on ensuring that assessment incorporates a range of levels and/or types of thinking. This change is directly linked to assessment requirements for demonstrating merit and excellence for achievement standards and teachers are actively looking for ways to develop this change further. Both mathematics and science teachers also say they are now spending more time teaching for understanding rather than for content coverage, but they are worried about “narrowing” the curriculum.

Mathematics teachers say they are now using fewer open-ended investigative tasks and less higher- order tasks than pre-NCEA. This accords with their assertions that it is not easy to fit rich mathematical tasks into their programmes now. However it may be that the difficulty of fitting such tasks into a mathematics programme has been exacerbated rather than arising as a new issue.

Science teachers say they are now using fewer strategies that help students to clarify their own ideas. Again, it may be that the NCEA implementation has exacerbated an existing tension in competing classroom priorities, rather than arising as a new issue.

The weight of responsibility that teachers feel when their students are assessed for qualifications may mitigate against some changes teachers would otherwise like to make in their classroom practices. For some teachers, this tension is exacerbated by internal assessment for qualifications because an explicit focus on preparation for such assessments is now spread through the year rather than being focused on a one-off end-of-year event. On the positive side, more time and attention are now being given to the practical course components that are internally assessed.

Best practice lessons/opportunities for professional development

In all 18 subject departments there is a strong focus on working collegially to implement the NCEA. Most schools have well-established internal moderation policies and practices. Some teachers have welcomed the strong focus on student learning and achievement, and say they have used this to rethink aspects of their teaching practice.

The achievement standards have made teachers more aware of differences in levels of student achievement for a range of aspects of learning. They are more focused on helping students develop the skills needed to demonstrate learning for merit and excellence, although there are some tensions between this teaching goal and current practice including:

- learning for higher-level achievement may be restricted to more able students, with “average” students expected to concentrate on gaining achievement grades;
- some teachers have responded to the likelihood that unfamiliar contexts will be used for excellence components of examination questions by trying to cover more content;
- formative assessment is often strongly associated with holding trial runs for summative assessment rather than with extending and deepening learning;
- teachers are reluctant to promote the use of self-regulated learning strategies except where these simply monitor students’ self-management of their overall progress; and
- some strategies that could help develop students’ thinking skills are being used less rather than more. There may have also been a decline in the use of meaningful contexts for learning.

These contradictory changes appear to be related to the imperative that teachers feel to “cover” the curriculum, and the time that preparation for new and unfamiliar types of assessments is taking from the overall learning programme. Both mathematics and science teachers would like to provide more stimulus materials that get students discussing ideas. There are opportunities to align professional discussions with the review of curriculum coverage in the curriculum stocktake. The provision of classroom-ready exemplars could support teachers to more closely match their practice to their priorities — especially if such tasks can be linked to the desire to help students demonstrate merit and excellence in their NCEA assessments.

These teachers favour the use of pre- and post- tests to monitor student learning, as did the mathematics and science teachers who responded to the national survey associated with the curriculum stocktake. The provision of exemplars of good formative assessment practice could help teachers to better involve students in the ongoing monitoring of their learning, and inform their next learning steps. Since students are beginning to make unilateral decisions about which assessments for qualifications they will undertake, handing them more such responsibility will be timely and may help with the associated motivation challenges. It could also help with the time pressures teachers feel when internal assessments are carried out under one-off examination conditions. However, students need good advice and support to choose appropriate pathways.

There is a strong focus on the development of literacy skills in both curriculum areas. While the secondary schools literacy initiative has contributed to this in at least some of the case study schools, teachers' awareness of literacy challenges appears to have been raised by the types of examination questions now being used in external standards-based assessments.

Policy and further research implications

This research can be used to inform the Ministry's ongoing work in a number of ways. These include:

- potential to monitor ongoing NCEA-related changes if the research is repeated in several years' time, and to use the teacher self-reflection instrument developed in this research for a larger-scale survey of teachers in these or other subjects;
- using the insights into the nature and range of interacting factors that impact on teachers' classroom practices when their students face high-stakes assessment for qualifications;
- informing the focus of any ongoing professional development initiatives that explicitly support the NCEA implementation, including the development of strategies that encourage teachers to revise their expectations of students perceived to be low- or under- achievers;
- auditing the work being carried out in other professional development to identify opportunities to create synergies that will enhance the likelihood of changes in classroom practice taking place;
- informing principals about such opportunities so that they can also make matches to any school-wide professional development that may be underway or planned;
- informing the revision of the suites of achievement standards already available for level 1 mathematics and science, and providing a basis for the discussion about the possible creation of new achievement standards; and
- aligning these findings with ongoing curriculum stocktake work, to encourage professional dialogue about the range of possible purposes for learning mathematics and science, and using these insights to develop new conversations — in addition to achieving success in assessment for qualifications — for teachers to draw on when motivating students to learn.

Section One

Introduction

This research identifies and discusses recent changes in the teaching of mathematics and science in relation to the introduction of the National Certificate in Educational Achievement (NCEA). The project has been carried out by two researchers. One of us (Hipkins) has a background in science education, the other (Neill), a background in mathematics education.

First introduced at level 1 in the 2002 year, assessment for the NCEA qualification is standards-based. Previously most[1] level 1 students sat a norm-referenced end-of-year external examination (School Certificate) that had some internally assessed components in some subjects, but not in either mathematics or science. There are now opportunities for students to gain credits towards their NCEA through internally assessed course components in all subjects, including mathematics and science. The initial research questions were directed toward the identification of changes in classroom practice that may have been triggered by this new opportunity to internally assess some aspects of the Year 11 mathematics and science courses.

The research questions

Three key questions underpinned the exploratory research reported here:

1. As a result of the introduction of the internally assessed achievement standards, are there identifiable changes in the content, structure, and balance within programmes for Maths and Sciences?
2. Are there identifiable changes in teaching and learning styles used within Maths and Science programmes that support the development of practical skills, or that allow teachers to address students' attitudes and values relevant to the subject area?
3. What case study/best practice lessons can be drawn from 2002/2003 practice in Maths and Science programmes for NCEA level 1?

[1] The exceptions occurred in schools that offered modular courses which were fully internally assessed with students' achievement moderated against a "reference test".

These questions were further developed as the research proceeded. It quickly became evident that the NCEA initiative has the potential to change existing balances in a range of counterpoised teaching and assessment practices, and that the actual changes that have taken place are influenced by teachers' values and curriculum beliefs. We report on all the shifts in balances we identified, including the balance between internally and externally assessed course components (question 1).

A note about the title: *Shifting Balances*

Teachers' classroom practice is complex. As we worked through the methodological issues described in Section Two we identified a range of aspects of classroom practice where one way of working or set of emphases could be balanced against another way of working/set of emphases. We anticipated that any of these sets of balances in classroom practice might potentially shift during the implementation of an initiative such as the NCEA. The list of classroom practices that we identified as potentially needing to be balanced against each other during classroom teaching in the senior secondary school is:

- time devoted to learning balanced against time devoted to assessment;
- use of internal assessment balanced against use of external assessment when assessing for qualifications;
- time devoted to developing new "content" knowledge balanced against time devoted to the development of new skills and/or the exploration of attitudes and values;
- a direct (acontextual) focus on concepts/facts/skills balanced against teaching that embeds learning in contexts of relevance to students' lives and interests;
- tool/methodology acquisition by direct "skill and drill" balanced against acquisition via open problem solving/investigations;
- participation in teacher-directed learning activities in which the teachers' ideas take precedence balanced against participation in activities that are student-led, in which students determine the pace and sequence of learning, or actively contribute their ideas;
- time when students learn as individuals balanced against time when they participate in group learning activities; and
- a focus on the cognitive/conceptual aspects of learning balanced against a focus on the metacognitive — that is students' thinking about their thinking and learning.

We foresaw that a shift of balance in one aspect might be reinforced by a related shift in another aspect — or it might equally well be effectively cancelled out by a compensating shift in another factor. Section Four reports our findings with respect to shifts in the balances of the alternatives outlined for these eight aspects of classroom practices.

The methodology we used to answer the research questions is described in Section Two, after some background on the changes brought by the NCEA, and some new material on teachers' views about the NCEA.

Background to the research

To be awarded a level one National Certificate in Educational Achievement students must gain a total of 80 credits from the subjects that they study. These must include 8 credits that demonstrate their literacy skills (i.e., credits from their English courses) and 8 credits from their mathematics courses that demonstrate their numeracy skills. The remaining 64 credits may come from any other courses and subjects for which credits are potentially available.

A National Qualifications Framework (NQF) underpins the NCEA reforms. This framework is intended to organise all credits achieved at a particular level so that they can be credited for one of the many national certificates available. For most school students it is anticipated that the credits they gain will count towards an NCEA award, although schools may also offer other certificates such as the National Certificate in Employment Skills (NCES).

Specifying standards: Achievement and unit standards

There are two types of standards in use for assessment of learning — achievement standards and unit standards. The credits gained from both contribute equally to the total of 80 needed to gain a level one NCEA, although there are some differences between the two types of standards:

- **Achievement standards** have been developed for all "conventional" Year 11 secondary school subjects as part of the NCEA initiative. These specify three levels of achievement: achieved, achieved with merit, and achieved with excellence.
- **Unit standards**, which were a fore-runner to achievement standards and have continued to co-exist alongside them, are competency-based, specifying the standard at a pass/fail level only.

While both types of standards can contribute credits, most "academic" courses that lead to an NCEA award are predominantly if not exclusively assessed with achievement standards. A wide range of unit standards is used to assess "alternative" courses that may lead to another national certificate — for example, an NCES[2] award. In mathematics a number of schools mix achievement and unit standards for "average" students, but this practice appears to be less common in science *(see* Hipkins, Vaughan, Beals, and Ferral, forthcoming).

The full suites of achievement standards available at level one in each subject are summarised below. Internally assessed standards are indicated as (I) and standards assessed in an end-of-year external examination are indicated as (E). These suites of standards are registered and maintained by the New Zealand Qualifications Authority (NZQA).

---

[2] National Certificate in Employment Skills.

Level 1 mathematics standards

There are nine mathematics achievement standards offering 24 credits if all are used to assess a course of learning. Six are examined externally and three are assessed internally.

1.1 (E) Use straightforward algebraic methods and solve equations (3 credits)

1.2 (E) Sketch and interpret linear or quadratic graphs (3 credits)

1.3 (I) Solve problems involving measurement of everyday objects (4 credits)

1.4 (I) Use geometric techniques to produce a pattern or object (2 credits)

1.5 (I) Use straightforward statistical methods to explore data (3 credits)

1.6 (E) Calculate relative frequencies and theoretical probabilities (2 credits)

1.7 (E) Solve straightforward number problems in context (3 credits)

1.8 (E) Solve right-angled triangle problems (2 credits)

1.9 (E) Use geometric reasoning to solve problems (2 credits)

There are many unit standards available to supplement or replace these achievement standards, as teachers choose. In a related piece of research we have found that the most popular mathematics unit standards in six case study schools (not the schools in the research reported here) have a decidedly practical flavour. They focus on various aspects of number and computation, using money, reading tables and graphs, and on the use of statistics *(see* Hipkins et al., forthcoming).

Level 1 standards in science

There are seven science achievement standards at Year 11, offering a total of 26 possible credits where the full suite is offered. Five are examined externally and two are assessed internally.

1.1 (I) Carry out a practical science investigation with direction (4 credits)

1.2 (I) Research, with direction, how science and technology are related (2 credits)

1.3 (E) Describe uses and effects of microorganisms and the transfer of genetic information (5 credits)

1.4 (E) Describe properties and reactions of groups of related substances (5 credits)

1.5 (E) Describe rocks and minerals (3 credits)

1.6 (E) Demonstrate an understanding of physical systems (5 credits)

1.7 (E) Describe spatial relationships in astronomy and their effects on space exploration (2 credits)

Potentially, science teachers have a great deal of freedom to construct quite different courses to meet the learning needs of different groups of students. At the time the achievement standards were being developed by the Ministry of Education's Qualifications Development Group (QDG), the decision was taken to continue the existing situation where levels 6–8 of *Science in the New Zealand Curriculum* co-exist with levels 1–3 of *Biology/Chemistry/Physics in the New Zealand Curriculum* documents. Thus, full suites of achievement standards were developed for all three of these subjects separately, in addition to the science achievement standards related to each of these three discipline areas.

In science, students are able to have their investigative skills internally assessed in two different contexts — if their teachers choose to offer them the opportunity. The relevant achievement standard (1.1) may be assessed for any two of science and/or biology and/or chemistry and/or physics. The same can be done for the research standard (1.2). Potentially then, although there are only two types of internally assessed standards, a course that offers 12 internally assessed credits from achievement standards could be designed (4+4+2+2 credits). Teachers may choose to have the biology aspects of the Year 11 science course externally assessed via the relevant science achievement standards, or via the most closely related biology achievement standard. The advantage of the latter choice may be that external assessments are spread over two examinations (Biology and Science), giving students more time to answer each set of questions than would be possible if all their questions were in one three-hour examination. Potentially the same can be done for science/chemistry and science/physics combinations.

The *Learning Curves* study shows that some teachers are mixing and matching achievement standards from the various science disciplines to create science courses to meet their students' learning needs (Hipkins et al., forthcoming). Three schools in the research reported here have created science courses with a specific type of focus (e.g., environmental science, or physical science) by this means.

There are also many science-related unit standards registered on the NQF. One suite of these was developed after the achievement standards for the explicit purpose of allowing NCEA credits to be gained from the alternate *Certificate in Science* course developed by the NZASE[3] for Year 11 students who cannot cope with a full academic course in science. Others predate the achievement standards, as in mathematics and English, and still others are related to the *New Zealand Certificate in Science*, which may also be offered as a polytechnic course. All unit standards are internally assessed.

---

[3] New Zealand Association of Science Educators.

Support for the reforms – data from the *NZCER 2003 National Survey*

As for any significant change in education, the active support of teachers is an important aspect of “making things work”. Many New Zealand teachers resisted the introduction of unit standards and some did not begin to engage with issues of standards-based assessment until compelled to do so with the arrival of the NCEA qualification. Others actively contributed to the development of the unit standards and also worked on the “expert panels” that developed the achievement standards, putting themselves in a position to feel more confident of the implementation requirements for the NCEA when the time came. The introduction of the NCEA has been supported by an ongoing series of professional development opportunities and many of the teachers whose work is documented in this report have also been active leaders for their colleagues in that learning process.

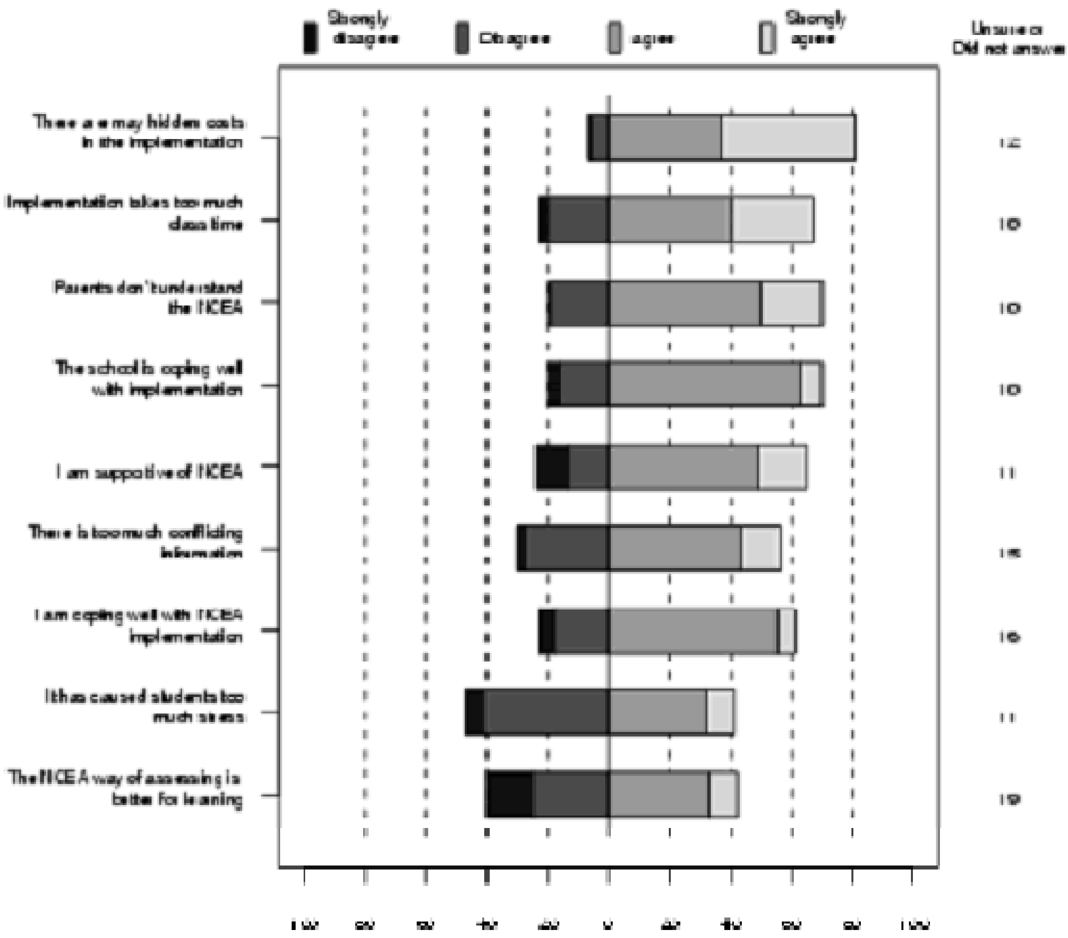
Data from the recently completed *NZCER National Survey* provides an overview of teacher views of the implementation of the NCEA. In mid–2003 NZCER conducted a national survey of secondary schools. Full responses from this survey will be published in early 2004. The survey covered a wide range of aspects of current educational practices and issues. Questionnaires tailored to each group were given to teachers, principals, trustees, and in some schools, parents.

The responses discussed here were made by 744 secondary school teachers from across a range of curriculum areas. Amongst the many topics canvassed, the survey investigated teachers’ views of the NCEA. The teacher questionnaire included nine statements about the NCEA and teachers were asked to respond using a four-point scale: strongly agree; agree; disagree; or strongly disagree.

Patterns of responses

Figure 1 shows patterns of responses to the nine statements. They are shown in the order in which they appeared on the survey form. Most teachers think that there have been hidden costs to the implementation of NCEA and many of them think that implementation takes too much class time. Their views are more divided on whether the NCEA has created too much student stress and on whether the NCEA way of assessing is better for learning. Overall, teachers were more likely to agree that the school is coping with the NCEA implementation than they were to agree that they are personally coping.

Figure 1 **Teachers' views about the NCEA**



Response patterns of mathematics and science teachers

Mathematics and science teachers differed from the other teachers in their views about the NCEA.[4] Somewhat fewer mathematics and science teachers agreed or strongly agreed with the statement "I am supportive of the NCEA" (60 percent compared with 68 percent of other teachers). Mathematics and science teachers were more than twice as likely to strongly disagree with the statement.

4 The tests of significant difference between the groups included only teachers who expressed an opinion one way or other. Teachers who taught mathematics or science along with a range of other subjects were included with "Other teachers", as they were not specialising in either discipline.

Only 33 percent of mathematics and science teachers agreed or strongly agreed with the statement "The NCEA way of assessing is better for learning" whereas 46 percent of other teachers agreed or strongly agreed. Again, mathematics and science teachers were twice as likely to strongly disagree with the statement compared with other teachers.

This means that these two groups of teachers do not think NCEA supports learning as much as do teachers in other areas of the curriculum. We cannot say if this is primarily an attitudinal difference, if it is reflective of some difference in the nature of learning in mathematics and of science, or related to an aspect of assessment such as the mix of internal and external assessment in subject courses.[5]

Mathematics teachers were also less likely to agree or strongly agree with the statement that "The NCEA has caused students too much stress" than are all other teachers. A third of mathematics teachers (33 percent) responded this way compared with other teachers (43 percent). Science teachers' responses were not statistically different from those of other teachers.

Relationship with school's decile rating

Response patterns to three of the nine statements shown in Figure 1 showed a relationship with the decile of the school. For this analysis the deciles were grouped into three clusters: low decile (1–3); medium (4–6); and high decile (7–10). Only schools where the decile was known and where the teacher offered an opinion one way or the other were included in the tests for significance.

Teachers in low decile schools showed a higher-level of support for the NCEA, with 75 percent percent of this group either agreeing or strongly agreeing with the statement "I am supportive of NCEA", compared with 68 percent of teachers in medium decile, and 60 percent in high decile schools. Teachers in low decile schools were also more supportive of the statement "The NCEA way of assessing is better for learning" with 56 percent agreeing or strongly agreeing with this statement compared with 40 percent of teachers in medium decile schools and 38 percent in high decile schools.[6]

The more favourable views of teachers in low decile schools may reflect the higher ratios of students with lower achievement levels in these schools. As Section Four reports, teachers see NCEA as being well suited to lower-achieving students. This year's *Learning Curves* report links such teacher perceptions to the availability of NCEA credits for unit standards-based courses, and also to the possibility of assessing all or most of the students' learning using internal assessments (Hipkins et al., forthcoming).

---

[5] Since different subjects had differing proportions of internal and external assessment pre-NCEA and continue to do so now, no generalisations are possible.

[6] chi-square = 12.16 with 2 d.f. (p. = 0.0023).

Fewer teachers in low decile schools agreed or strongly agreed with the statement “The NCEA has caused students too much stress”: 33 percent of these teachers, compared with 40 percent of medium decile teachers and 46 percent of high decile teachers.[7] Teachers’ perceptions of student stress also increase with the decile rating of the school. During the interviews in our *Shifting Balances* schools we noted that some teachers in high decile schools are anxious because they perceive that they now have to teach more to ensure they have successfully anticipated “excellence” level questions that might be asked. This could be a contributor to the higher stress levels teachers in high decile schools perceive.

Modelling decile and teacher discipline together

Both decile and the teachers’ discipline (mathematics/science or other) show a significant relationship with some of the attitudes towards NCEA when tested individually. To investigate the interrelationship between decile and discipline and to test if each of these effects remains significant in the presence of the other, they need to be modelled together. To do this a log-linear model that included the decile effect (“low”, “medium” or “high”), the discipline of the teacher (“maths or science” or “other”) was fitted to each of three variables where either decile had an effect or where mathematics and science teachers differed from other teachers. These three variables were “I am supportive of the NCEA”, “The NCEA has caused students too much stress”, and “The NCEA way of assessing is better for learning”.

For each of the three NCEA response questions, both decile and discipline had a statistically significant effect (as indicated above) but there was no statistically significant interaction between them — that is, mathematics and science teachers’ views are not affected by the socio-economic decile of the school they are teaching in.

Other questions related to NCEA

The *National Survey* included four other questions where the NCEA was mentioned in some of the response options provided. For each of these four questions, cross-tabulations were performed to see if mathematics and science teachers differed in their responses from teachers of other disciplines. In none of the four cases was there a statistically significant effect, indicating mathematics and science teachers have similar patterns of responses to other teachers on these questions. The relevant responses for these questions were as follows:

- *“Do you feel there are any barriers to your making changes to the curriculum you teach?”*

---

[7] chi-square = 9.51 with 2 d.f. (p. = 0.0086).

Of the thirteen options provided, the most frequently selected response was “Lack of time” (50 percent). The second most frequently selected response was “The time taken for NCEA implementation” (40 percent).

“*Where have the most useful ideas for your teaching programme come from in the last 2 years?*”
Of the thirteen options provided, the most frequently selected responses were “Other teachers in the school” (56 percent) and “One-off courses/conferences/professional development” (52.0 percent). The “NCEA” was the sixth most frequently selected response (29 percent).

“*What do you think are the three major issues confronting your school (if any)?*”
Of the twenty categories listed, the most frequently selected was “Funding” (38 percent) followed by “Student behaviour/discipline” (32 percent). “NCEA implementation” ranked third (24 percent).

“*What do you feel are your three main achievements as a teacher in the last 3 years?*”
Of the nine options provided, “Implementation of NCEA” ranked first equal (46 percent) with “Positive/improved learning environment” (46 percent). These were closely followed by the option “Increased my own knowledge/skills” (45 percent). “Improvements in student achievement” was ranked eighth (17 percent).

When considered together, these findings from the *National Survey* present a mixed pattern of concerns and feelings of personal achievement. Across the curriculum teachers are feeling the pressure of the time needed to implement the NCEA and see this as a barrier to making curriculum-related changes in their teaching. But they also view their implementation of the NCEA as a main achievement in their teaching and some teachers perceive that it has given them useful ideas for their teaching. Perhaps reflecting the ambivalence teachers express about whether the NCEA is better for students’ learning, they do not feel as successful in improving student achievement as they do in actually implementing the NCEA.

Section Two

Methodology

This section describes the research design and explains how specific challenges and constraints were overcome as they were encountered. Data was gathered using a case study approach and was exploratory.

Preliminary scoping of the project

In mid-2003 groups of facilitators appointed to implement NCEA training for the teachers met for their own preliminary training and discussions. Most of these facilitators are experienced secondary school teachers, and some are school support advisers. We attended the one-day meetings in Auckland and in Christchurch.

Christchurch and Auckland/Northland were chosen for the potential to locate the case studies in a mix of urban and small town/rural areas. The regions provide contrasts of population composition, and reflect North Island/South Island differences. No *Learning Curves*[8] schools are located in either of these regions.

While at the planning meetings we interacted with small groups of mathematics and science facilitators, listening to their discussions as they planned the professional development they were about to deliver, and talking with them during the breaks. The facilitators were asked to share their impressions of the extent of innovative level 1 practice, and their views on enabling factors and barriers to change in level 1 mathematics and science classrooms. This discussion was used to make a preliminary identification of likely case study schools.

---

[8] A related study of another aspect of NCEA implementation also being carried out by NZCER. In the event, when the geographic area was extended *(see* below) one teacher from one *Learning Curves* school was invited to join the sample, on the recommendation of the regional facilitator.

As we listened to the flow of ideas we identified likely teachers/schools with whom to work and we pursued the topic of changes in classroom practice with the facilitators during their breaks. This feedback allowed people who would not be directly involved in the case studies to tell us about ideas they have tried and/or issues they have encountered when making changes based on the introduction of the internally assessed achievement standards. It provided a wide perspective on the extent of innovative practice at this stage of the NCEA implementation process and informed the development of the interview protocols.

Many of the participants at the regional training days thought the research was somewhat premature. They had the impression that many teachers were temporarily neglecting some aspects of their classroom teaching while learning how to implement standards-based assessment. This view was shared by some of the principals whom we subsequently approached to invite participation in this research. This view also accords with the findings of the *NZCER National Survey* (which were not available at the time this scoping exercise was carried out). The NCEA implementation is taking a great deal of time and this can act as a barrier to making other sorts of changes in the classroom. Teachers in the *Learning Curves* research similarly report heavy workloads and time pressures related to the NCEA implementation that may be preventing them from thinking more creatively about the new possibilities the NCEA is opening up in the area of subject choice (Hipkins et al., forthcoming).

Notwithstanding this somewhat discouraging tone, facilitators were able to come up with interesting and/or promising aspects of NCEA implementation in level 1 science and mathematics classes when prompted to think about all aspects of classroom practice and curriculum implementation. We realised at this point that the focus of the research needed to be somewhat widened beyond the initial three questions if it was to be productive.

The case study schools

With the help of the facilitators and the regional science and mathematics advisers, selected teachers of each subject were invited to participate. While it was initially intended that these teachers be selected for the potential to demonstrate innovative pedagogy in mathematics and science teaching with respect to level one internally assessed standards, the feedback we had received prompted us to widen our sights. We invited participation from teachers of the two subjects who were known to be coping with any aspect of NCEA implementation in interesting ways.

We had anticipated that choice would be possible within each region, allowing selection to take account of diversity of school types and to spread the choice of science teachers across the specialist science disciplines. In the event we struggled to elicit participants. After an initial informal approach to the relevant teacher, we made a formal approach to the principals of a range of potential case study schools. Some teachers and some principals turned us down, giving a variety of reasons that were not all NCEA-related. Eventually we were able to work with nine mathematics and nine science teachers, including several from regions outside the two initially targeted.

Mathematics study schools

The nine “mathematics” schools have a range of decile ratings with a slight predominance of mid-range deciles. (One of the two low decile schools that we had initially identified was not able to take part in the research.) There are two all boys’ schools and two all girls’ schools. The remaining five are co-educational. The schools range in size from 700 to 1450 students. One is located in a provincial town, three are in provincial cities, and five are in major cities. Of this latter group, one is a central city school and the other four are suburban. All nine are state schools, and all enrol students from Year 9 to Year 13.

School A is located in a provincial town. It is a decile 7 co-educational school with a roll of about 1150. It draws students from within the town and the surrounding area. The student population is largely European with a small number of Mäori but very few Pasifika students. In this school, both the HOD and co-ordinator of Year 11 mathematics took part in the research.

School B is an all boys’ school. It has a decile rating of 6 and has a roll of about 1300. Its students are predominantly European, with some Pasifika students, and fewer Mäori students. Some students are attracted from the area to a number of traditional schools near the city centre. The HOD was interviewed in this school.

School C is a co-educational school located in the suburbs of a major city. It has a decile rating of 8 and with a roll of 1450 students, is the biggest school in the mathematics sample. While it has a number of Mäori and Pasifika students, it draws students from a largely European catchment with a number of Asian students. The HOD was interviewed in this school.

School D, with a roll of 900 students, is the second smallest of the schools in the mathematics study. It has nearly 100 international students included in this total and the other students are mainly of European origin. It is a co-educational school located in the suburbs of a major city, and has a decile rating of 7. The HOD was interviewed in this school.

School E is a boys’ school located in a provincial city. It has a roll of 1100 students, mainly of European origin. The school has boarding facilities used by both New Zealand and overseas students and has a decile rating of 8.

School F is located in the suburbs of a major city. With a roll of 700 it is the smallest school in the mathematics study group. Most of its students come from Mäori or Pasifika backgrounds. It is co-educational and has a decile rating of 1. The HOD was interviewed in this school as well as the person who takes the middle- and upper-streamed mathematics classes. A brief opportunity was given to observe a lesson.

School G is a girls' school located in a central city area. There are 1350 students, from a range of ethnic backgrounds, and the school has a decile rating of 4. The HOD was interviewed in this school with a short conversation with another teacher occurring early in the interview.

School H is a girls' school located in a provincial city. The school roll is about 1200 students, from a range of ethnic backgrounds, and the school has a decile rating of 5. The interview was with a senior teacher as the HOD was absent overseas at the time of the study.

School I is located in the suburbs of a provincial city. The school currently has about 1250 students, from a range of ethnic backgrounds, with a significant proportion of Mäori students. It is a co-educational school with a decile rating of 5. The HOD was interviewed in this school.

In all but one of the mathematics schools, the interview was conducted with the HOD. Two of the schools also involved another teacher in the interviews. In one school this person had responsibility for the Year 11 programmes in mathematics. In the other the teacher was responsible for the classes being assessed by achievement standards, rather than the classes who were primarily assessed using unit standards.

The interviewees were predominantly very experienced teachers. The time spent as a teacher ranged from 11 to 40 years with an average of 25.5 years. The time spent at the current school ranged from 4 to 26 years, with an average of 11.1 years. This means that all those interviewed had experience of both School Certificate and NCEA in their current school.

Science study schools

The nine "science" schools vary widely across the full range of decile ratings. There are three girls' schools but no boys' schools were recommended for inclusion in the sample. The other six schools are co-educational and range in size from 150 to 1450 students. Two are located in provincial towns, one is in a provincial city, one is an area school in a rural location, and five are in major cities. Of this latter group, three are central city schools and the other two are suburban. Seven are state schools and two are private. Six schools enrol students from Year 9 to Year 13. The two private schools and the area school span all year levels from 1–13.

School A is a girls' school located in a central city area. The students have a wide range of ethnic backgrounds. There are 1350 students at the school which has a decile rating of 4. The science teacher in charge of Year 11 programmes was interviewed in this school.

School B is a private girls' school in a central city area. It has a decile rating of 10 and draws its 1300 students from all over the metropolitan area. There are some boarders. In this school we interviewed the HOD and spoke with four other teachers as they came through their shared workroom at various stages of the visit.

School C is also a private girls' school, located in another city. The decile rating is again 10 and the roll is around 600. In this school students may choose between studying for the NCEA and studying for an international baccalaureate award. We interviewed the HOD and a first year teacher, as well as sharing a morning tea conversation with two other science teachers in the shared workroom.

School D is a decile 1 co-educational school in a suburban area of a major city. Students are predominantly from Mäori or Pasifika families. With 450 students, this was the second smallest school in the sample. In this school we interviewed the HOD science and briefly shared time with the “alternative” Year 11 class as they worked on a science research project.

School E is also decile 1 and is an area school in a remote location. The roll spans Years 1–13 and with 150 students this is the smallest school in the overall sample. Most students are Mäori and even those with mixed Mäori/Pasifika backgrounds are seen to relate most strongly to their iwi affiliations. The teacher we interviewed teaches science at all levels in this school.

School F is located in the suburbs of a provincial city. The 1250 students are from a range of ethnic backgrounds, with a significant proportion of Mäori students. It is a co-educational school with a decile rating of 5. The HOD was interviewed in this school.

School G is a co-educational school located in a rural town. It has a decile rating of 3 and a roll of 600 students, many with Mäori backgrounds. In this school we interviewed the HOD and one other teacher. We also interacted with one Year 11 class during a science lesson.

School H is co-educational school located in another rural town. It has a decile rating of 5 and a roll of 600 students. In this school we interviewed the HOD who has taught all three science disciplines at senior level but is currently teaching senior physics.

School I is a co-educational school located in the suburbs of a major city. It has a decile rating of 8 and with a roll of 1450 students, is the biggest school in the science sample. While it has a number of Mäori and Pasifika students, it draws students from a largely European catchment with a number of Asian students. The HOD and one other teacher were interviewed, and we attended part of an after-school meeting of all the science teachers.

In all but one of the science schools, the main interview was conducted with the HOD. At the school where this was not the case, the interviewee was the teacher in charge of Year 11 science in the school. Five of the schools also involved one or more of the other science teachers in the interviews, either formally or informally. This widened the range of views we were able to capture, including perceptions from a Year 1 teacher who has never taught School Certificate classes, and those of an overseas-trained Year 6 teacher who had no experience of assessment in New Zealand schools prior to the NCEA.

As in the mathematics sample, the interviewees were predominantly very experienced teachers. The time spent as a teacher ranged from 3 to “30+” years with an average of 17.6 years. The time spent at the current school ranged from 1 to 18 years, with an average of 8.1 years, somewhat less than for the mathematics teachers. Six of those interviewed had experience of both School Certificate and NCEA in their current school, three did not.

The design of the data-gathering instruments

We have already noted the early finding of a perception that “things are going backwards” in some classes as the NCEA implementation beds in. In the absence of any recent, large-scale studies of actual classroom practice in secondary school mathematics and science teaching in New Zealand, there were no available “base-line” data against which we might measure change in classroom practice.

These twin challenges posed an interesting dilemma when designing the data-gathering instruments for the school case studies. The lack of base-line data meant that we had to rely on teachers’ own perceptions of changes they had made. We anticipated a risk of being seen to distort the data by over-identification of positive changes if we only sought and reported on these. On the other hand, the data could have been as easily distorted the other way if we had invited teachers to focus on factors they perceive to be problematic about the NCEA[9] that could inhibit change or push it in negative directions.

An observation made by one of the key participants in a very large Science in Schools (SIS) professional development initiative being carried out in Victoria, Australia seemed to us to pose a third challenge. Tytler (2003a) observed that, when invited to identify and discuss changes in their classroom practice, teachers tend to say there has been no change unless they can report on some very substantial difference. That is, small incremental changes are typically overlooked.

The design of the self-reflection instrument

With all three types of challenge in mind, we designed a self-reflection instrument that could capture changes in a range of classroom practices. The instrument drew teachers’ attention to the multitude of smaller and larger changes that might potentially have happened, thereby meeting the challenge posed by Tytler.

Drawing on the findings of the Science in Schools (SIS) project we selected 19 succinct descriptors of best practice. These descriptors were developed during the SIS research to illustrate how eight key “SIS Components”[10] of best practice science teaching might play out in classrooms (Tytler, 2003b). We adapted these descriptors where necessary so that they would be applicable in both mathematics and science classrooms.

These eight key “SIS components” of best practice are:

---

[9] We do not mean to suggest that the significant workload issues which we have reported elsewhere (Hipkins et al., forthcoming) should not be taken seriously. The focus here is on changes in the classroom and we did not want workload issues to distort the teachers’ perceptions of these.

[10] http://www.scienceinschools.org

1. Students are actively encouraged to engage with ideas and evidence.
2. Students are challenged to develop meaningful understandings.
3. Science is linked with students' lives and interests.
4. Students' individual learning needs and preferences are catered for.
5. Assessment is embedded within the science learning strategy.
6. The nature of science is represented in its different aspects.
7. The classroom is linked with the broader community.
8. Learning technologies are exploited for their learning potentialities.

The first of these components, with its associated descriptors, is provided as an example in Figure 2. The 19 modified descriptors used for the self-reflection sheet are shown in Table 1.

Figure 2 **An example of a "SIS component" with its key indicators**

1. Students are actively encouraged to engage with ideas and evidence.

1.1 Students are encouraged and supported to express their ideas, and question evidence.

1.2 Student input (questions, ideas, and expressions of interest) influences the course of lessons.

1.3 Students are encouraged and supported to take some responsibility for the design, conduct, and analysis of science investigations.

Table 1 **The descriptors used for the teacher self-reflection sheet**

Number assigned	Descriptor as modified from SIS research
1	Providing stimulus materials that challenge students' ideas and that encourage discussion, speculation, and ongoing exploration by groups of students working together.
2	Moving away from a strong focus on content "coverage". Moving towards a focus on ensuring understanding and meaningful learning of a reduced amount of content.
3	Encouraging students to make their own decisions in practical investigations concerning hypotheses to be explored, experimental design, measurement and recording techniques, analysis and interpretation.
4	Including frequent open-ended investigations or short-term open explorations.
5	Ensuring higher-order tasks involving the generation, application, analysis, and synthesis of ideas, are well represented.
6	Encouraging students to actively clarify their own ideas, and to think about their learning processes. (E.g., by using concept mapping, model making, learning journals, exploration of alternative strategies, etc.)
7	Using students' personal interests (sports, hobbies) and social/ethical concerns as the context of mathematics or science topics and involving them in making choices about their learning.
8	Setting a variety of types of tasks during each unit.
9	Using a variety of methods to assess student understandings, at various points in a unit (e.g., open-ended questioning, checklists, project work, problems, practical reports, role plays).
10	Involving students in decision making about what should be assessed, and when and how assessment should be carried out.
11	Ensuring assessment incorporates a range of levels and/or types of thinking.
12	Probing student understandings and perspectives early in a learning sequence to help plan subsequent lessons.
13	Ensuring students have ongoing feedback which indicates their strengths and weaknesses and their next learning steps.
14	Using a variety of types of experiment to exemplify scientific/mathematical methods and principles, including measurement techniques, variable control, survey work, modelling, and open exploratory designs.
15	Including discussion of mathematical/scientific evidence contributing to contemporary science/mathematics-related public issues that are of interest/importance to students.
16	Discussing and developing understanding of language conventions of science/mathematics.
17	Basing sequences of work around local community projects, such as environmental maintenance or studies of local industries.
18	Using learning technologies to support quality learning behaviours such as exploration, conjecture, or collaboration (e.g., spreadsheets, Internet, data loggers, graphics calculators).
19	Exploring different values and perspectives that students bring to their science/mathematics learning.

The "SIS components" are *principles* for best practice. Their descriptors mix and match various aspects of classroom practice related to those principles. For our purposes, we rearranged the descriptors so that we could "un-mix" certain aspects of learning from aspects of assessment, for example. This provided six distinct groupings of the selected descriptors, as shown in Table 2. The descriptors for each theme were presented in a random order on the reflection sheet.

To provide a perspective against which to weigh teachers' views of the actual changes that had taken place, we first asked them to assign a *priority* to each practice using a five-point scale: very high; high; moderate; low; very low. This gave us an insight into the types of changes that the responding teachers would be most likely to value.

Next we captured their perceptions of actual changes. To avoid leading their responses we used two scales. One recorded teachers' perceptions of how often they carried out each of the described practices pre-NCEA, the second how often they did these things now. We used a four-point scale: hardly ever/never; occasionally; often; all/most of the time. The differences between the two sets of responses provided us with the data on actual changes — both positive and negative — that we report in Section Three. A copy of the full self-reflection instrument is provided as Appendix 1.

We felt it was important that teachers had time to ponder their responses rather than making judgments on the spot. Accordingly, the self-reflection sheets were sent to teachers ahead of the arranged interview time, as were the additional interview questions we intended to ask *(see* below).

Table 2 **Themes addressed by self-reflection descriptors**

Theme	Sub-themes	Descriptor numbers
Assessment	Formative assessment	9, 12, and 13
	Variety of assessment tasks	11
	Student input into assessment decisions	10
Rich tasks	Types of rich tasks	1, 4, 5
	Variety in tasks	8
Practical work	Own investigations	3
	Nature of science/maths	14
Learning	For understanding (vs "coverage")	2
	Metacognitive skills	
	Use of language	6
		16
Use of contexts for learning	Personal interests and values, local and public issues	7, 15, 17, 19
Use of new technologies		18

Setting our findings in context

Teachers' classroom practices are embedded in the wider contexts of their departmental team and its practices, their school and its policies and practices, and their community with its particular characteristics and expectations. We felt it was important to embed the findings from the self-reflection sheet within the range and variability of teaching contexts. We designed an interview schedule with 16 open-ended questions to cover at the outset of each interview *(see* Appendix 2). We then worked through the self-reflection sheet, capturing teachers' comments about the ratings that they had made.

Interviews took between 2 and 4 hours in total. Longer visits typically included conversations with other teachers and/or brief classroom visits. Interviews were audio-taped to assist us in reviewing our notes. Teachers were assured that the tapes would not be transcribed.

The time we spent in each school varied. The shortest visits took around 2 hours. Some visits took 4 hours. Schools were provided with money to pay for teacher release time, and the teachers were also offered some financial compensation for the time they spent preparing for the visit. Teachers frequently commented that they had found the self-reflection sheet thought provoking. Some had given copies to other teachers in their subject area. Most teachers had prepared notes before the interview and some said they had returned to the research instruments on a number of occasions over several days ahead of the visit. All of the teachers seemed to enjoy the opportunity to discuss professional issues related to the NCEA implementation regardless of the tenor of their actual views and feelings about it.

Exemplars of student work

To address the "best practice" aspect of the initial research questions, we invited teachers to preselect interesting examples of students' work that illustrated changes they had made in their classroom practice. The intention was to discuss these at the time of the interview and to select some for inclusion in the final report.

In the event, teachers were unsure what would interest us or thought they had nothing particularly innovative to share. Most had not gathered actual student work ahead of the interview. They were more likely to give us examples of modified worksheets and/or test questions. A few teachers did show us examples of innovative tasks that they had already generously shared with other teachers in their area. Inclusion of these would have made them identifiable to some readers of this report and we had promised confidentiality so we have not included these.

In view of these responses, this report shares less of this type of evidence than we had anticipated being able to include. Examples we have been able to include are discussed in Section Five.

Section Three

The nature and extent of teachers' reported changes in their classroom practice

This section reports on patterns of teachers' responses to the self-reflection sheets that were mailed to them ahead of the scheduled interview *(see* Section Two). Teachers' perceptions of the value that should be attached to the various classroom practices identified on the reflection sheet are compared with their perceptions of actual changes in classroom practice. The section begins with a short discussion of the collation and analysis of the teachers' responses.

In some schools more than one teacher had completed a reflection sheet. In total we received nine completed sheets from mathematics teachers and 13 completed sheets from science teachers. Two science teachers had completed the priority and frequency of current practice sections but had no experience of preparing Year 11 students for School Certificate examinations. Responses from these teachers could not be used when calculating perceptions of the extent of classroom changes.

When considering the patterns reported, it should be noted that the findings are based on a sample of just nine mathematics teachers and thirteen science teachers and cannot be generalised. With this small sample size, the average scores and variance are susceptible to being affected by just one or two respondents. For this reason, most of the analysis is done on ranked scores and the statistical tests done are non-parametric, which are robust, and make no assumptions about normality.

Quantifying responses to the provided scales

Teachers were asked to assign a priority to each descriptor of teaching practices using a five-point scale from "very low" to "very high". Some teachers ticked on the lines between boxes rather than in the boxes provided. We took account of these responses by collating them on a nine-point scale rather than the five we had initially designed:

1 = very low; 2 = on the line between 1/3; 3 = low; 4 = on the line between 3/5; 5 = moderate; 6 = on the line between 5/7; 7 = high; 8 = on the line between 7/9; 9 = very high.

Responses to frequency of classroom practice were similarly collated using a seven-point scale:

1 = hardly ever/never; 2 = on the line between 1/3; 3 = occasionally; 4 = on the line between 3/5; 5 = often; 6 = on the line between 5/7; 7 = all/most of the time.

Once all responses had been collated numerically, the scores for each descriptor were averaged. The average scores were then ranked from 1 for the practice rated as the highest priority to 19 for the practice rated the lowest priority. The same process was followed to rank teachers' perceptions of the frequency of each practice before and after the implementation of the NCEA. The results are summarised in Table 3 for mathematics and Table 4 for science.

The mathematics teachers' responses

Priorities

Table 3 shows the average score (S_p) and rank (R_p) for the priority assigned to each descriptor, as well as the perceived average score and rank of practice pre-NCEA (S_b/R_b) and after-NCEA (S_a/R_a) implementation (1 = highest ranking, 19 = lowest ranking). The final column gives the difference in average score between current practice and pre-NCEA practice (S_a – S_b) — that is, it provides a means of quantitatively reporting the actual changes the responding teachers think they have made. Each descriptor is explained in Table 2 of Section 2. The results in Table 3 can be analysed by the themes of assessment, rich tasks, learning, practicals, context, and technology. The overall distribution of rankings across the six themes is summarised visually in Figure 3.

Table 3 **Mathematics teachers' perceptions of priorities and changes in practices**

Descriptor	Average priority (S_p)	Priority rank (R_p)	Av. Pre-NCEA (S_b)	Pre-NCEA rank (R_b)	Av. Post-NCEA (S_a)	Post-NCEA rank (R_a)	Change ($S_a - S_b$)
1	7.50	2	3.00	10=	3.13	11	0.13
2	6.13	12	3.50	8=	4.25	6	0.75
3	6.45	9=	3.50	8=	3.88	8	0.38
4	6.45	9=	3.67	6=	3.22	10	-0.45
5	6.56	7=	3.00	10=	2.67	15	-0.33
6	6.45	9=	2.78	13	3.33	9	0.55
7	5.78	14=	3.00	10=	3.11	12	0.11
8	7.11	4	5.00	1	5.22	2	0.22
9	6.67	6	3.89	5	4.22	7	0.33
10	3.88	19	1.89	18	2.11	18	0.22
11	7.00	5	3.67	6=	4.78	4=	1.11
12	7.56	1	4.89	2	5.45	1	0.56
13	7.23	3	4.56	4	4.89	3	0.33
14	5.78	14=	2.63	14=	2.83	13=	0.20
15	5.11	16	2.57	16	2.57	16	0.00
16	6.56	7=	4.67	3	4.78	4	0.11
17	4.33	18	1.33	19	1.33	19	0.00
18	6.00	13	2.25	17	2.25	17	0.00
19	4.89	17	2.63	14=	2.88	13=	0.25

Figure 3 **Ranking of maths priorities for the 19 descriptors according to themes**

Assessment	12		13		11	9													10
Rich tasks		1		8			5		4										
Learning							16		6			2							
Practicals									3					14					
Contexts														7		15	19	17	
Technology													18						
RANK	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**

Figure 3 shows that the highest priorities in mathematics are given to assessment-related practices. Four of the top six descriptors relate to the assessment theme. Descriptor 12 (formative assessment early in a learning sequence) was accorded the highest priority rank, with descriptor 13 (feedback to students — another plank of good formative assessment) ranked third. Another descriptor related to formative assessment (9 — using a variety of methods to assess students at various points) ranked sixth and "ensuring assessment incorporates a range of levels or types of thinking" was ranked fifth. During the interviews, teachers often linked the latter descriptor to the focus on the levels of achievement now specified in the achievement standards (achieve, merit, excellence). The impact of this association is explored in some detail in Section Four of the report. Interestingly, the only other assessment descriptor (10 - involving students in decision making about what should be assessed) was ranked last.

Questions relating to the theme of rich and varied tasks within the classroom were afforded the next highest set of overall priorities. The second rank went to descriptor 1 (providing stimulus materials that challenge students' ideas). The two other rich task-related descriptors were number 5 (use of higher-order tasks) and number 4 (open-ended investigations) and were ranked somewhat lower at seventh and ninth equal respectively. "Setting a variety of tasks" (descriptor 8) was ranked fourth.

Descriptors related to learning issues made up the next group. "Literacy requirements" (descriptor 16) ranked seventh equal, with metacognition (descriptor 6) ranked ninth equal, and learning for understanding (descriptor 12) ranked twelfth.

Technology (descriptor 18) ranked thirteenth, and is seen as a somewhat separate dimension than the overall learning process.

The two descriptors of practical or experimental tasks ranked next. Descriptor 3 ("encouraging students to make their own decisions in practical investigation") ranked just ninth equal while 14 ("using a variety of experimental tasks" — including measurement and hinting at statistics) rated quite lowly at fourteenth equal. These responses are surprisingly low given the emphasis the internally assessed achievement standards place on the development of practical skills.

As a group, the four descriptors relating to contexts rated lowest. Personal interest (7) ranked fourteenth equal, values students brought with them (19) ranked seventeenth, local issues (17) ranked eighteenth, and public issues (15) ranked sixteenth.

Current practice

Figure 4 shows the distribution of rankings assigned to current practice in the second year of the implementation of the NCEA at level 1. With the exception of eliciting students' own input to assessment decisions, assessment-related activities are perceived to be used more often overall than the practices related to the other themes. It does appear that teachers perceive that aspects of assessment dominate their classroom practice.

The biggest change from the ranking assigned to priorities is for descriptors related to the theme of rich tasks. While they still rate highly, these have moved down to now rank third overall along with practical activities, although using a variety of tasks (descriptor 8) is still rated second highest for current practice. Teachers value them but it seems they are not using rich tasks in their teaching as often as they would perhaps like to be. Descriptors for the theme of learning move up to be ranked as the second most common group of current practices. The use of contexts for learning remains as the least often used type of classroom practice in Year 11 mathematics. The use of technology is unchanged.

Figure 4 **Ranking that Year 11 mathematics teachers assign to their current practice according to the identified descriptors/themes**

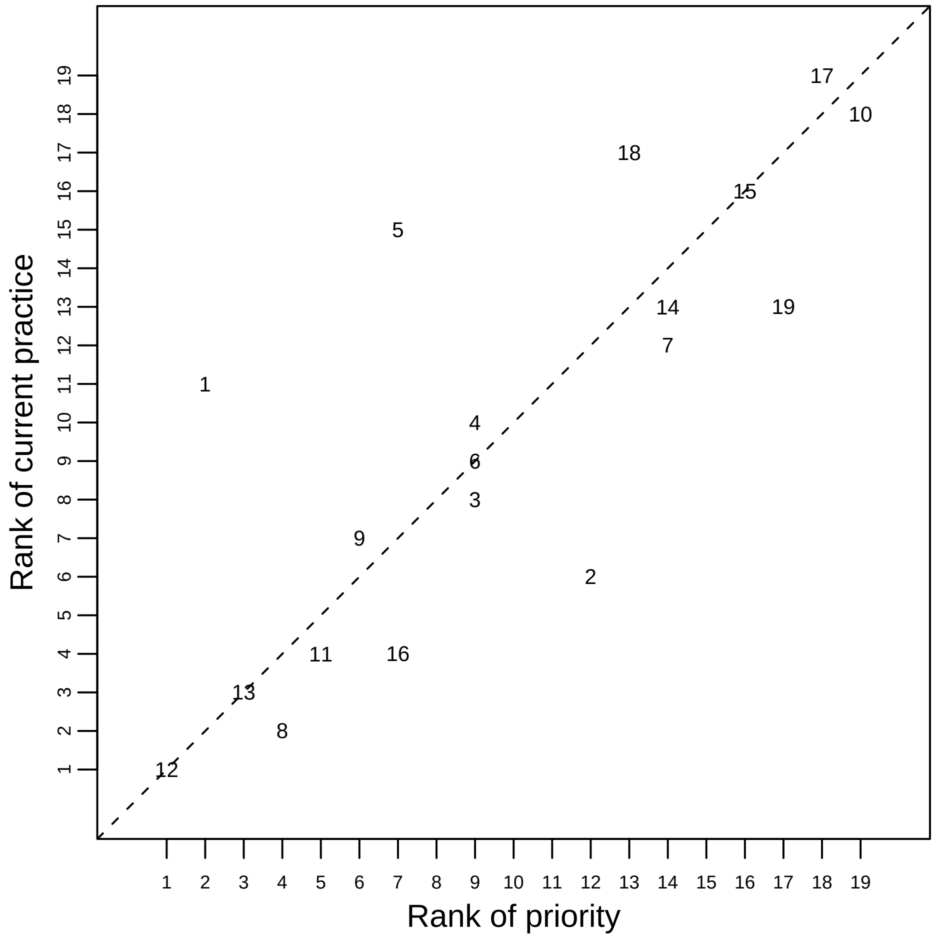
Assessment	12		13	11			9												
Rich tasks		8								4	1				5				
Learning				16		2			6										
Practicals								3					14						
Contexts												7	19			15			17
Technology																	18		
RANK	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**

Comparing priorities with current practice

Figure 5 plots the priority ranking the teachers gave to each descriptor against their ranking of how often they carry out these practices at present. Many small changes in ranking can be assigned to random fluctuation so the discussion centres only on those aspects of practice that have changed by 4 rank points or more. Nevertheless, some interesting patterns emerge.

Points lying on or very near the diagonal line identify descriptors for which the priority rank is the same as the current practice rank — that is, the teachers think they do these things as much as they feel they should do them, or want to do them. Points above the line indicate practices where the priorities teachers would like to assign are ranked higher than their rankings of their perceived current practice — that is, these are things teachers think they do not do as often as they would like. Points below the line show classroom practices teachers currently carry out more often than they would prefer.

Figure 5 **A comparison of mathematics teachers' rankings for priorities and current practice related to the 19 provided descriptors**



Three descriptors differed by five or more ranking points: descriptors 1, 2, and 5. Descriptors 1 and 5 both relate to the use of rich or varied mathematical tasks. Descriptor 1 (providing stimulus material) had a far higher priority rank (second) than either its pre- or post-NCEA rank (tenth and eleventh respectively). However, descriptor 5 (higher-order tasks) has dropped further from its priority rank (seventh), with its pre-NCEA rank of tenth dropping to fifteenth post-NCEA. This confirms that these nine mathematics teachers perceive they use rich tasks less often than they would like, though only some of this mismatch is related to NCEA *(see* Section Five).

Descriptor 2 relates to teaching for understanding rather than for content coverage. With a relatively low priority ranking of twelfth, it has moved up from being ranked eighth equal pre-NCEA to sixth post-NCEA. It appears that teachers perceive they have to do more of this now than they would like, at the expense of some of the other practices they were asked to consider.

The priority of two other descriptors differed from current practice by four ranking points. These were descriptor 18 (using learning technology) where the priority was higher than pre- or post-

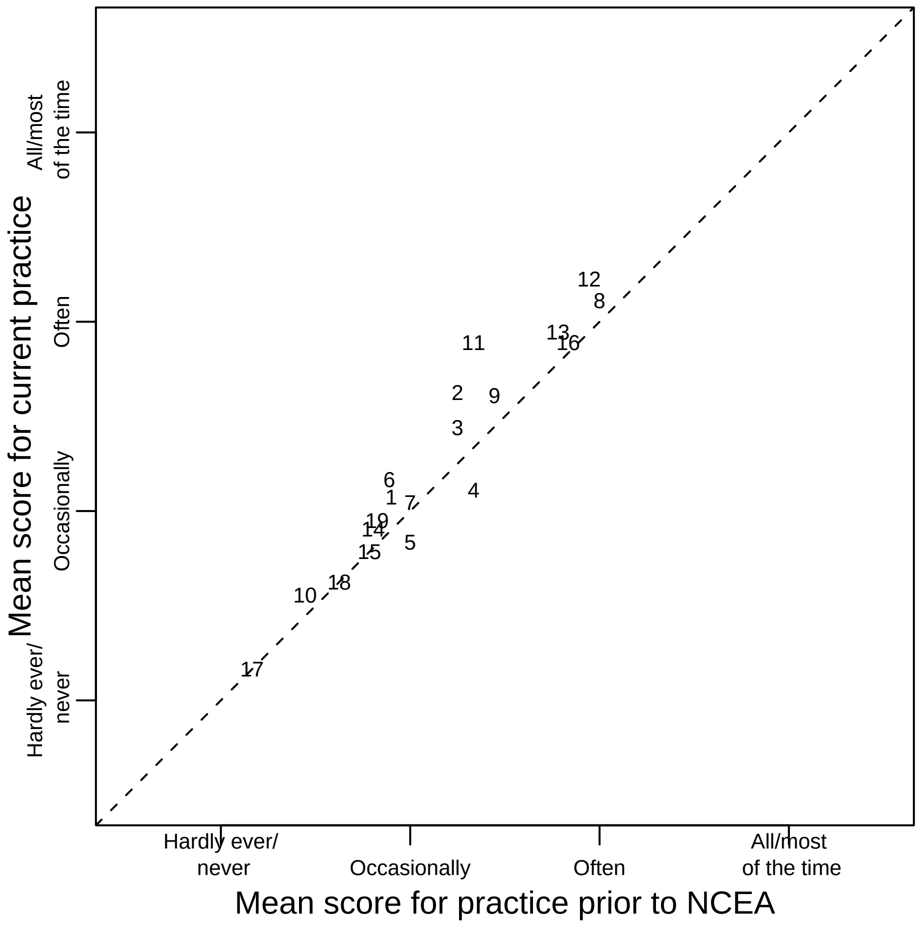
NCEA practice (for which there was no change in ranking), and descriptor 19 (exploring student values) where current practice was higher than its ranked priority, although pre- and post-NCEA ranking were very similar. It seems that mathematics teachers do not use learning technologies as much as they would like and when they explore student values, they do so reluctantly.

Changes in mathematics teaching practice pre- and post-NCEA

Overall the responses show a modest increase in teachers' practice for the majority of the nineteen descriptors, with none of them changing by a statistically significant amount. Figure 6 compares teachers' perceptions of their practice before NCEA and their current practice, based on the average score for each set of responses. Points on the diagonal line represent practices that have not changed with the introduction of the NCEA (for example descriptor 17 — using local community projects as contexts. The study teachers say they hardly ever did this pre-NCEA and they are no more likely to do so now). The movement in score is shown by the vertical distance from the line of the descriptor number. (Data are also shown as the column labelled "S_a – S_b" in Table 4.)

All but two of the 19 practices are scored as being more common now than prior to NCEA. These two (descriptors 4 and 5) both relate to the use of rich mathematical tasks. Both were ranked with substantially lower practice now compared with pre-NCEA. This further reinforces the previous comment which shows that rich mathematical tasks are getting less emphasis than these teachers would like.

Figure 6 **Mathematics teachers' perceptions of changes in practice**



The biggest positive change is for descriptor 11 (ensuring assessment incorporates a range of levels and/or types of thinking). This practice registered an increase in average score of 1.11. While this was not statistically significant (using the Wilcoxon signed-rank test for matched pairs (Wackerly, Mendenhall, and Schaeffer, 1996)), it does represent a move from this being done "sometimes" to being done "often". Teachers rated it relatively high in their priority rankings. As discussed in Section Four, this practice is now seen as helping students to prepare to demonstrate merit/excellence achievement levels in their NCEA assessments.

The second biggest shift was for descriptor 2 (teaching for understanding rather than content coverage). This practice changed by a score of 0.75. However, while teachers say they now do this more often, they do not assign high priority to the practice. In Section Five we note some mathematics teachers' unhappiness with what they perceive to be a "narrowing" of the curriculum because topics/assessments are now seen to need more time and some course components have been dropped.

The science teachers' responses

The average score and the rank for the priority the science teachers assigned to each descriptor are shown in Table 4, along with their average ratings of their practice pre-NCEA and their current practice for each descriptor. As in Table 3, the final column of Table 4 shows the change in the average score pre- and post-NCEA.

Table 4 **Science teachers' perceptions of priorities and changes in practices**

Descriptor	Average priority (S_p)	Priority rank (R_p)	Av. Pre-NCEA (S_b)	Pre-NCEA rank (R_b)	Av. Post-NCEA (S_a)	Post- CEA rank (R_a)	Change ($S_a – S_b$)
1	7.17	7	3.22	9=	3.50	11	0.28
2	7.15	8=	3.50	6	4.69	4	1.19*
3	7.15	8=	3.00	11	3.42	12=	0.42
4	6.08	15	2.70	15	2.64	18	-0.06
5	7.54	3	2.80	14	4.17	7	1.37
6	7.85	2	3.91	3	4.00	8	0.09
7	6.00	16	3.27	8	2.92	15	-0.35
8	7.46	4=	4.80	1	5.15	2	0.35
9	6.85	11	3.90	4	3.69	10	-0.21
10	5.33	19	1.80	19	2.39	19	0.59
11	8.33	1	3.22	9=	5.42	1	2.20**
12	7.15	8=	2.89	12	3.92	9	1.03
13	7.46	4=	3.45	7	4.58	5	1.13
14	6.50	13	4.13	2	4.18	6	0.05
15	6.73	12	2.88	13	3.36	14	0.48
16	7.20	6	3.57	5	4.80	3	1.23*
17	5.82	18	2.43	17	2.70	16=	0.27
18	5.83	17	2.67	16	3.42	12=	0.77*
19	6.20	14	2.29	18	2.70	16=	0.41

* significant at the 5 percent level, ** significant at the 2.5 percent level

Priorities

In science, the separation between the groups of themes was not quite so clear cut as in mathematics. The overall pattern is shown in Figure 7. It can be seen that the rankings for three themes (learning, assessment, and rich tasks) have strong overlaps.

Figure 7 **Ranking of science teachers’ priorities for the 19 descriptors according to themes**

Assessment	11			13				12			9								10
Rich tasks			5	8			1								4				
Learning		6				16		2											
Practicals								3					14						
Contexts												15		19		7		17	
Technology																	18		
RANK	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**

Learning was the theme accorded the highest average priority in science (whereas in mathematics the assessment theme was top rated). Descriptor 6 (metacognition — encouraging students to actively clarify their ideas) was ranked second, the discussion of language conventions (descriptor 16) was ranked sixth, and teaching for understanding rather than content coverage (descriptor 2) was ranked eighth equal. The slightly greater importance that the science teachers accorded these practices may reflect the recent focus on constructivist research in science education, including New Zealand’s internationally recognised *Learning in Science* (LISP) research *(see*, for example, Osborne and Freyberg, 1985). The LISP research was seen as important to her practice by at least one experienced science teacher in our sample *(see* Section Five).

The science teachers’ second highest grouping was for the practices grouped in the assessment theme, although the priority order differed from that of the mathematics teachers. Top ranking was given to the provision of variety in assessment (descriptor 11), feedback to students (descriptor 13) was ranked fourth equal, and the two descriptors relating to formative assessment (12 and 9) were ranked eighth equal and eleventh respectively. (One of these — probing for understanding early in a learning sequence — was top ranked by the mathematics teachers.)

Rich tasks were assigned the third highest priority of the six themes but rankings for the descriptors within this theme were widely spread. The use of higher-order tasks (descriptor 5) was ranked third, having a variety of types of tasks (descriptor 8) was fourth equal, and the use of stimulus materials (descriptor 1) was seventh. However, descriptor 4 (including frequent open-ended investigations) was ranked down at fifteenth place. The mathematics teachers also rated this practice as the lowest of the rich task priorities, but they placed it ninth.

For the lower-ranked descriptors, science teachers seemed to be more in agreement with their mathematics counterparts. They ranked practical tasks (descriptors 3 and 14) eighth equal and thirteenth respectively (for mathematics teachers they were ninth and fourteenth). Like the mathematics teachers, the science teachers gave descriptors related to the theme of teaching in context a low priority. Personal interests (descriptor 7) were ranked sixteenth (mathematics = fourteenth). Personal values (descriptor 19) were ranked fourteenth (mathematics = seventeenth). Local issues (descriptor 17) ranked eighteenth (as this practice did in mathematics) and public issues (descriptor 15) ranked twelfth (compared with sixteenth in mathematics). The slightly higher ranking of the latter factor may reflect the cautious introduction of socio-scientific issues into some science courses.

The use of technology was ranked as seventeenth priority compared with thirteenth in mathematics. The difference may reflect the growing use of graphics calculators in mathematics.

Current practice

Teachers' descriptions of their current practices were broadly similar to their priorities, with the groups of descriptors within each theme being accorded the same relative order of importance, with the exception of rich tasks. (Practices within the theme of rich tasks have become even more variable, and have dropped back to the same average ranking as the practical tasks.) The average rankings the science teachers currently give to 19 identified practices in the second year of implementation of the level 1 NCEA are shown in Figure 8.

Figure 8 **Ranking that Year 11 science teachers assign to their current practice according to the identified descriptors/themes**

Assessment	11				13				12	9									10
Rich tasks		8					5				1								
Learning			16	2				6											
Practicals						14						3							
Contexts														15	7	19	17		
Technology																			
RANK	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**

Comparing science teachers' priorities with current practice

Figure 9 plots the science teachers' priority rankings against their ranking of how often they do these practices now. Numbers lying on the diagonal line represent descriptors for which the priority rank is equal to the practice rank — that is science teachers perceive that they do these things as often as they say they want to. Points above the line indicate descriptors where the assigned priority is higher than the current practice, whereas points below the line indicate the priority is lower than current practice. As in the mathematics responses, there are many small changes in ranking due to random fluctuation.

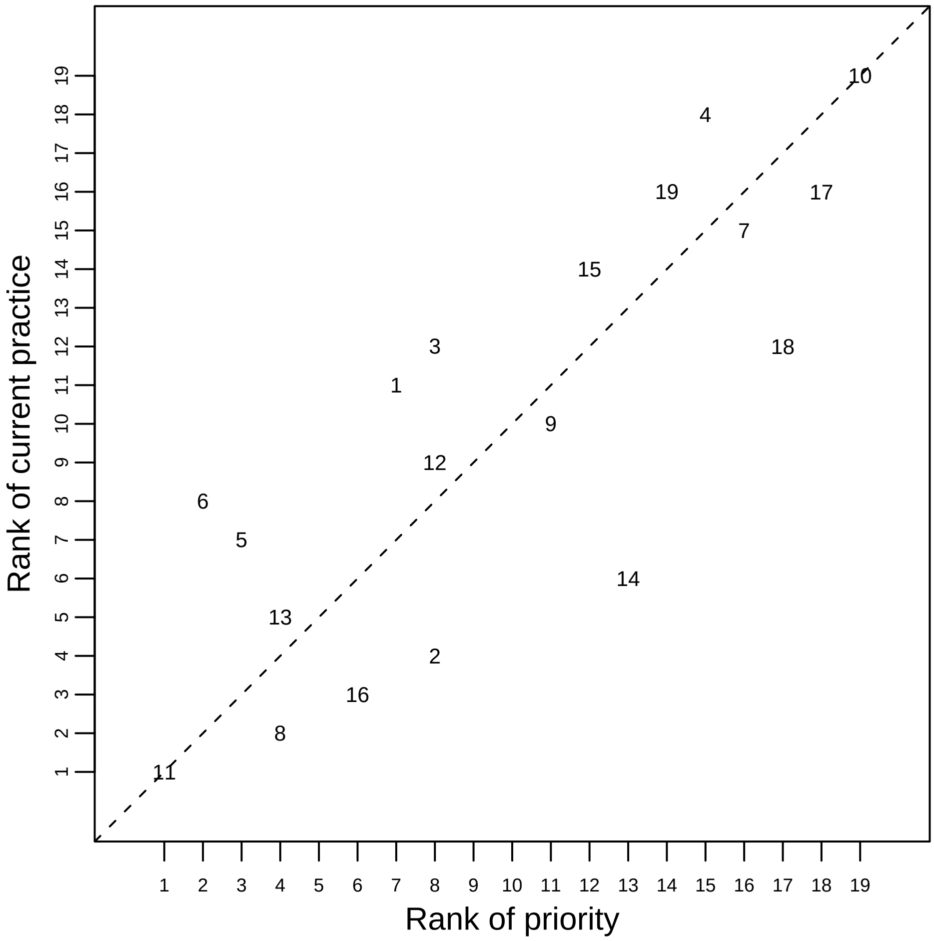
Figure 9 shows large differences between the priorities that teachers gave three of the descriptors and their estimations of their current practice where the differences in ranking are of at least 5 rank points. Practices associated with developing students' metacognitive skills (descriptor 6) are seen as a high priority (second) but were ranked eighth for current practice. PreNCEA the rank for this descriptor was in line with priority.

Conversely both descriptors 14 (using a variety of types of experiments to exemplify scientific methods) and 18 (technology use) were assigned much lower priorities than the science teachers' current practices reflect. These changes may relate to a shift in focus with the introduction of the internally assessed achievement standards in investigative skills (descriptor 14) and research (descriptor 18) although not all the schools are assessing the research standard *(see* Section Five).

Four other descriptors had differences of four rank points between the priority and current practice. These were descriptor 1 (providing stimulus materials), descriptor 3 (students making decisions in practicals), and descriptor 5 (higher-order tasks). Each of these had a higher priority than the current practice — that is, these science teachers want to do these things more often than they feel they can. Descriptor 2 (teaching for understanding) had a lower priority than current practice.

For all five descriptors where there was a mismatch between priority and pre-NCEA ranking there was a better match with current practice. Both descriptor 5 (higher-order tasks) and descriptor 11 (assessment incorporating a range of levels) were underrepresented in practice prior to NCEA but priority and current practice are more closely aligned. Descriptor 7 (using students' interests) and descriptor 9 (using a variety of assessments) were both more common than their priority ranking pre-NCEA but are now more closely matched. (In the case of descriptor 7 it seems that teachers may have been drawing on students' interests more than they wanted to and may be quite happy to do little of it.) Descriptor 14 (using a variety of types of experiments) has moved closer to its priority rank of thirteenth, down from a pre-NCEA practice rank of second to sixth currently, indicating relatively less of this may be happening post-NCEA.

Figure 9 **A comparison of science teachers' rankings for priorities and current practice related to the 19 provided descriptors**

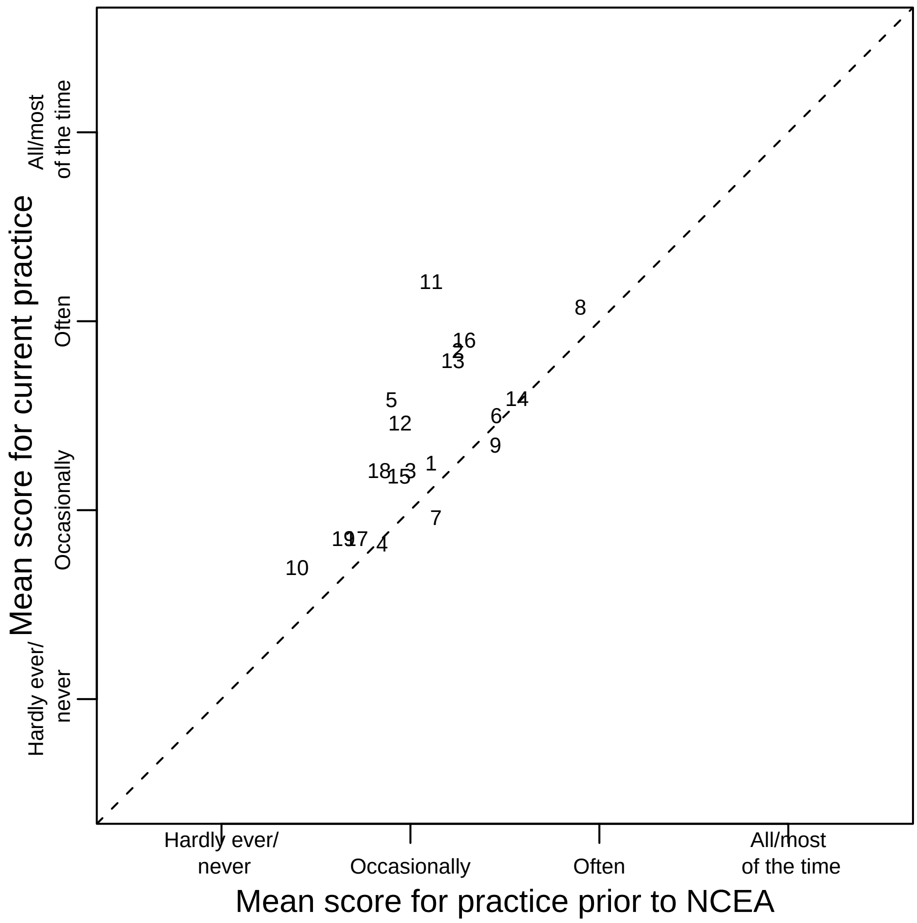


Changes in science teaching practice

Figure 10 plots teachers' average descriptor scores for their perceptions of their practice pre-NCEA against their average scores for how often they do the same things now. Points on the diagonal line represent descriptors of practices that the teachers perceive to be carried out to the same extent now as they were before the introduction of the NCEA.

Any movements in overall scores are shown by vertical distance from the line. (This data is also listed in the column labelled "$S_a – S_b$" in Table 5.)

Figure 10 **Science teachers' perceptions of changes in practice**



In this analysis, most of the 19 practices fall above the line, showing that the science teachers perceive they do these things more often now than they did prior to the implementation of NCEA. The reported increase in practice is statistically significant (using a one-sided Wilcoxon signed-rank test for matched pairs) for four of the descriptors.

Descriptor 11 (ensuring assessment incorporates a range of levels) shows a significant increase in practice at the 2.5 percent level. It has moved from happening "occasionally" pre-NCEA to happening "often" now. It has also increased from ninth rank pre-NCEA to top rank post-NCEA. This change is discussed in Section Five, where it is related to teachers' awareness of the need to help students prepare to demonstrate their achievement at merit and excellence levels. Whereas in the past the most able students demonstrated this by accumulating more marks than other students, they now need to be able to demonstrate their academic ability in qualitatively different ways.

The other three practices that show significant increases are learning for meaning rather than for content coverage (descriptor 2), developing understandings of use of language conventions in science (descriptor 16), and use of technologies (descriptor 18). The latter shows only a small absolute shift but reaches statistical significance because *all* the teachers who recorded a shift in practice say they are using more technology now. While descriptors 5 (use of higher-order tasks), 12 (probing student understanding), and 13 (giving ongoing feedback for learning) all showed larger absolute shifts in the average score, at least one teacher in each case had said they were doing less of that practice now. This made these overall shifts non-significant. Descriptor 5 has, however, a much higher practice rank now than it did pre-NCEA.

Two descriptors had a lower level of practice now than pre-NCEA. These were descriptors 7 (using students' personal interests) and 9 (using a variety of assessment methods). Both of these, along with descriptor 6 (encouraging students to clarify their ideas) had a much lower-rank on current practice than they did prior to NCEA.

Of these changes, teachers may be least happy about the increased use of technology since they rank it as a low priority but say they are doing more of it. Because this practice was not a direct focus of the research, this is a pattern that remains to be more fully examined.

Significance of main patterns

In Section Two we described the development of the teacher self-reflection sheets as a direct response to the perception that "things have gone backwards" in mathematics and science classrooms since the implementation of the NCEA began. To test this perception we selected descriptors that have been identified by other researchers as representing a range of aspects of best practice and we took care not to lead indications of change in the practice of these descriptors in either positive or negative directions.

The analysis reported above shows the general perception of "backwards" changes to be unduly pessimistic — at least as far as the practice of this small group of teachers is concerned. There are some differences between the ranking of priorities and practices reported by the mathematics and science teachers, but the broad thrust of the responses is similar. The patterns generated show a mix of changes pre- and post-NCEA implementation at level 1. Some are perceived to be positive by these teachers and other changes are perceived to be negative.

Both mathematics and science teachers say they are now spending more time on ensuring that assessment incorporates a range of levels and/or types of thinking (descriptor 11). As Section Five shows, this change is directly linked to assessment requirements for demonstrating merit and excellence for achievement standards and teachers are actively looking for ways to develop this change further. Both mathematics and science teachers also say they are now spending more time teaching for understanding rather than for content coverage. They appear to be less happy with this change, with both groups assigning it a lower priority ranking than current practice ranking. It seems unlikely that these teachers would not want their students to understand their work. We read this pattern as representing concerns about what some teachers see as a "narrowing" of the curriculum — that is, in responding to descriptor 2 they have emphasised the "less content coverage" part of the descriptor.

Changes that might be construed as "going backwards" differ for the two groups. Mathematics teachers say they are using less open-ended investigative tasks and less higher-order tasks now (descriptors 4 and 5). This accords with their assertions that it is not easy to fit rich mathematical tasks into their programmes now *(see* Section Five). However it should also be noted that the teachers rated such tasks a higher priority than their pre-NCEA practice ranking. It may be that the difficulty of fitting such tasks into a mathematics programme has been exacerbated rather than arising as a new issue.

Science teachers say they are using fewer strategies that help students to clarify their own ideas (descriptor 6). While they rate these a high priority, both pre- and post-NCEA practice are ranked lower (although only just so for pre-NCEA). This is a similar pattern to that shown for rich mathematical tasks. Again it may be that the NCEA implementation has exacerbated an existing tension in competing classroom priorities, rather than arising as a new issue.

Both mathematics and science teachers would like to provide more stimulus materials that get students discussing ideas than they currently do (descriptor 1). This is clearly compatible with their desire to use rich mathematical tasks and/or ideas that develop students' metacognitive awareness. It may be that provision of classroom-ready exemplars could support teachers to more closely match their practice to their priorities here — especially if such tasks can be linked to the desire to help students demonstrate merit and excellence in their NCEA assessments.

The next two sections put the flesh on these quantitative bones by reporting and discussing the teachers' qualitative responses to the open interview questions.

Section Four

Summary of findings from interview questions

This section summarises the key responses made to each of the 16 interview questions *(see* Appendix 2 for interview questions). The findings describe selected teacher beliefs and departmental/school policies and practices related to the NCEA implementation. These findings provide a context for the in-depth discussion of teacher decision making that follows in Section Five.

Familiarity with, and use of, unit standards (Q1)

1. Eight of the nine mathematics teachers and seven of the nine science teachers had used unit standards in the past. The other two science teachers had taught modular courses and so were also familiar with internal assessment procedures, if not standards-based assessment.
2. All the mathematics teachers said their schools currently use unit standards within their NCEA programme, especially for the low-ability students taking MAP courses (mathematics applied programme). Seven of the nine science teachers said their schools use unit standards-based courses for lower-achieving students, in some cases in combination with one or both internally assessed level 1 achievement standards. In both subjects, unit standards-based assessment is seen to give these students recognition for their work and a sense of achievement on work that is accessible to them.
3. "Average" students may study courses that are assessed with a mix of achievement and unit standards, although this is more common in mathematics than in science.
4. The support materials developed by the NZASE for unit standards-based science assessment are popular and used by the majority of schools in the sample.
5. The internally assessed nature of unit standards is viewed positively.

6. The flexibility of course design using a mix of unit and achievement standards is a positive feature of NCEA. In all but the smallest schools, courses can be tailored for the learning needs of different groups of individuals.
7. The perceived lack of comparability between unit and achievement standards causes some concern in mathematics as they are seen as easier to obtain than achievement standards. The lack of "excellence" or "merit" for unit standards was mentioned by some mathematics teachers.
8. Some teachers expressed the view that unit standards can be "pedantic" or "picky" although they also think they have improved over time and that experienced teachers now understand how to use them more effectively.

Preparation for NCEA assessments (Q2)

1. The use of practice tasks or exams is the major form of preparation for formal assessments in both subjects. Most teachers now focus on the quality of answers, preparing students to display the critical thinking skills that can take them beyond "achieve" to the "merit" and "excellence" levels of the various standards.
2. Examination techniques for external standards, especially time management skills, are more important now than previously. This was particularly true for one mathematics school which had previously done all School Certificate assessment internally.
3. Practice assessments are being used as a formative assessment tool. However, formative assessment focuses on the learning to come, not to prepare for the summative assessment to come.
4. Literacy in both its more general sense, and literacy skills specific to mathematics and/or science are seen by some teachers as key factors required for success. The style of NCEA assessment makes this a high priority for targeted teaching.
5. "I can do" sheets assist students to understand course requirements.
6. Preparation for external examinations is hard for both students and teachers because the newness of the NCEA initiative means they are uncertain what to expect.

Teaching style (Q3)

1. Practically based internal assessments are leading to more meaningful hands-on mathematics in the classroom. Science investigations now pay more attention to analysis and interpretation phases than may previously have been the case.

2. More in-depth teaching and an emphasis on teaching for understanding are being given more prominence in both curriculum areas. There is an increased emphasis on teaching students to explain and justify their answers in mathematics and to link recall of content to explanations and discussion in science. However the content of the internally assessed standard for measurement is arguably of too low an academic level for level 1 of the NCEA.
3. Teaching to the assessment is common in mathematics because of its well-defined, relatively closed and predictable nature. Schools in this study are doing fewer investigative, exploratory-style teaching activities in this subject.
4. Streaming of classes is common in mathematics. In both subjects, teaching may focus on the level of attainment that is seen as realistic for different groups of students. "Top" students may be challenged with "excellence" work but other students will be guided to focus on answering to "achieve" level only.
5. Aspects of teaching related to "coverage" of traditional curriculum content remain much the same as they were pre-NCEA, especially for subjects that are externally assessed.
6. Some teachers say they are less able to be spontaneous and to follow issues or interests that arise in science because of the narrowly defined focus of the achievement standards. Five mathematics teachers and several science teachers think the change in assessment regime has "taken the fun away".

Impacts on students (Q4)

1. Many positive effects of the NCEA implementation were noted, including improved self-regulation, thinking processes, and a wider skill base in mathematics.
2. The emphasis on learning for understanding is seen to have a positive effect in science, and the transparent nature of the standards and of teacher expectations boosts student confidence.
3. For some students, there is a strong emphasis on collecting credits, together with a subculture that "achieved" is good enough. This does not set Year 11 students up for future mathematical success, as "achieved" is very skill-based. Science teachers are concerned that students with "achieve" level passes will struggle with level 2 of the discipline-specific science curricula.
4. The pressure of time needed to complete all learning and assessment tasks is seen as having a negative effect on some students. Some teachers see this as largely the fault of an overfull curriculum, not the assessment system. There is more pressure to fully complete a standard. If parts of the standard are ignored, the students cannot get "merit" or "excellence".

Balance between internal and external standards (Q5a)

1. Schools had a variety of views on whether these are in balance or not, and they are making curriculum decisions that result in variations of this balance from school to school *(see* Section Five).
2. Schools report various ways of allocating time to the different standards. These include using historical data, allocating time evenly, or basing it on either the content level or the number of credits in the standard.
3. Most schools have either already restricted the number of standards being offered, or are considering doing this in the future.
4. Some schools are assessing internal standards in Year 10 to reduce the load in Year 11. These are in either measurement 1.3 or geometry 1.4 in mathematics and investigation (1.1) in science.
5. Topics and/or skills that are internally assessed are usually done early in the year, or one per term for the first three terms.

Curriculum coverage (Q5b)

1. All schools commented on the lack of linkages between the strands of mathematics and the compartmentalisation of the subject.
2. The majority of schools say fewer mathematical processes are being taught or assessed except perhaps at the "excellence" level. Communicating mathematical ideas is being taught and assessed.
3. There is a perception in six schools of a narrowing of the mathematics curriculum. In four schools science content is now divided between different courses, so that not every student will do every topic as they may have in the past. Some teachers say reducing content coverage allows for greater depth in the teaching.
4. Geometry standards, both internal and external, are the ones most likely to be dropped from Year 11 mathematics courses, leaving geometry underrepresented. However, some schools see this topic as being useful to mathematics, with students enjoying it and experiencing success.
5. The externally assessed standards most likely to be dropped from science courses are earth science and astronomy. Four schools are not offering the internally assessed research standard as part of their science programme.

Time spent on assessment (Q6)

1. Almost all schools see assessment, and particularly internal assessment, as taking up more time now compared with teaching. Mathematics teachers tend to think this is more of an issue than do the science teachers, who may only be assessing one standard internally. Some see this as being only a temporary occurrence, related to their initial unfamiliarity with the new assessment methods.

2. Two-hour long formal internal assessments are seen by some mathematics teachers as causing particular issues of workload and manageability. Some are reverting to one-hour classroom-based assessments. Those schools that assess science investigation skills under examination conditions also face issues of manageability and preparation time.

Reassessment (Q7–8)

1. Schools have a wide range of reassessment policies and practice on reassessment. In mathematics, two schools do not offer any and one offers it rarely. In science, two schools do not offer reassessment and two offer it sometimes. These schools say they are encouraging a "do it once, do it well" attitude amongst the students. Of the schools who always offer it, most reserve the opportunity for "not achieved" students in mathematics. In science, some schools allow students to try for higher-levels of achievement in a second investigative context although many students are not interested in improving on the initial award. In some schools reassessment is compulsory, while in others it is optional.

2. Schools offer different opportunities to prepare for reassessment. Some have well-structured tutoring systems available. One school has a separately funded provision for reassessment, with extra staff employed to do it.

3. The timing of the reassessment varies from about two weeks to several months from the original assessment. The latter may use exam times or NCEA jumbo days.

4. The use of oral evidence to re-grade the initial assessment was mentioned by two mathematics teachers and by one science teacher. Policies on this may well differ.

5. Some mathematics teachers expressed diffidence about the appropriateness of reassessment. Science teachers were more likely to comment on its practicality (for example finding an appropriate time, laboratory preparation, and so on).

Other assessment issues (Q9)

1. The manageability of internal assessments was a commonly mentioned issue. Equipment, conducting outside/field trip assessments, and timing or timetabling were challenges mentioned by both mathematics and science teachers.
2. Task security was seen as an issue by some mathematics teachers. Keeping a task secure before administration, and invigilation during the assessment to ensure the work was the student’s own were both seen as problematic for internal assessments.
3. Finding quality assessment tasks was a common issue for both mathematics and science teachers but informal sharing of assessment resources is common. A larger bank of moderated and trialled resources, for both internal and external assessments, would be useful. The constantly “changing playing field” with respect to appropriate tasks is seen as unhelpful and teachers in both curriculum areas would like the ongoing modifications to the web-based resources to keep abreast with policy decisions.
4. Increased administration was cited by most with increased record keeping and student tracking being mentioned. Storage of student work is an issue, as are constant meetings. Some teachers think these aspects will settle down with time.
5. Photocopying, especially for assessment materials, has escalated in many schools.

Moderation procedures (Q10)

1. Across the schools a range of different moderation systems for checking tasks and marking internal assessed standards is employed. Each school aims for reasonable internal consistency and there is considerable teacher interaction. However, teachers in the smallest schools have to seek that interaction with teachers from other schools.
2. Careful consideration of grade boundaries is needed, with teachers sometimes seeing these as “grey areas”.
3. Some mathematics teachers voiced reservations about the consistency between-schools. One thought that the different tasks had more variation between them than the variation of within-task marking, making task selection a key parameter.
4. Some mathematics teachers also commented on procedures for making the task administration and verbal instruction to students as uniform as possible between classes.
5. One science teacher commented on differing interpretations of the same [Internet-sourced] task made by different external moderators and another science teacher said that unit standard moderation had been more “user friendly”. Delays in moderation can cause problems when feedback about the task design is received after the task has been used.

6. While science teachers feel they have gained in confidence and consistency when moderating level one tasks and marking, those in smaller schools are less confident about level 2 if they have no-one else on staff with whom to discuss their specialist discipline.
7. One science teacher is planning a course that incorporates aspects of the **Putaio** document as well as aspects from *Science in the New Zealand Curriculum*. He anticipates moderation challenges because he will only be able to work with other teachers who also have a science/te reo/nature of science combination of knowledge and skills.

Impact of moderation discussions (Q11–12)

1. Moderation meetings are seen as having positively influenced teaching practices in most schools.
2. Marking practice tasks helps teachers clarify expectations and can be used to coach students for the formal assessments to come. One science teacher said it "makes everyone aware of what is important".
3. Diagnostic feedback on common student misconceptions also assists teaching and learning. Often this will not be until the following year, or for reassessments. Moderation meetings on practice tasks allow issues identified to be fed back to students.
4. The focus on achieve/merit/excellence has made some science teachers more aware of the need to pose open-ended questions, especially in assessment tasks.

Impact on target groups (Q13)

Low achievers

1. Almost all mathematics and some science teachers think the NCEA has improved the situation for these students. Spreading assessments over time, while working at a slower pace to ensure understanding and focusing on a smaller number of standards are ways of making courses achievable. Students like the recognition of learning via gaining a nationally recognised qualification.
2. Internal assessment works well for low-achieving students. However, some teachers are anxious about what they perceive to be a lack of equivalence between unit and achievement standard credits.
3. For low achievers who take externally assessed achievement standards the situation can be worse as the work can be too hard and they then get no recognition. The more open format of some science questions disadvantages these students.

4. High rates of absenteeism can compound learning disadvantages — especially where students miss out on opportunities for internal assessments.
5. Low achievers in mathematics may be too focused on obtaining just the eight numeracy credits needed to obtain level 1 of NCEA.

Average students

1. Some schools found NCEA more achievable for this group, saying the initiative had made it clearer to them what they had to do to succeed, and/or that gaining credits in advance of the external examinations boosts their confidence.
2. Some teachers feel students in this group are "doomed to be forever achieved" because they lack the ability to move up to merit level, while others see striving for merit as an achievable goal.
3. In any case, significant numbers in this group may aim for "achieved" only and this is usually insufficient for further mathematics advancement. Some focus too much on merely accumulating credits.
4. The high literacy demand of more open science questions is seen as disadvantaging some students in this group.

High achievers

1. The incentive of "excellence" awards works well for the very top achievers in science. It is not seen as trivial and science tasks that were used for School Certificate often have to be taken to a higher-level to provide an opportunity to demonstrate ability to meet specified excellence standards. With the new emphasis on thinking skills, one teacher thinks there has been some change in the type of students who now achieve excellence in science.
2. Some mathematics teachers see excellence as too hard, or as requiring too much work for insufficient reward (since it generates no more credits than "achieve"). A common reaction was that some students who deserved "excellence" in mathematics are not getting it for "picky" reasons. This is de-motivating, and some students settle for "merit".
3. One science teacher thought that last year's examination did not provide sufficient chances for students to demonstrate excellence, which was "not fair".

Mäori / Pasifika students

1. Views were varied, and many schools would not be drawn either way. Some saw it as more of a socio-economic issue, with more of this group from low decile backgrounds.

2. In the two “science schools” with the highest Mäori and Pasifika student numbers, the teachers appreciate the opportunity to introduce contexts relevant to their students, and feel students have a better chance of gaining credits from internally assessed tasks because expectations are clearer.

Gender

1. Only some schools commented. They were cautious in their statements, but the consensus was that the NCEA suits girls better and boys less well.
2. Girls are seen as more organised, mature, steadier workers with better language skills.
3. Boys are typified as more minimalistic and goal focused. One teacher commented on the lack of Mäori male role models to help make achievement seem more acceptable to boys in the school.

Other impacts on students (Q14)

1. Student stress is seen as an issue by some teachers of both subjects. Continuous assessment, the assessment regime starting so early in the year, and the downward push of standards-based assessment to Years 9 and 10 are all seen as issues. This leaves less time for non-curricular activities in or out of school. There is some burnout of serious but not high-achieving students. One mathematics teacher commented “It is crushing for such young students. There is too much too soon. The formal assessments are too early in the year and the pressure is on very early.”
2. Strategies for ensuring students get the eight credits they need in numeracy to attain level 1 of NCEA were discussed by some schools. Some finish getting these in Year 12 courses.
3. Students and parents continue to need quality information on the NCEA. Some students are struggling because of not coming to terms with the new system, but others have “caught on fast” and may now be very “credit focused”. It is problematic if they see gaining credits as the main purpose for their learning.
4. Subject compartmentalisation has the upside of allowing a fresh start for students who may have failed previous standards. This works if they taste success early.
5. Problems need to be addressed for students who change schools. Credits may be offered at different times or in different ways.

Impact on teachers/ other issues (Q15–16)

1. In both subject areas, teacher workload has increased significantly as a result of the NCEA, especially in administration and assessment. Some teachers see their focus being deflected from teaching and learning to making the system work. However, making refinements to the NCEA is seen as preferable to abandoning the system.
2. In science the flow-down of standards-based assessment to Years 9 and 10 has added significantly to workload pressures.
3. Support by and wide consultation with colleagues is the main way of handling the increased workload and the NCEA training days are useful in bringing schools together.
4. Some teachers feel they have more ownership of the new system, and feel it increases teacher accountability and expertise. They are becoming less anxious and more realistic in their expectations.
5. The design of the TKI website could be made more user friendly.
6. Opportunities to develop more cross-curricular courses now exist, but student learning needs must be balanced against other pressures such as university entrance requirements.
7. The NCEA works best if the whole staff is supportive. Students and parents need to be given consistent positive messages by staff who show a united front.

Features of the NCEA implementation context in the study schools

In all eighteen subject departments there is a strong focus on working collegially to implement the NCEA. All schools have modified the curriculum they offer in mathematics or science, and more such decisions are pending, with the aim of better accommodating the perceived learning needs of students in the new assessment environment. Teachers would like to have access to a wider range of examples of assessment tasks. In most schools they are now redesigning Year 9 and 10 tasks to standards-based formats.

Most schools have well-established internal moderation policies and practices, and teachers value the professional discussions that take place. Redesigning tasks to standards-based formats, internal assessment, internal moderation, and reassessment (where this happens) have added considerably to teacher workloads.

The findings summarised above suggest that, in these schools, students in science and mathematics classes with different types of learning needs may be experiencing the NCEA differently:

- students are likely to be in streamed classes, especially for mathematics;

- “less able” students are likely to have a higher proportion of their course internally assessed and to be assessed with unit standards rather than achievement standards;
- “able” students are more likely to be encouraged to try for merit or excellence level awards, especially for externally assessed standards;
- students in some schools are well supported to try for reassessment but those in other schools get “one shot” at internally assessed achievement standards; and
- learning needs related to differences in cultural backgrounds are no more or no less likely to be accommodated now than they were pre-NCEA.

Section Five

Reporting on shifting balances

> Teaching is not just a collection of individual features. It is a system composed of tightly connected elements. And the system is rooted in deep-seated beliefs about the nature of the subject, the way students learn, and the role of the teacher. Attempts to change individual features are likely to have little effect on the overall system. The changes often get swallowed up or reshaped (Stigler and Hiebert 1997).

Has the NCEA actually changed classroom practice, or have aspects of the initiative with the potential to impact on learning simply been "swallowed up or reshaped"? Section One described how we teased out the three initial questions to identify a series of "balances" that could potentially impact on the type of learning students are now experiencing in mathematics and science classrooms. Section Three reported the overall pattern of changes that teachers say have taken place in Year 11 mathematics and science teaching since the NCEA was introduced at level 1. Section Four reported on teachers' perceptions of the impact of the NCEA on students' learning and documented their curriculum decisions and assessment practices related to its implementation. In this section we expand on these findings by analysing them with respect to the key balances we identified. These balances are:

- time devoted to learning balanced against time devoted to assessment;
- use of internal assessment balanced against use of external assessment when assessing for qualifications;
- time devoted to developing new "content" knowledge balanced against time devoted to the development of new skills and/or the exploration of attitudes and values;
- a direct (acontextual) focus on concepts/facts/skills balanced against teaching that embeds learning in contexts of relevance to students' lives and interests;
- tool/methodology acquisition by direct "skill and drill" balanced against acquisition via open problem solving/investigations;
- participation in teacher-directed learning activities in which the teachers' ideas take precedence balanced against participation in activities that are student-led and/or in which students determine the pace and sequence of learning and/or actively contribute their ideas;

- time when students learn as individuals balanced against time when they participate in group learning activities; and
- a focus on the cognitive/conceptual aspects of learning balanced against a focus on the metacognitive — that is, students' thinking about their thinking and learning.

In the sub-sections that follow, shifts in these balances are documented if and where they have been detected, with some exploration of what these balances mean for mathematics and science teachers.

Balancing learning and assessment for qualifications

> The talk in our staffroom is about dropping the number of credits that we set so that we're not driven by assessment and we can get back to delivering the curriculum. So we'll still teach the whole curriculum, we'll just choose to assess less of it. …
>
> On the one hand, the ministry itself is saying perhaps schools should be assessing less; on the other hand, the minister is congratulating those students who have gained the most achievement standards, who have got the most credits – and that in a way is going to set up a new league table, which is really difficult (Margaret McLeod, Principal of Wellington Girls' College, as reported by Welch, 2003, p. 20).

There has long been critical comment that assessment can be the "tail that wags the curriculum dog". The article cited above (Welch, 2003) links the "overstuffed" curriculum to a market ideology from the early 1990s that privileged knowledge as a commodity. Accordingly, some academics urge teachers to resist what they see as neo-liberal agendas that would have school education serve the self-interests of the competitive "market" and those who benefit from it *(see*, for example, Gale and Densmore, 2003). However, assessment for qualifications carries high stakes for students and teachers and resistance can be easier said than done. For example, despite the many assumptions and misleading patterns, "league table" comparisons of schools' overall achievement profiles have continued to be published in the standards-based regime, just as they were when assessment was norm-referenced. Furthermore, individual teachers can now be provided with much more specific data about the patterns of achievement of their students for the various individual components of their courses. Their "success" as teachers can be quantified for each individual part of their courses and compared with that of their colleagues — in the next door classroom or even, potentially, in the "school down the road". There is evidence that schools are already using such analyses to report to their Boards of Trustees (Hipkins et al., forthcoming).

On the other hand, the introduction of the NCEA has provided schools with the means, at least in theory, to assess less and hence free up time for teaching and learning that they consider will be of most value for their students. (This assumes of course that "value" is seen as something other than, or in addition to, passing examinations and accumulating credits for their own sake.) The Minister of Education has signalled that credit reduction should be considered (Mallard, 2002; Minister of Education, 2002). In the *Learning Curves* schools we have found that principals are relatively more keen for their staff to consider credit reduction than the staff are to oblige (Hipkins et al., forthcoming). The stumbling block seems to be that teachers are accustomed to "selling" school learning via the assessment opportunities generated — "you'll need to know this — it could be in the exam". In the NCEA regime this form of extrinsic motivation seems to have translated to credits that can be gained. Students, too, like the novelty of being able to accumulate credits, at least initially. Again the *Learning Curves* study shows that this may wear off at Year 12, with consequences that are worrying for schools.

To what extent do the very real pressures generated by assessment for qualifications impact on the decisions teachers make about what their students will learn and how much time they will devote to various aspects of their courses? Have teachers been able to make changes that move away from an "overstuffing" of content to focus on deeper and more meaningful student learning? Against the background sketched above, we report on changes to the balance between various types of learning and assessment activities in our case study schools.

Is more time now being devoted to assessment for qualifications at the expense of learning?

In both subjects teachers noted the general problem of finding sufficient time to cover all that needs to be done in class. Although they held varying opinions as to the extent of the impact of assessment on teaching time, most of the 18 teachers said that assessments are taking more time since the introduction of NCEA. No teachers responded that assessment is taking "less" or "much less" time. Responses reported in Section Three show that both mathematics and science teachers rate various types of assessment activities as a high classroom priority.

Table 5 **Teachers' estimations of the increase in class time taken for assessment**

Degree of increase	Much more	More	Hedged between more and same	Same
Mathematics responses	4	4	1	0
Science responses	1	6	1	1

The mathematics teachers expressed more concern than the science teachers about the increase in time taken for assessment. One suggested the increase was as much as three times, while another thought it was up by 50 to 100 percent. However one mathematics teacher suggested the increase applied to internal standards only and said there was no change for topics assessed externally. Five of the mathematics teachers pinpointed the internally assessed achievement standards as a large contributor to the increased assessment load. Comments such as “Time for an internal assessment can be quite long and hard to manage” were typical. Two schools mentioned that having two-hour assessments on each of three internally assessed standards added to the workload, especially given the need to do practice tasks which could also take this same time to administer. One mathematics teacher commented that the regularity of assessment was an issue, rather than the overall increase in time taken. The same point was made by one science teacher.

Perhaps because most of them do not assess a large proportion of their courses internally, the science HODs tended to be more sanguine. One said “There is not much change really” in the time spent on assessment but noted that the *quality* of assessment has improved and each one is more comprehensive. These teachers did say that investigations take much longer than they had initially anticipated but one noted that she would want to spend the time developing students’ investigative skills as the school now does, regardless of whether the NCEA was a factor or not. Another commented that the line between assessment and learning is, in any case, blurred, saying that “Formative assessment *is* learning”. A fourth science teacher noted a shift in his approach to formative assessment. Whereas he had previously interpreted this as a “practice test” the feedback to students now needed to say more about their thinking and understanding, not just to tell them whether they were “right or wrong”.

One science teacher hesitated before answering the question and could not decide between the “same” and “more” responses. She said “It feels like it’s more but it may be that the past system was so familiar. You built it into your teaching then, for example by saying “this is a classic School C question”. A good teacher is ultimately preparing kids for exams.” Thinking further along the same lines, three of the mathematics teachers went on to say that they expected the situation to improve as they became more familiar with the new assessments and as more resources become available. One has spent less time on assessment this year than in 2002. Another commented that “There is less load for HODs on writing exams and the load is spread more evenly between teachers.” This teacher also commented that “In-class time on assessment is the same as they always had end-of-unit tests. Now they use unit standard or achievement standards tests instead.”

Time taken for reassessment

One aspect of the NCEA with the potential to impact on available teaching and learning time is the process of reassessment when students do not achieve a standard, or in some cases do not achieve to the level they would like. In order to gauge the impact of this standards-based practice we asked teachers about their reassessment policies and processes. Six of the mathematics schools and five of the science schools offer reassessment for the internally assessed achievement standards. The range of reassessment practices differed widely between-schools.

Reassessment practices

In one mathematics school it is compulsory for students who have "not achieved" to re-sit the assessment. However, while this may seem like "a reasonable idea", the teacher commented that "You can lead a horse to water but you can't make it drink. Being compulsory defeats the purpose."

A range of methods is used to prepare students for reassessment in mathematics. Three schools use tutoring sessions outside school hours. In two of these schools students must give evidence that they have prepared for the reassessment by attending the tutoring sessions, or undertaking some equivalent preparation. One of these schools has money available to pay for a suitable non-staff member, with the teacher commenting "You can't do it and full-time teach as well." Another school mentioned the availability of informal tutoring but noted "a need to set up individual reassessment opportunities in teacher time". Another school currently uses class time. One other teacher spoke out against this practice, saying, "Any programme which leads to individual attention for some of the class means the teacher must split into 2, 3,…, 6 pieces and have a programme of work for each student. Often the student who needs help dominates the attention that can be given to the other students."

The timing of the reassessment in mathematics differed. In three schools it was as close to the original assessment as possible, usually within about two weeks. In other schools it was done during convenient times such as exam slots, NCEA jumbo days, etc. In these cases the elapsed time to the reassessment could be quite large. Three schools mentioned that the reassessments were done outside class time, while three were not specific on when they were done. One school mentioned that some students were using the local technical college for reassessments.[11]

Teachers in two mathematics schools discussed the issue of seeking oral feedback from a student to clarify a grade. One of these said they "Do oral checks with students on the original assessment as long as it is not teacher-directed. This is more a resubmission rather than reassessment. At times we get a student to repeat just a part of the initial assessment." They commented that the acceptability of this practice was not always clear to them and that "It is a national qualification but schools are doing it different ways. It's hard to know where to draw the line on what is acceptable in reassessment/resubmission." One school that did not offer reassessment made a comment that seems to hint that the use of oral feedback or resubmission would be a potential solution. They said "There are so many minor errors that the students make that they can correct themselves with no extra teaching."

To make teacher workloads more manageable, one science school has scaled back its reassessment opportunities, after having offering "multiple" chances last year. Students in this school may be reassessed to improve their level of achievement. Another offers reassessment

[11] We did not gather data that would allow us to comment on students' success rate when reassessed. This could be interesting to investigate further.

"sometimes" where students fail to achieve the investigation standard, but for time reasons students are never reassessed to improve their level of achievement.

Some mathematics schools offer reassessments for only some of the internal standards. For example, one does not offer it for Statistics 1.5, while another only does so rarely. This latter school does not offer reassessment for Measurement 1.3. One school reassesses in rare cases where a student is close to reaching the "achieved" standard. There was consistency on the number of reassessment opportunities in mathematics schools — these are limited to one. Two mathematics schools also mentioned the opportunity for reassessments for unit standards. One science teacher noted that reassessment may not be possible where the initial investigation was part of a one-off event such as a field trip.

Two mathematics schools and two science schools do not offer reassessment at all. One mathematics school did so in 2002 but has changed its practice. School-wide policy not to offer reassessment in any subject precludes this in the other of these schools.

Formative assessment practices

Comments about formative assessment were most often made during the discussion of teachers' priorities and shifts in practice related to these *(see* Section Three). However there seems to be a common misconception that practice tasks equate with formative assessment. Where formative assessment is used at the beginning of a unit, it seems to be interpreted as a "pretest" that can be compared with a similar, or even identical "post-test" at the end of the unit. More usually, "practice runs" when nearing the end of a unit are seen as a type of formative assessment.

When these teachers conduct "practice runs" for internally assessed tasks, their focus is primarily on achievement in the summative *assessment* to come. By contrast, formative assessment, as this term is used in assessment theory and research, is focused on the *learning* to come (Harlen, 1998) and actively involves students in making judgments about their learning progress (Black and Wiliam, 1998).

All nine mathematics schools in the sample mentioned the use of practice tasks to prepare students for the final assessment of internally assessed standards. Three schools said they made the tasks as close as possible to the actual assessment task. On the other hand, three schools mentioned that they used practice tasks that did not mimic the assessment task and that they gave as wide a range of tasks as possible. One teacher found it useful to mark and discuss the practice task in class. One school integrated the practical aspects into the class-work instead of doing a practice task before the final assessment task. Practice exams were the main preparation for external standards other than ordinary course work.

Only two mathematics schools mentioned formative assessment as a process that is distinct from practice tasks which are more summatively focused. One teacher had been exploring formative assessment theory while the other mentioned "Using formative assessment to identify where students are on the continuum". However, teachers in another three schools commented that feedback to students is much more timely now. One teacher said “The immediacy of teaching and assessment is good” while another said, “The quick feedback on practice tasks is good.” This seems particularly true of the internally assessed standards. A warning was made by one teacher that feedback “could lead to test orientation and complacency if students do well at an early assessment”.

One science teacher who said he used to think of formative assessment as a “practice test” now recognises that more attention needs to be paid to students’ thinking. He felt that teachers do not do enough probing of Year 11 students’ understanding early in a unit of work because they still “have coverage at the front of their minds”. Yet he said that “if it’s good enough for Year 9, it’s good enough for Year 11”. In his view teachers will devise better strategies for giving students more feedback about their learning as they become more familiar with NCEA requirements, especially with respect to achieving at merit and excellence levels.

Impact of assessment for qualifications on curriculum coverage

The “chunking” of mathematics to match the achievement standards is seen as an issue in all nine mathematics schools. This chunking is problematic when it leads to a lack of linkages between the strands — variously called “Compartmentalisation” or “lack of integration”. One teacher suggested this is a radical change and commented “There is less time to explore curriculum links. This is accentuated by achievement standards but was already evident in unit standards. School Certificate was almost too much the opposite.” However, another teacher saw this as merely continuing an already established trend, saying “Compartmentalisation is higher under achievement and unit standards. It was already there in School Certificate.”

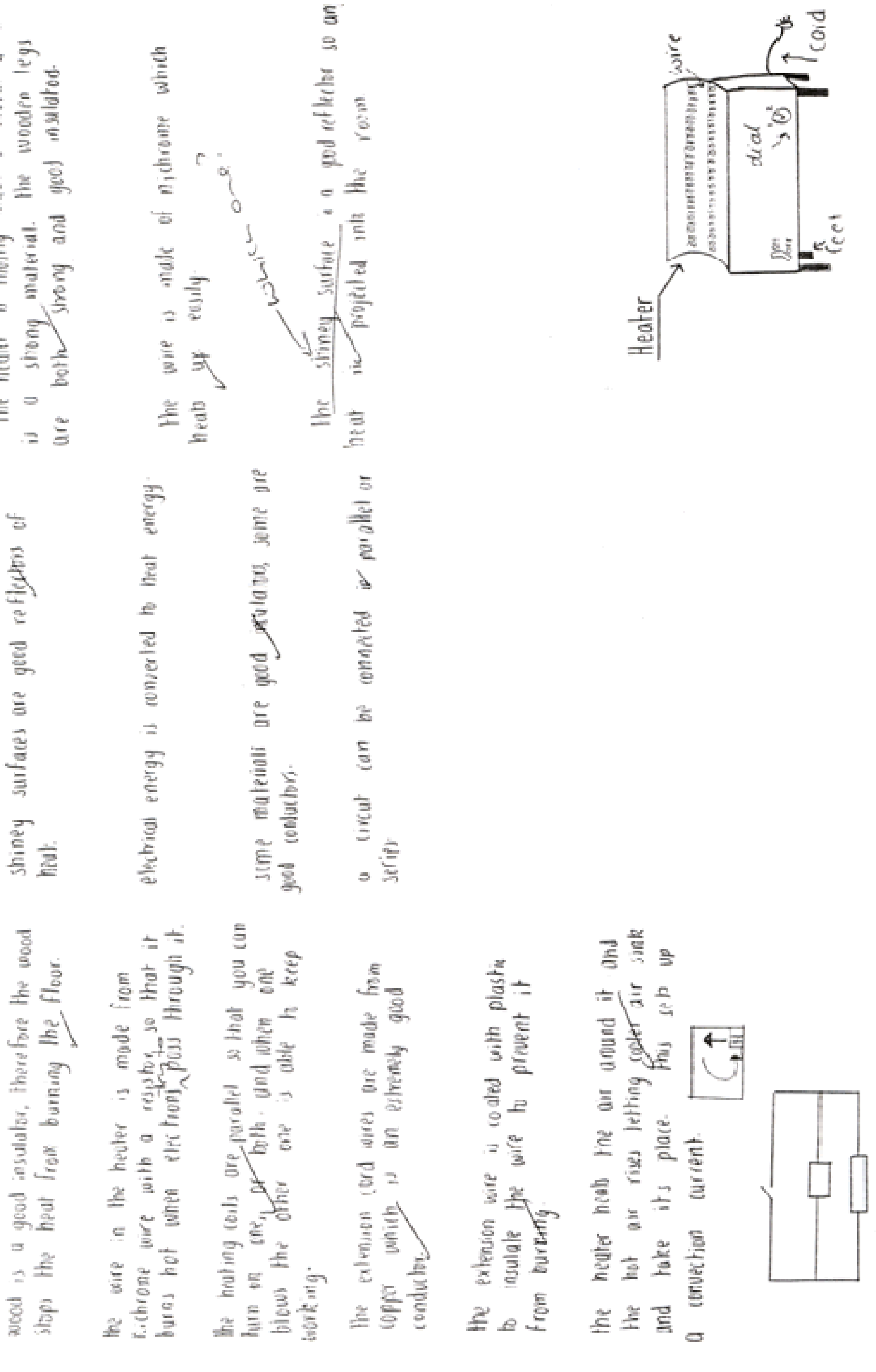
Six mathematics teachers described a “narrowing down” of the curriculum. One believed that this was short-term only and would broaden out. Another thought the narrowing related to the teachers’ ability to more easily predict what would be included in examination questions now that standards are specified, whereas previously students needed a wider preparation to cover what might be in the examination. One felt comfortable with the narrowing because they see opportunities for more in-depth teaching. Two teachers noted that some material has been moved up to level 2, specifically trigonometric functions, the sine and cosine rule, and other non-quadratic functions. One saw this as “a systematic “dumbing down” that was already occurring and NCEA is continuing this”. One teacher said “The statistics assessment is very narrow”, and another commented “There is little difference between Year 9 and 11 statistics. Statistics gets left to the end.”

One mathematics teacher thought “merging topics could be done in teaching to get integration between curriculum areas”. A science teacher who made a similar comment noted that the taught curriculum now more closely matches the intention of *Science in the New Zealand Curriculum* because teachers are paying more attention to the integrating strands (which are assessed via achievement standards 1.1 and, where it is used, 1.2). Another science teacher described the process her team has developed to explicitly teach students how to demonstrate links between science and technology in their research task, as specified by the first integrating strand of *SNZC*.

An example of student work that demonstrates this approach is included as Figure 11. Students summarise the key science ideas for a selected technology on the left hand side of the page. They summarise the technological ideas on the right hand side and write explanations in the centre that link the two sets of ideas together. The format makes explicit the requirement to link scientific and technological ideas — a link that is specified in science achievement standard 1.2 (the research standard).

It was clear from comments at various parts of the interviews that teachers think students have an increased requirement to *communicate* mathematical ideas. Similarly, most science teachers see a greater emphasis on the development and communication of investigative skills in science. Both could be interpreted as testing ideas *about* mathematics or science — that is, wider aspects of literacy — as well as in mathematics or science, if the relevant parts of each curriculum were to be given sufficient emphasis.

Figure 11 **Example of one student's work for science achievement standard 1.2**



Balancing curriculum "coverage" and deeper aspects of learning

In Section Three we reported that the biggest positive shift registered by this research, for both mathematics and science teachers, was for descriptor 11 — "ensuring *assessment* incorporates a range of levels and/or types of thinking". The necessity to make this shift appears to have been brought to teachers' attention by the challenge of preparing students to demonstrate merit and excellence in ways that are now made explicit for the various achievement standards. This change involves the re-balancing of a number of factors and seems to us to be the most complex and — potentially — the most powerful, of the changes we can report. However there is a set of tensions to resolve if these benefits are to be realised.

Identifying the tensions

One issue mentioned by five mathematics teachers was that mathematical processes of problem solving and logic and reasoning were being used less now. There seemed to be two different types of justification for this comment:

- Two teachers believed that mathematical processes are now assessed in separate "excellence questions" in the examination. (However, as responses summarised in Section Four showed, "average" students may not be encouraged to prepare to answer such questions.)
- Another mathematics teacher commented "There is no achievement standard in process skills. There shouldn't be less emphasis on these because it should come in 'merit' or 'excellence', but they are missing." In this view, the mathematical processes are not now being assessed at all.

Taking a different view, one mathematics teacher said that students now develop "more process (problem solving, logic, etc.) skills. They will use these in later life."

Several sets of tensions we detected in teachers' positions towards the formal assessment of the non-"content" aspects of the curriculum may have contributed to this difference of interpretations of the same set of external examination questions.[12]

The first set of tensions concerns interpretation of, and uncertainty about, the intent and/or design of examination questions. Should skills such as problem solving and critical thinking be assessed via separate identifiable questions, or are they the "value added" aspects of learning that differentiate merit and excellence from achieved as different levels of attainment within nuanced and layered questions?

---

[12] These were set and sat for the first time in the 2002 year. As with any very new practice, there seemed to be some inconsistencies of interpretation across the curriculum. We interviewed teachers mid-way through 2003 so they had not yet seen a second set of external examination questions that might help resolve issues of interpretation.

The latter view has been encouraged by NZQA examiners. They now talk of the need to make holistic judgments about the worth of each student's overall performance for a standard. Perhaps as a result of their long experience of making tick/cross/add-up types of judgments, some teachers clearly still expect to see distinct testing of discrete skills. However others, like the science teacher quoted below, have taken up the challenge of thinking quite differently about the process of making assessment judgments and the implications of this change for learning:

> It [NCEA] gets away from nebulous concepts of 'facts' and compartmentalised learning and moves toward a more holistic/process focused approach. Certainly provides a sharper focus as to what is required and therefore what must be learned. (Comment from one science HOD).

The second set of tensions relates to the necessity to make room in a crowded curriculum for the teaching needed to prepare students to demonstrate their learning more holistically — once this need has been recognised. This challenge involves finding new balances in learning, not just in assessment practice tasks. Pressured content "coverage", for subsequent external examination *recall*, needs to be re-balanced against more carefully paced exploration of topics, in a manner that also emphasises other facets of learning such as skills development. Students need time to learn to *think more critically* and independently, making connections to other parts of their learning and/or application to related issues and contexts. Part of this process of thinking more critically can be related to a growing awareness of the nature of knowledge development in the subject. What does it mean, in a critical sense, to "do" mathematics or science? These questions imply shifts in the balance between content and skills. Some aspects of such a shift do appear to be taking place, although teachers may not wholeheartedly welcome them. For example, Section Three suggested that both mathematics and science teachers may perceive a "narrowing" of the curriculum when they spend more time on teaching for understanding than they have in the past.

Adding another dimension to the shifting balances to be explored in this area, a recent review of research that sought to identify effective pedagogy for raising achievement in science identified increasing students' metacognitive awareness of their learning as an important aspect of effective science teaching (Hipkins, Bolstad, Baker, Jones, Barker, Bell, Coll, Cooper, Forret, France, Haigh, Harlow, and Taylor, 2002). Thinking more critically about their learning requires students to be more aware of their own thinking — that is the balance between the cognitive and the metacognitive must shift. Since teachers cannot actually *do* the students' examination thinking for them, the balance must shift from teacher-directed to more student-directed learning if students are to aspire to achieve at merit and excellence levels. Tellingly, however, mathematics teachers feel they are spending *less* time on rich mathematical tasks, and science teachers feel they are spending *less* time on thinking strategies such as concept mapping.

Furthermore, questions that address these higher-levels of achievement typically require students to demonstrate transfer of knowledge and skills to new questions and contexts. If students are to develop the skills to transfer learning flexibly and confidently, shifts are also implied for the balance between abstracting the learning of concepts and embedding such learning within "real" contexts and integrated topics. However, Section Three reported that aspects of teaching in context are neither prioritised nor often used by most of these teachers.

How are these seemingly contradictory findings to be resolved? We turn now to a discussion of how teachers see their role in preparing students to answer questions at merit and excellence levels. This discussion could help explain some aspects of the tensions and contradictions outlined here.

Developing the focus on deeper levels of thinking: mathematics

Teachers in over half of the mathematics schools mentioned the need for explicit teaching and discussion of how to gain achieve/merit/excellence levels in assessments, using focused discussions and teaching time for this purpose. One mathematics teacher said that "bottom-end students need coaching in how to meet the standard" but worried that "this may become a 'do these steps in order to get achieved' approach which may just train students to get it without understanding". By contrast, three teachers said that giving the "answer only" is no longer sufficient to obtain merit or, particularly, excellence. They focus students' attention on some of the key words and vocabulary that give information on how to answer questions. The descriptors that define the achieve/merit/excellence are interpreted along the lines of Bloom's taxonomy and the teachers are debating questions such as:

> "What do 'explaining', 'describing', 'analysing' mean?"
>
> "What should the answer look like if words like these [describing/explaining/analysing] are in the question?"

One of these teachers practises "quick recall and repeating of answers and explanations" orally, as a way of embedding these skills and ways of answering questions.

Three teachers saw this change of emphasis as leading to a significant change in their teaching. One commented that they could "go deeper into topics and get deeper understanding and also have the opportunity to explore applications". This emphasis on teaching for understanding leads to more effective teaching than just "how to do it". However they did also say that this was possible partly because the content has narrowed, with fewer areas of the curriculum being addressed. One teacher specifically commented on students' increased understanding in algebra where "graphing has changed hugely from just drawing graphs to understanding concepts such as gradient and intercept and what they mean". They went on to comment that algebra needs more emphasis and that the new system allows this. Another teacher commented that the emphasis on explaining rather than just giving closed answers to questions was a significant shift.

Some teachers saw positive changes for the students as learners, especially those who are mathematically able. Three teachers commented on the depth of student thinking now required:

> "Students think more deeply about results in both internals and externals."

> "There is a deeper appreciation of mathematical things."

> "It is good for those who gain deeper understanding."

Not all of the mathematics schools reported these positive shifts. One school reported big differences for teaching units assessed by internal standards *(see* next sub-section) but saw teaching style as largely the same for units assessed by external standards. Teachers in three schools expressed a view that their teaching was largely the same as prior to NCEA, and that the basic teaching methods were not affected because their assessment followed on after the teaching. One of these schools did qualify this by saying that they now integrate practice tasks into their teaching. A fourth teacher thought that teaching was largely the same for top streamed classes, but that for the others it was more focused on passing rather than on understanding. This teacher saw some changes in external assessments including "different emphases, the changes in lesson plans to equip students for exams, and being more rigidly tied to time to ensure all students were prepared for everything".

There are several comments here that suggest that, at least in some schools, only the more able mathematical students are benefiting from the focus on higher-levels of thinking needed to gain merit or excellence passes. This supports the comments made about "average" learners reported in Section Four.

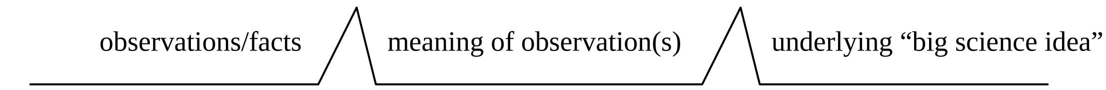
Developing the focus on deeper levels of thinking: science

Seven of the nine science teachers also said they are spending time preparing students to recognise the different types of answers now required to achieve merit and excellence levels. In all of these schools students are given examples of answers that meet these different levels so that they can become more familiar with what is required. One teacher noted that the emphasis on the *quality* of the answer given is "the big change with the NCEA". Another reflected that the types of thinking required to answer questions to excellence level were probably ignored in the past when there was "more focus on the 40 percent kids" [that is — what were perceived as the needs of "average" learners took precedence].

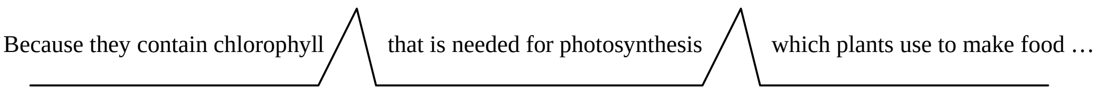
As for some mathematics teachers, strategies may focus on verbs that can be related to Bloom's taxonomy:

- "achieve" questions ask students to: describe, list, draw, name, interpret diagram or table, (described by one teacher as "remembering");
- "merit" questions ask students to: explain how, explain why, give reasons (described by one teacher as "understanding, application"); and
- "excellence" questions ask students to: discuss, apply (described by one teacher as "analysis, synthesis").

Two teachers also described specific strategies that they have developed to help students become more conscious of the depth to which they have shaped answers to questions. Reflecting on her deepening professional expertise in this new teaching and assessment focus, one teacher commented that she is very careful now with her use of the word "explain" in questions. In her experience "explanations" can actually be very complex and the word "discuss" can better reflect the type of answer anticipated for "merit" level (even though it is typically associated with excellence). Rather than focusing on the verb used in each question, students in this school are being shown how to use a visual strategy called a "bridge map". They sketch a shape that helps them to organise their knowledge to make two types of connections that progressively take their answers to the deeper levels of abstraction/wider connections needed to write "excellence" answers:



To illustrate, the teacher used the impromptu question "Why are leaves green?" and quickly illustrated the necessity to make "two links" by writing the following:



Another science teacher explained that she finds as many ways as she can envisage to reformat her stock of existing revision questions and to design new ones, so that students regularly practice responses that address all three levels of thinking. By exposing last year's students to "lots of different formats" she hoped that "nothing would throw them" in last year's external examinations. She noted that the students did indeed achieve at levels above the national average — in her words, "even in the problematic biology questions". She noted that "a huge range of kids achieved".

Most science schools have also begun to teach Year 9 and 10 students to attend to the different qualities of answers required for achieve/merit/excellence. Typically, the teachers in these schools have rewritten existing junior tests in standards-based formats that match the types of questions included in external examinations of science achievement standards. They say they use their professional judgment and knowledge of the typical achievement patterns of the students to determine the levels at which to pitch the standards they set. One teacher commented that this practice is helpful for the learning of both students and teachers because teachers are developing their skills in writing standards-based assessment tasks in "low-stakes" contexts. By contrast, the teacher in one decile one school reflected that he was "not sure if Year 11 kids are ready at this school". He saw the development of higher-level thinking skills as a higher priority at levels 2 and 3 of the NCEA.

Two science teachers commented that expectations for various levels of achievement are now much more transparent. Students have a clear idea of what is expected of them and one teacher noted that it is easier to communicate these expectations to parents as well. By contrast, one teacher said that he does not put an "assessment needs" focus on the merit/excellence types of links but rather uses these links to help students see the "logic and structure of knowledge" in the topic.

A younger teacher who had recently come to teach in New Zealand talked about a strategy she uses routinely to help students develop a deeper understanding of chemistry. While her immediate focus is not on NCEA assessment, this strategy has provided benefits for students' NCEA achievement levels. This teacher described the powerful learning she had experienced during her own teacher education in classes led by Tom Russell, an internationally recognised expert in pedagogy for effective science teaching. In particular, he had made extensive use of the Predict-Observe-Explain (POE) strategy. In her own chemistry classes, this teacher taught her students to use this strategy to work out what they thought about what they expected to happen and they did this before every practical investigation. Initially, this was very time intensive, which caused her to worry whether she would be able to get through the planned course. However, she persevered, and found that her students now use the strategy routinely. Any potential learning challenges become apparent so quickly — both to her and to the student themselves — that they are learning more concepts, in more depth, than she had hoped to achieve. Not only does this strategy serve as a very effective means of formative assessment, it also helps the students to make the sorts of links between theory and practice needed for merit and excellence questions.

A science teacher in another school commented the *timeliness* of changes that build on the conceptually focused research from the 1990s that led to the development of metacognitive strategies such as the POE. She said that the "groundswell from other research and professional development" had "met top-down with the NCEA" and they had "gelled" somewhere in the middle. Noting that science teachers have the advantage of examples such as the Nuffield teaching materials, she added that the Learning in Science Project (LISP) had also prepared the way. LISP findings were, she said, "there, bubbling, waiting for us".

Literacy demands and new types of learning skills

Section Three reported that both mathematics and science teachers give a relatively high priority to the need to develop language and literacy practices associated with each discipline. In at least two cases the teachers' awareness of these issues has been sharpened by participation in school-wide literacy initiatives.

Four mathematics teachers said that their discussions of learning and assessment range far more widely than mathematical issues. They see a need to focus on the literacies that are required in mathematics, especially now with the NCEA model of assessment. For some this means increasing basic literacy requirements. One said that mathematics "now has far more words than numbers". Several mathematics teachers commented on the positive impact of the need to *communicate* mathematical understanding on students' learning. This is seen as stimulating more self awareness that leads students to make linkages between various aspects of their mathematics, to develop a growing statistical literacy, and to become more aware of their wider-ranging skill-base. One teacher noted that NCEA assessment "requires full reasons in the explanations in answers" and said that this helps students to see the purpose of lessons.

One mathematics school has developed a school-wide literacy programme as a consequence of involvement in the *Literacy Leadership in New Zealand Schools* initiative. The mathematics teacher at this school commented that "language is the key" for passing assessments, whether in School Certificate or in NCEA examinations. To increase students' mathematical literacy skills the school started by "identifying questions, the mathematical content, and the skills required". The teachers explored how an understanding of the meaning of questions involves reading and interpreting instructions. They are encouraging students to look at the features of diagrams and graphs, emphasising visual literacies, and linking these with the written text. Issues of vocabulary, and especially mathematical vocabulary, have also been addressed. An example of an analysis of the types of questions students need to be taught to help them address a typical mathematics question in an external examination is shown as Figure 12.

One science school is also involved in the secondary literacy leadership project. At the time of our visit the teachers were preparing to trial "guided reading" literacy strategies in science, with a view to subsequently modelling these strategies for other departments in the school early in the 2004 year. A science teacher at a different school said he now focuses more on the teaching of "academic vocabulary". He pointed out that students who use the word "bugs" in their everyday lives do actually know what bacteria are. They had fewer vocabulary problems in School Certificate where much of the technical language was provided in the question materials and students' answers were shorter. However the requirement to write more open-ended answers means that students may indeed need to know when to say "bacteria" instead of "bugs". For his predominantly Pasifika and Mäori students, the correct, unprompted use of such vocabulary could be a real challenge.

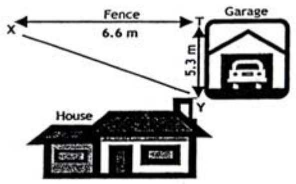
One science teacher said she uses word lists as a type of formative assessment at the start of each unit. Students tick the terms they have heard of and say what they know about these. This beginning teacher says she is very conscious of language and using language strategies such as key word lists because she herself is "relearning" what she knows.

Figure 12 **Drawing students' attention to the wording of instructions**

MANU AND

QUESTION ONE

At Manu's house there is a tree that has to be cut down. It is in the corner of the section labelled with an X.



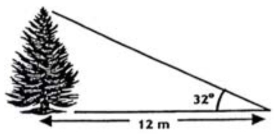
a) What is the tallest tree that could be cut down and not hit the house or the garage if it falls diagonally?

Calculate the length XY.

b) Manu's father wants to know the height of the tree to see whether it is safe to cut it down. He asks Manu to use trigonometry to work out the height of the tree.

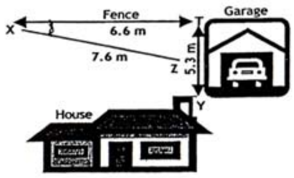
Manu stands 12 m from the tree and measures the angle to the top of the tree.

This angle (angle of elevation) is 32°
Find the height of the tree.



c) When the tree is cut down it falls along the line XZ.

What is the angle the tree makes with the back fence XT?



Questions from NCEA Level 1 Practice test papers

1. Look at the layout of the question.
   IS there a diagram? What information does it have? How does the diagram link with the written text?
2. What are the key words? mathematical / instructions?
3. What information is given?
4. What do you have to do to complete this question?
5. What skills are required?
6. Is there enough information here? What is missing? Do I have to work out something else first? Is there more information in the previous question?
7. Are there any (explicit / implied) instructions about how I should write the answer?

Possible resolutions for some seeming contradictions

In both curriculum areas some teachers have attempted to help their students meet the merit/excellence challenge by teaching more detailed content, seemingly in trying to *anticipate* the links that students might be required to make. Four mathematics teachers said that they need to cover more content in class to prepare students to meet the excellence or merit component of externally assessed standards. One said "In the past you could do 90 percent of a subject, but now the coverage needs to be for excellence. This turns some students off." A science teacher discussed taking a similar stance, saying that teaching time is more pressured now because teachers want to take students to the "discuss level" [i.e. excellence level] in each topic. Until teachers are clearer about the "depth of treatment" that will be tested in examinations, she does not see any easing of this pressure that has teachers "running headlong" in their NCEA classes.

This type of response could help explain the seeming contradiction that some teachers want to help students develop their thinking skills, yet say they are spending less time on some of the very types of activities that might help them achieve this aim. (Strategies that focus on the development of literacy skills seem to be an exception to this trend — at least in some schools.) Where any teacher tries to second-guess the content of questions that are intended to introduce new areas so that students can demonstrate their own ability to think, their curriculum choices will inevitably remain crowded and pressured.

During the interviews we identified several other factors that may deflect teachers from their stated goal of helping students develop higher-level thinking skills. Several science teachers worried that the more open style of questions in external examinations disadvantages students with weak reading and particularly with weak writing skills. Notwithstanding her earlier comment about not "spoon-feeding" one teacher said she had helped students to prepare an essay that they could "tweak" to a specific question in their practice examination because she was so worried about their "poor English". She expressed particular concerns about the open-ended questions for last year's biology achievement standard, saying that "you'd have needed a degree" to answer to the level that appeared to be expected.

During a morning tea break in one school, a different type of dilemma was discussed by the three science teachers present. They see a lost opportunity to make "excellence" type links between different parts of the course because each standard is assessed in a separate suite of questions. In their view, requiring students to build links between different areas of knowledge would be the best way to assess for excellence and they wonder if they are actually "not teaching higher-level skills, but teaching how to answer higher-levels questions in an exam".

Cynicism about teaching as coaching for assessment took a somewhat different focus for some mathematics teachers. Time management is seen as a major issue for teaching of examination techniques in mathematics. One teacher commented that "Time management issues are huge, you need exam techniques much more than School C. It is harder now as there are too many externals (up to six) in a three hour exam." These pressures may have contributed to comments such as the following, made by one of the mathematics teachers: “There is a feeling of teaching kids to pass exams. This is not mathematically sound. Teaching for understanding is better.” Here again we see a tension between teachers’ desire to develop higher-level thinking skills, which they acknowledge students now need to achieve at merit and excellence levels, and their desire to coach and guide students through problems that they believe accompany the changes in the assessment regime.

When students take assessment decisions into their own hands

One teacher commented that students tend to avoid certain parts of the mathematics examination if they perceive these to be too hard. This had affected achievement patterns for the algebra standard in particular, which was unfortunate as this is an important skill for level 2 mathematics. A science teacher made a similar type of comment about being selective about dropping harder topics:

> Some aspects of subjects have been sacrificed to get time for the depth needed in others. [It is] no longer helpful to get ‘40 percent in a weak topic’ – it is perceived by students (and staff) as a waste of time (From one science HOD’s written notes for the interview).

Teachers in six mathematics schools thought that students are too focused on gaining credits which results in them being “more selective of what they learn”. One said “Gaining credits is an obsession” and noted that students are asking “What are easy credits?”, “Is it in the standard?”, “Does it count?”, “Is it achieved, merit, or excellence?” are now commonly asked questions. This latter question is of particular concern as more than one mathematics teacher commented that students often aim for “achieved” only, or aim for the minimum (except the most able students). This teacher believes it is acceptable for students whose realistic maximum is “achieved” but not for others. One teacher commented that “Some students are giving up and this is worse than previously” and for others “They aim for “excellence” initially, but they are delighted with “merit” and will settle for “achieved.” These teachers worry that “achieved” is often not sufficient preparation for further mathematical study and can constitute significantly less than a pass in the previous system.

One science teacher commented that able students are strategically choosing which of their assessments will count towards their final record, dropping internally assessed standards if they do not achieve excellence. Another noted “A disturbing trend among students to value work by “credits” and “instant credits” at that. US are preferred by students to AS, even by better students. [We have] no answers as yet.”

Such unilateral action on the part of individual students has the potential to compound the tensions discussed above. For example, where an “average” student concurs with a teacher’s expectations that it would be a waste of effort to try for anything beyond an “achieve” the benefits of encouraging the development of thinking skills are unlikely to be realised. On the other hand, where teachers are able to really engage their students with deeper and more meaningful learning, the external motivator of credits to be gained may begin to lose some of its allure (whether real or imagined).

Shifting balances for internally assessed course components

Changes and strategies outlined above refer to changes made in learning that prepare students for *external* assessments. We now turn to changes in the balance between content "coverage" and other aspects of learning in parts of mathematics or science courses that are internally assessed. The new imperative to differentiate between achieve, merit, and excellence levels of attainment has also impacted on the assessment of the internally assessed standards. We report next on these changes, and also on the degree to which teachers have shifted the balance between the content and skills aspects of these parts of their courses.

Learning in internally assessed components of mathematics courses

Four mathematics schools said there was now an increased focus on practical aspects of mathematics teaching. One teacher said, "The internals lead to clear changes in what is done in the classroom. They are more hands-on with the emphasis on measuring, collecting data, and designing. We need to search out things that are relevant to the school." Another teacher commented "on doing more teaching outside the classroom", going on to say that the internal standards "force students to do practical work". This teacher noted that this situation has arisen from a combination of influences that include the NCEA, but also the availability of unit standards and the curriculum model specified by *Mathematics in the New Zealand Curriculum*. Another teacher also picked up this theme — saying that the NCEA more accurately reflects the intent of the mathematics curriculum.

One teacher commented how students now "use more sophisticated strategies" in the internally assessed area of statistics and statistical thinking. This school has been working in conjunction with the University of Auckland to develop their teaching and learning in this area. They now use an approach that models a more metacognitive analysis process as students learn to use focusing phrases such as:

"I notice … from the data"

"I wonder …"

The school has referred to this as "the TTRC approach to statistics". (This acronym stands for tools, trials, results, calculations.)

Other comments made by the teachers suggest that the internal standards that assess some practical skills aspects of the Year 11 course may need ongoing refinement. One teacher who had first commented that an increased emphasis on practical work is a good feature of the NCEA initiative went on to lament that "It gives an undue emphasis on skills that are too trite. For example, the measurement standard is more a primary school objective. This gives disproportionate emphasis to aspects which are not central to maths and compromises academic depth." Another teacher also commented that "the measurement internal is trivial". On the other hand, a school with a much lower decile rating found the measurement task useful as they saw this skill as one which their students could achieve.

Not all of the mathematics schools saw a positive shift in the balance between practical aspects of the course and more theoretical work. Teachers in five of the mathematics schools said that they are undertaking fewer investigative, exploratory-style teaching or fun tasks related to the learning *(see* Section Three, which quantifies this effect). Two of these teachers said they can still incorporate the investigative approach for their "top classes", but not for the others because of time pressure in the new system, especially with the weaker students. In their view, the skills developed during more investigative learning are assessed in "excellence" questions — for which only "top" students are being prepared.[13] These teachers say that this change has made mathematics classes blander and "less enjoyable" for both students and teachers. Teachers in two schools were of the opinion that the old School Certificate-type investigations were better than those currently being offered.

Reinforcing these types of decisions, a number of the mathematics schools use a streaming process that leads to a differentiation in the types of learning experiences being offered to students who are perceived to have differing mathematical abilities. In this year's *Learning Curves* report we describe the nature of this streaming in detail for each of the six case study schools (Hipkins et al., forthcoming). As in the Learning Curves study, all mathematics schools offer the lower-ability students MAP-type courses that are primarily unit standards-based, with the possible inclusion of the internally assessed achievement standards. Some schools offered a mix of unit and achievement standards to students of average ability. Teachers of both mathematics and science

[13] One science teacher who likened learning chemistry to learning mathematics also said that she only encouraged "top" students to think about "excellence" type questions and had other students focus on the "achieve" level.

see internal assessment as advantageous for lower-ability students. For the other streams, realistic expectations of likely student achievement (achieved, merit, excellence) have influenced some schools to largely teach mathematics to the level they perceive to be appropriate to that class. For example, one school "teaches to 'achieved' because of the school profile".

Here is evidence that supports the explanations already proposed for the tensions in teaching externally assessed course components. It seems that the *able* students are most likely to be the positive beneficiaries of the new imperative to develop deeper thinking, with the associated prospect of richer types of classroom explorations, even for learning that is assessed internally by achievement standards.[14]

Teachers in six of the mathematics schools were aware that their teaching for internally assessed course components was assessment-led rather than mathematics-led, or that it was aimed at reaching criteria or practising for assessments rather than learning. This was a source of tension for them. One said that "The idea was to make teaching more innovative, but schools teach to the assessments, especially the internals because there is prior knowledge of what these standards and assessments are." When the focus is on outcomes, and the means of assessment is transparent, the pressure to "teach to the test" is obviously hard to resist, even when teachers want to do so.

In Section Six we propose that professional development in the area of formative assessment may help teachers to rethink the way they use the transparency of the internally assessed standards. If a greater proportion of teachers' and students' attention can be shifted from outcomes to learning processes, a wider range of students might benefit from the advantages that can potentially accrue from more extensive use of internal assessment for qualifications.

Learning in internally assessed components of science courses

One science teacher had written notes on her copy of the interview schedule prior to our visit. In response to Q3 she wrote:

> Often we concentrated on learning a body of information and made little effort to extend this to any application or higher-level thinking. Experiments were well designed and carried out but the data collected was often left unprocessed or discussed, the relevance of a graph or table ignored, or we ran out of time. Lesson ends (Science HOD's comment in interview notes).

---

[14] This comment may not apply to those students who are learning in "Mathematics Applied" courses, typically assessed by unit standards. Such courses were not the focus of this study and further research would be needed before comparisons of types of learning opportunities can be made.

The juxtaposition of external/internal assessment situations in this short comment links changes in the degree of critical thinking now expected of students to both types of situations. During the interview, the teacher elaborated on this link, saying that the school has introduced a "thinking skills" curriculum at Year 9 that will be progressively extended to the senior levels. The teachers are "training the kids not to be expect to be spoon-fed" and retraining their own expectations at the same time. She reiterated the change to "less recipe/more analysis" in practical work and said that students are given extensive formative practice because the school has a policy of not reassessing. In her view, students "love" the changed emphasis in their practical work, seeing it as "real science". She also said that the shifts in actual practice are manageable for the teachers because they are small. They are changes in emphasis rather than implementing something completely new. Another teacher also noted that the shift in focus on investigations had already begun prior to the implementation of the NCEA.

Explaining why she gives a very high priority to encouraging students to make their own decisions in practical investigations, another teacher said "How else would they learn to do 1.1?" This teacher also said she "hates the fill-the-boxes format" of the reporting sheet for this achievement standard. In her school, students in all Year 11 classes are encouraged to write more unstructured reports for practice assessments. Another teacher said that he had always encouraged students to make their own decisions during practical investigations. However he sees some NCEA-related changes in that, whereas he had only used relatively open and unstructured short episodes on an opportunist "ad hoc" basis in the past, he is now attempting to plan for such opportunities. He noted that the "what ifs" provide the foundations for capturing students' interest in experimental design and hence help build the skills for practical investigations.

While the above comments indicate a shift in the *focus* of practical investigations, other teachers commented on the *frequency* with which they offer these and the time that they take. For example, one teacher said he had not put as much effort into investigations before saying, "We would have only done one. So now investigation is more valued and takes more class time." Several teachers noted that work for internally assessed standards takes longer to complete than they had anticipated. One said that in 2002 "neither the students nor I were prepared/experienced in the type of assessment and we took on too much". Several schools have rationalised the content areas covered this year so that the practical aspects can be addressed more appropriately. One teacher who thinks that internally assessed components are taking a disproportionate amount of time expects this situation to ease as the NCEA "beds in".

A teacher in one of the smaller schools said she was "between a rock and a hard place" in trying to accommodate the principal's call for credit reduction in all courses. She is trying to find a balance between investigations, that offer a tangible achievement of credits to less able students, and externally assessed topics that provide the "groundwork to carry on" to higher-levels of the subject.

Like their mathematics counterparts, some science teachers commented on the *manageability* of internally assessed science investigations *(see* Section Four). Several schools allow students to plan and carry out their investigations in small groups, but require them to write up their work individually in formal examination-like conditions. By contrast, two teachers commented on the importance of having a really good technician to assist with preparation of materials for investigative tasks that are carried out as part of the mid-year examination cycle. In both schools, all materials must be ready for all Year 11 and 12 students to carry out their investigations under examination conditions within a relatively short time frame. One teacher noted that "Preparation begins weeks before, while normal work goes on."

When they assess achievement standard 1.1 (the internally assessed science investigation) science teachers do not appear to be differentiating between learning experiences offered to students of differing ability levels to the same extent as their mathematics colleagues. Indeed some of the most supportive comments about science achievement standard 1.1 were made by the two teachers in decile one schools, where overall achievement levels have been historically lower than in higher decile schools. These apparent differences, between teachers in the two discipline areas, and between science teachers' practices and perceptions for internally and externally assessed achievement standards, bear further investigation.

Balancing teacher-directed and student-directed learning – no change?

To what extent are students being given responsibility for their own learning? Teachers from several mathematics schools said that the NCEA assessment regime is making students more independent, self-regulated learners. Several mathematics teachers noted that clarity of criteria helps students to focus their own learning with comments such as:

> "The students know what to do to pass and are more focused on this."

> "Students are taking control of their learning more."

> "Students have more confidence to stretch themselves."

> "Students can work at their own level."

One teacher commented "There is more questioning of the significance of what the teacher says." Another teacher also noted this, but also said that doing so in class was not easy, and so students are encouraged to make appointments with the teacher out of class time. Here it seems the imperative to cover content may be crowding out valuable opportunities for discussions that could lead to the development of students' metacognitive skills.

Five mathematics schools mentioned the use of "I can do…" sheets. These are guides that specify what students need to do to obtain achieved, merit, or excellence. One school had revised these to include a study guide with textbook references, and for the externals, a timeline of what to cover. Another school had adapted these sheets for use with unit standards as well. *See* Figure 13 for an example of an "I can do" sheet. These are available from NZAMT.

At one science school, students are also given the opportunity to self-assess their learning, with a focus on self-management. As the example in Figure 14 shows, self-assessment in this case is explicitly linked to progress in assessment of achievement standards for a course in "Biological Science" that mixes biology and science achievement standards.

Figure 13 **An example of an "I can do" sheet for mathematics**

STUDENT SKILLS SHEET : Topic 4

Geometry 1.4 — **Credit, Merit, Excellence**

GEOMETRY: Use geometric techniques to produce a pattern or object — **Credits 2**

Skills I need to know	**Have tried**	**Can do**	**Need help**	**Where can I find this skill?**
• Know vocabulary associated with locus and constructions				Gamma Ch 9
• Construct right angles, parallel and perpendicular lines, circles, simple polygons, medians, mediators, altitudes and angle bisectors				Gamma Ch9 p151-154 MW p294-309
• Draw the path of a point moving equidistant from two points, equidistant from a point and a line				Gamma p155-161
• Use isometric paper to draw 3D representations (more complex than cuboids) • Investigate nets (more complex than cuboids)				Gamma Ch 21 p372
Use properties of :- • reflection, • rotation, • translation(described by a vector)				Gamma Ch 1 (vectors) Gamma Ch 12 p199-207 MW P310-320 P344-349
• Revise Y10 work on enlargement and extend to negative and fractional scale factor				Gamma Ch 12 p 208 MW P331-334
• Draw and describe the results of combining transformations				Gamma Ch 12 p216-220
CEDIT: 'Produce a pattern or object using representations **MERIT : Design and produce a pattern or object using a combination of representations** **EXCELLENCE: Design, produce and critically discuss a pattern or object.**				

NCEA LEVEL 1 2002

Figure 14 **An example of a student self-reflection sheet for science**

Faculty of Science
Individual Summary

Name:

Biological Science

Internal Assessments

Investigation One:	M
Investigation Two:	M
Final Investigation (4):	M
Ecology (3):	M

External Assessments

	Common Test	Practice Exam	External Exam
Genetics (3):			
Human Biology (4):			
Plants (4):	M		
Micro organisms (2):	E		

Self Assessment

I come to class with all my gear	1	2	3	4
I am a good listener	1	2	3	4
I follow instructions	1	2	3	4
I complete all the work set in class	1	2	3	4
I complete my homework	1	2	3	4
I ask and/or answer questions	1	2	3	4
I get involved during practical work	1	2	3	4
My bookwork is tidy and easy to read	1	2	3	4
My behaviour helps my learning	1	2	3	4
What I do in class is helpful to others learning	1	2	3	4
I work well with others	1	2	3	4
I can stick to a study plan	1	2	3	4
I find my own work to do when no homework is set	1	2	3	4
I am satisfied with my progress	1	2	3	4

1 - Never 2 - Sometimes 3 - Usually 4 - Always

Strategies

I am going to employ the following strategies in order to achieve my goal marks in the examinations...

While not explicitly linked to the NCEA initiative, the science teacher who uses the Predict-Observe-Explain strategy extensively also actively encourages students to reflect on the learning strategies that work best for them as individuals *(see* example in Figure 15). This illustrates a different aspect of self-management — developing an awareness of personal metacognitive strengths. However, in all three examples the focus is on *monitoring* learning progress, while decisions about the processes of learning appear to remain essentially teacher-directed.

One finding from Section Three is interesting in this context. Descriptor 10 (offering students opportunities to have input into their own assessment) was prioritised last of all 19 descriptors by both mathematics and science teachers. Science teachers also rated it the least frequent of their classroom practices, both pre- and post-NCEA and mathematics teachers rated it second to last, again with no pre/post change.

It may be that this emphatic rejection of student-directed input into assessment — beyond some responsibility for self-management — is another manifestation of the pressure teachers feel to ensure their students do well in high-stakes assessments for qualifications. Although they cannot actually sit the students' assessments for them, they do everything possible to make sure all bases are covered, and indeed may well be seen by students (and parents) to be negligent if they do not do so. The use of formative assessment to monitor *assessment* rather than to inform next learning steps is also likely to mitigate against valuing students' own input into their *learning* progress.

Balancing learning in meaningful contexts with acontextual "skill and drill" learning

Both *Mathematics in the New Zealand Curriculum* and *Science in the New Zealand Curriculum* model the teaching of topics in contexts that relate to students' lives beyond school. Teaching "in context" can be balanced against teaching concepts and skills as abstractions separate from contexts linked to students' lives. In Section Three we reported that the teachers in both subject areas rated the four descriptors for practices in this area as a fairly low priority in their teaching. Reasons for this may well be interwoven with those just outlined for not valuing student-directed learning approaches.

One teacher who had been using contexts for her science teaching said she was no longer doing so post-NCEA saying, "I've just not found ways to use contexts still" and they have been "squeezed out by other needs". She did however say that "excellence" level questions could potentially bring contexts back and she noted that this had happened when she was teaching genetics — an area in which she feels confident of her own expertise. Another science teacher said that the use of contexts can "cloud an issue" and take time from "thinking and linking".

Figure 15 **A self-reflection sheet on preferred method of learning in science**

Year 11 (Form 5) Science **Learning Checklist**

Enjoyment Ratings	Learning Ratings
5 – My favourite!	5 – A very effective way for me to learn.
4 – I like it.	4 – A good way for me to learn.
3 – Not great, not terrible.	3 – An okay way for me to learn.
2 – I don't like it.	2 – Not a very good way for me to learn.
1 – My least favourite.	1 – A very difficult way for me to learn.

Learning Activity	Enjoyment Rating	Learning Rating	WHY This is My Learning Rating
1) P.O.E.			
2) Lab / Practical			
3) Classroom Discussion			
4) Jumbled Notes			
5) Notes / Examples on the Board			
6) Group Work			
7) Think-Pair-Share			
8) Picture Dictation			
9) Solving Problems from the Textbook / Workbook			
10) Worksheets			
11) Vocab List / Word Wall			

A third science teacher who rated the use of contextually relevant stimulus materials as the "highest of priorities" commented on how hard this is to do in practice because his cultural background is very different from those of his students. Being relatively new to his decile one school, he is still learning what will engage his students' interest. He reflected that there is an "element of risk" that students will reject the teacher's investment in the resource development. However, in his experience, if the resource looks good, and they can see a good reason for the learning, students are likely to engage. This teacher noted that the purchase of a digital camera had allowed him to put familiar images into student assignments — much to both his and their enjoyment.

Two mathematics teachers commented on the difficulty of sourcing good contextual tasks *(see* moderation discussion – p. 82). However their perception that they are now using fewer rich tasks than pre-NCEA (*see* Section Three) suggests that, even if they were doing so pre-NCEA, they are less likely to be doing so now.

In seeking explanations for the small amount of evidence of negative change reported here, it is worth briefly outlining the origins of the idea that teaching should be embedded in meaningful contexts. This idea originated as advocacy — it was a response to findings from constructivist research. While some successes have been registered *(see*, for example, the findings of Hipkins et al., 2002) the difficulties of using contexts for teaching are now being acknowledged, especially as student groups become increasingly diverse and assumptions about what might engage them can be problematic (Ninnes and Burnett, 2001). Teachers were not themselves taught in context and may lack the pedagogical content knowledge needed to do so effectively (Hipkins and Arcus, 1997). Conflation of assessment in context with learning in context compounds all these issues (Hipkins, 1997, Boaler 2003). Given these difficulties it seems unlikely that any benefits that might be perceived for teaching in context would outweigh considerations of the difficulties of actually doing so, especially in situations where students' achievement will be subjected to high-stakes scrutiny.

The most recent School Certificate science examinations did use contexts to assess students' learning. Concepts from different strands of the curriculum were assessed within the same question. Although not specifically mentioned by any teacher in this study, it is possible that the curriculum "chunking" they perceive also mitigates against the use of contexts for learning. Whereas learning and assessment events are related to individual achievement standards, the use of real contexts would require learning to range more "messily" across the range of curriculum areas.

Balancing internally and externally assessed course components

The preliminary questions that shaped this research asked whether the availability of internally assessed achievement standards has impacted on various aspects of teaching and learning in Year 11 mathematics and science. The question itself presupposes that teachers are actually including the internally assessed achievement standards in their assessment programme. Accordingly, the internal/external standards mix is another balance on which we report.

Although schools have made some similar *types* of decisions we have not found any one potentially generalisable trend. Our case study schools have each made a range of curriculum decisions that interact to produce variations in the balance of externally and internally assessed credits in the courses they offer to different groups of students. Most of the nine mathematics schools have dropped or are considering dropping some of the achievement standards from their Year 11 courses. All of the nine science schools have selectively dropped one or more achievement standard from their courses.

Selective use of internally assessed standards in mathematics

There are three available internally assessed level 1 achievement standards in mathematics. Together they offer 9 credits from a possible total of 24 — that is just 37.5 percent of the available credits if the full suite of achievement standards is used to shape the course. However, selective replacement of externally assessed achievement standards with equivalent internally assessed unit standards can alter this balance one way, as could the selective exclusion of externally assessed standards if a decision is taken to reduce total course coverage.

Most mathematics schools are including the internally assessed achievement standards in their courses but if one is dropped it is most likely to be the geometry standard. One teacher commented that their school gives all three internally assessed achievement standards more emphasis because the students see that they can achieve when assessed this way. However, two schools have already dropped the internally assessed geometry standard (1.4) and another school is considering doing so. One school is offering geometry 1.4 in Year 10 with a chance for reassessment in Year 11. The teachers in this school believe that the students are easily able to cope with this, as “the skills required are not high”. At least one other school is consider offering some of the internally assessed achievement standards in Year 10 next year. Another school has considered this but does not want it to interfere with their existing Year 9–10 mathematics programme, which is seen as being very effective.

While unit standards tend to be more frequently used to assess the learning of lower-achieving mathematics students[15] some individual unit standards are also being substituted for particular externally assessed achievement standards in courses for higher-achieving students. For example, one school has done this for level 2 trigonometry, and is considering doing so for level 1 statistical investigations in 2004. Selective replacement of achievement standards by equivalent unit standards in “academic” mathematics courses has also taken place in the schools in the *Learning Curves* study *(see* Hipkins et. al., forthcoming, for details of these substitutions).

[15] Such courses are typically given names such as “Mathematics Applied” (MAP). *See* Hipkins et al., forthcoming, for descriptions of the type of content included in such courses.

Internal assessment via unit standards is seen as “allowing a better focus” for low-achieving students. Seven of the mathematics schools commented on the extra flexibility they provide. One teacher said, “It is good to develop courses and pathways that are more meaningful and appropriate to students’ needs.” Some of the students in this school are gaining pretrade credits as part of their Mathematics Applied (MAP) course. Schools now had the opportunity to design courses with a mix of unit and achievement standards. One teacher said “We can tailor courses for individuals, building on what they know, with a flexible number of credits, and a focus on achievable credits.”

Notwithstanding these advantages, the relationship between unit and achievement standards caused comment. One teacher said “The credit value of achievement standards does not reflect the (extra) amount of work required”, and another that “Unit standards are not of equivalent worth to achievement standards. Generally they are easier … so there is no comparability.” One teacher commented that “Students don't see unit standards to be as important as the exam”, while another commented that unit standards are seen as “an alternative to ‘real’ maths”.

Selective use of externally assessed standards in mathematics

One school has dropped, and another is considering dropping probability (AS 1.6).

Two schools are not offering students the opportunity to be externally assessed for geometry (AS 1.9). One teacher lamented the geometry decision, saying “Geometry is not getting a good deal. Mid-level kids like it and can do it because it is structured, diagrammatic and more concrete. It is the first to go.” In another school with mainly lower-achieving students the interviewed teacher also reflected that geometry was both useful and liked by the students who could achieve at it.

The positive side of dropping some externally assessed achievement standards is that it gives more room for the remaining standards. One school said that they “now have more time for statistics and algebra, and are doing probability better”.

Selective use of internally assessed standards in science

Whereas there are three internally assessed level 1 achievement standards in mathematics, there are only two in science. Together they are worth a total of 6 of the potentially available 26 credits — just 23 percent if the full suite of standards is used to assess a course. All nine schools in the sample offer the science investigation standard (AS 1.1). Some of them effectively offer it twice because they also use the equivalent achievement standard from one of the three separate disciplines (biology, chemistry, physics). In this way, the proportion of available internally assessed credits is increased in these schools.

The research standard (AS 1.2) is only offered at five of the nine schools. Teachers who have chosen not to assess research skills in science typically see this as unnecessary duplication of other parts of the curriculum saying that research has been “done to death” or asking “how many times can you ask someone to research?" We found the same tendency to drop research in the *Learning Curves* study, not just in science but right across the curriculum (Hipkins et al., forthcoming).

Last year one of the larger schools offered a course that assessed the investigation standard twice (as noted above), the research standard twice, and another internally assessed biology standard. For this course, most of the assessment was internal and was achievement standards-based. This is particularly interesting in the light of above findings about the pressure that teachers feel to "cover content". In a course such as this the relative lack of emphasis given to external examinations could well help resolve the tensions generated by the coverage imperative. However, this year the extent of the internal assessment in this innovative course has been scaled back because the school has gone to a semester structure and all courses have been streamlined. Nevertheless, internal assessment still features strongly in these new shorter courses in this school.

As in mathematics, unit standards are most likely to be used to assess the learning of lower-achieving science students. Five of the nine schools offer the full unit standards course developed by the NZASE to at least one class, with one teacher noting that the materials developed by the NZASE to support the standards are "better than the NCEA exemplars". Science courses for mainstream students at these schools are less likely to include unit standards-based assessment components than are the mathematics courses. However we have already noted that science teachers have a large pool of achievement standards in the various science disciplines from which to draw. They can, should they choose, avoid internally assessed components altogether.[16]

The two lowest decile schools offer courses that are mixtures of unit standards and the internally assessed achievement standards — in the case of the very small area school, this is the only science course available. In one private girls' school no unit standards are offered to *any* students.

Selective use of externally assessed standards in science

Four of the nine schools do not offer earth science, assessed by achievement standard 1.5, and three do not offer astronomy (AS 1.7). Two other schools have considered dropping astronomy. In one the HOD wanted to drop the topic but the other teachers objected. In the second school it was the other way around, although the HOD felt she might soon need to capitulate because of time pressure to get through other components of the course. In every case the creation of additional time for the disciplines seen as foundational to the specialist sciences offered in Years 12 and 13 is used to justify the decision to drop these two topics. It seems that curriculum "narrowing" is a reality for the students in these four schools.

---

[16] While none of the teachers in this study have chosen to do so, they perceive that this happens in "other schools".

Balancing teaching time for internally and externally assessed course components

Three of the mathematics schools felt that aspects of the course assessed internally by achievement standards were taking up proportionately more teaching time than were those aspects assessed externally. Two of these schools believed this was primarily because of the assessment component of these standards, rather than the teaching time involved, with the externals taking less preparation time for assessment. One teacher said "The balance of teaching time is still the same but the time for internal assessment takes too long (writing, administering, marking)." One other school said that the internally assessed course components had a more prominent place with their weaker students (measurement 1.3 and geometry 1.5 in particular), because they could achieve these standards, whereas their stronger students have a greater focus on the externally assessed course components.

Three mathematics schools believed that teaching time was well balanced between the internally and externally assessed course components. Two schools thought that the external standards took proportionately longer.

The science schools all reported spending more time on preparation for internal assessments than they had in the past, with one teacher noting that investigations have to be "taken seriously" now that they are formally assessed. Where all the externally assessed standards for the traditional science disciplines are offered (i.e. in courses that still resemble School Certificate science minus the Planet Earth curriculum component) several schools reported spending between 25 percent and 35 percent of their teaching time on internally assessed components. However as several teachers noted, internally assessed investigation actually "boosts" the time spent on one of the topics that will subsequently be examined externally.

In both mathematics and science some schools are balancing the time spent by using historical data on how time has been allocated to different curriculum topics. Others divide the year into equal slots for each of the standards offered, regardless of differences in credit levels. One mathematics teacher noted that the content of the achievement standard dictated the amount of time spent, with another saying "the content and credit levels don't always match". Six mathematics teachers said that the algebra standards[17] need more time because they lay a good foundation to prepare for level 2, and because "There is more in-depth questioning in them, especially in algebraic graphs" (AS 1.2). Some mathematics teachers try to keep the balance indicated in the mathematics curriculum but six of the nine science teachers in this study are inclined to ignore, or actually ignore, one whole contextual strand of their curriculum (the Planet Earth strand).

---

[17] This emphasis on the foundational importance of algebra also emerged in the *Learning Curves* study this year.

In both subjects strategically timing the assessment of internal standards is a strategy for obtaining balance in the overall programme. Three mathematics schools stated they start off with internally assessed topics, with one saying this helps them formatively "to check if students are at the appropriate level". Another saw it as "Trying to get through internals as quickly as possible and give more time to externals to get coverage for them". Three mathematics schools explicitly stated that they did one internally assessed topic per term to keep a balance between the internal and external achievement standards. Others may well do this also. The smallest science school spends the entire first term on internally assessed course components, to motivate students and "get them off to a good start".

Some schools are assessing selected internally assessed standards in Year 10 to reduce the load in Year 11. In mathematics these are either measurement (AS 1.3) or geometry (AS 1.4) and in science the investigation (AS 1.1) standard. Most schools are now offering practice assessments at Year 10. Where these are summative, the possibility of reassessment at Year 11 is typically offered.

Moderation, teacher learning, and classroom change

In the related Learning Curves study, one principal commented early in 2003 that the imperative for staff to work in departmental teams as they moderated NCEA work had been a "wonderful plus" (Hipkins et al., forthcoming). In this section we have noted a number of tensions that teachers need to resolve as they make decisions related to the implementation of the NCEA in their teaching and learning programmes. These tensions are complex, and making decisions related to them requires finding new balances between a number of inter-related aspects of teaching and assessment practices. To what extent we wondered, have opportunities for working together in collegial teams helped teachers to deal with these tensions and rethink long-established teaching practices?

We asked the teachers we interviewed for this study to comment on the link between moderation, teacher learning, and classroom change. To set the scene we first asked them to describe how moderation is actually carried out in their school/department. A range of processes emerged and these are summarised next.

Methods of moderation – task preparation in mathematics

Moderation begins with seeking consistency in the design of assessment tasks. While examples of tasks are available on the TKI website, these must still be moderated, and in any case, teachers often say they need to amend aspects such as context and language to meet the needs of their students.

Three mathematics schools mentioned sharing resources between-schools. One of the mathematics teachers is the key regional person in establishing a resource-sharing site that is now used by a large number of schools. This teacher said that it is difficult to get uniformity between the excellence components of different questions that are contributed to the site. However the co-ordinator also commented that “There is great sharing between-schools. You can model what another school does. The schools use these [examples] to discuss what is required for assessment.”

Notwithstanding such sharing, finding relevant and easy-to-manage assessment resources is seen as an issue for many mathematics schools. One teacher would like to see “a bank of moderated and trialled assessment tasks covering a wide range of contexts”. Another wanted “more teaching material, especially in statistics and literacy”. Finding tasks with contexts that are meaningful to students was seen as an issue by two of the mathematics teachers. One of them said, “Students don’t have sufficient experience to be able to draw truly authentic experiences that they can relate to and that are mathematically appropriate.” Only two schools explicitly mentioned making their own assessment tasks. Time precluded some others from doing this.

Teachers at two mathematics schools said it was important to ensure teachers were aware of policy and website changes. “There is a need to keep these official website resources updated with any small policy decisions, and that all these decisions should be easily found in one place.” Consequential changes to tasks need to he highlighted so teachers can see them.

Several mathematics teachers described preassessment meetings when tasks designed by one team member are discussed and modified. This process may be accompanied by limited trialling of a new task before its use for formal assessment.

Methods of moderation – task preparation in science

Science teachers were less likely to say they had shared resources between-schools. Several mentioned issues that had arisen for them when they used tasks sourced from the TKI site, only to find that some aspects of these were deemed inappropriate by NZQA moderators after they had used them for internal assessments.

Several science teachers mentioned processes for formal premoderation of tasks within the school before their use. In one school a panel of senior staff with expertise in assessment vets all tasks, with their associated marking schedules, from all curriculum areas before their use for NCEA assessments, thus ensuring a high standard of assessment practice across the curriculum. Teachers from two other schools described a process of collaborative discussion of science tasks and mark schedules once these had been developed by one of the team. One science HOD said the first year teacher in her team writes the best new tasks, having not been “polluted” by the type of assessment and marking that has prevailed in recent years.

While not such an issue at level 1, science teachers face moderation challenges at levels 2–4 if they work in smaller schools where there may be only one specialist teacher in the different discipline areas (biology/chemistry/physics/science).

Methods of moderation – marking internally assessed tasks

Teachers described a range of types of procedures to moderate the internally assessed achievement standards within each school:

Model One – Common marking

One teacher, or a small team of teachers (two or three), do all the marking for a particular achievement standard. Three mathematics schools use this model. Consistency is achieved through the constant interaction of the small group. This is maximised with just one marker, but this approach lacks the synergy of bringing more markers together. One of the schools commented that "people have specific expertise in particular areas of maths and types of assessment and we utilise this". While not used for assessing level 1 science standards, one school described a similar process for assessing level 2 biology internally assessed standards, with individual teachers marking a whole task each.

Model Two – Moderation meetings

This involves each teacher marking their own class's work, followed by departmental discussion with the aim of reaching consistent decisions via a sharing of professional judgments. Three mathematics and four science schools follow this model, with each school using a slightly different process. Two teachers noted that this process is particularly useful for learning together how to make the more holistic types of judgments now required in science.

Model Three – Moderation meetings with cross-marking

This is similar to model two except that after marking their own class, teachers mark a number of scripts from another class (five and six respectively in our sample), typically those that are borderline or have unusual features. This ensures consistency of the application of the marking schedule between markers and is done in conjunction with departmental meetings. Two mathematics schools use this process.

Model Four – The "marking panel"

HODs in two science schools run marking panels in which they check-mark a sample of each teacher's scripts, just as they do when they are running external assessment panels. In a third school a similar, if less formal process is followed, with one teacher marking samples from across all science classes to ensure consistency.

Model Five – Whole-group marking with strip-marking

Only one mathematics school uses this system. The whole department meets together to mark and discuss issues as these arise. Small teams of two to three teachers mark particular parts of all scripts. The schedule can be amended if necessary, and whole-group discussions held. This means that consistency is assured in a way analogous to Model One where all students are marked by two to three teachers working closely together. This school used Model Two for marking tasks assessed by unit standards.

Model Six – Individual accountability

Two science schools have only one teacher of NCEA science so none of the above models would apply. One teacher moderates his judgments with the help of a colleague in another school, visiting on weekends when he can make the long drive. At other times he relies on e-mail communication. The other small school has two science teachers so they can discuss moderation issues as these arise for the NCEA class's work.

The impact of moderation on classroom practice

Teachers in eight of the nine mathematics schools and all of the science schools described ways in which moderation processes have impacted positively on teaching practice in their schools. The one exception in mathematics was the school that uses a marking process in which one teacher marks all the scripts, with no wider discussion between teachers. Common themes were:

Clarifying key concepts and task expectations

One mathematics teacher said "Moderation helps us get a better grasp on the deeper concepts. Specific terms can be emphasised to ensure students are familiar with them and their meanings. For example we emphasise the difference between spread and central tendency."

In a similar vein, a science teacher said "Collegial debate over both assessment tasks and assessment schedules/criteria is healthy [and is] passed on as more focused instructions to students — less ambiguous." This teacher also noted that the team had made "subtle changes in emphasis or clarification of direction for topics repeated from 2002-03". Another science teacher noted that staff in her school have become more aware of writing tasks that include open-ended questions, so that students can answer at levels above "achieved".

A third science teacher mused that after years of experience, the theory needed for School Certificate had been "taught on auto" by teachers in his team. However, the implementation of the NCEA appeared to have unsettled their (largely unexamined?) curriculum assumptions. He said his team was looking more carefully at course schemes and paying more attention to developing process skills, not just content.

Identifying misconceptions

Four teachers in mathematics schools and three teachers in science schools said they use moderation discussions to identify and discuss common errors and misconceptions. One mathematics teacher noted "We have always analysed patterns of results and fed them back into teaching", but as a science teacher commented "This is more structured now." However, as noted by a mathematics teacher, this feedback only affects the next year's teaching and so has little impact on current learning.

Enhancing awareness of teaching for different levels of achievement

A science teacher commented that moderation is "educative — especially about levels of achievement — making that explicit" and said that this had helped the transition to standards-based assessment. She described a process of constant "evolution" of tasks as the teachers work together. One mathematics teacher used the term "reflective marking" and another said pithily: "Going through assessments emphasises what to teach and how to teach it." One science teacher noted that opportunities to discuss standards collegially were not usually available in the past and that teachers are now much more aware of the need to "teach to the standard".

Teachers in several of the mathematics schools say they use moderation to look at the grade boundaries carefully and they may also adapt scoring schedules in response to students' answers. When these processes are used for practice assessment tasks teachers can feed the results back to students to prepare them for the formal assessment. Comments made about "coaching for assessment" included:

> "Give students the standard and spell out what is required for getting A/M/E for that standard", "Teach that 'Justify' does not mean write an essay but 'show working'".

> "Standards are a bit picky so students must be informed of these, e.g., write '$3 million' not '3'". "Leave out the bits that are less likely to be assessed".

> "Considering the type of question that is likely to be asked can increase performance by 15–20 percent".

Some mathematics teachers also reported learning, or not, from between-school moderation processes. One said "Between-school standards are hard to gauge. External moderators are being a bit tough." Another teacher thinks "External moderation is useless. The views of moderators are different. They are not consistent between-schools." One interesting comment was that between-task variation is greater than within-task variation. The corollary of this is that it is not so much how the school moderates internally that matters, but which task they choose to administer that affects both the levels of performance and the consistency of results.

Developing collaborative teacher learning

One mathematics teacher commented that it was good to have input from a variety of schools at the Jumbo Days. For this teacher, the general professional discussions were the best part of these days and some follow-up collaboration was envisaged: "We could have workshops for schools without facilitators, which would force schools to come together. The onus is on schools to find what other schools are doing. Clusters of schools help this. You need to know how you compare with other similar schools."

One science teacher reflected that moderation has made staff more open to acknowledging their own mistakes to each other, and said that this had been "really important for our learning". Another said "More heads are always better than one. Other teachers often have new ideas and different resources and there is an opportunity to reduce workloads here." One of the less-experienced science teachers has found moderation helpful for his professional growth saying "Some-one else's call is always interesting, especially when explained. It exposes your weaknesses and over emphases." A fourth science teacher commented that collegial moderation could be seen as very threatening at first if a teacher is not used to it and feels that their assessment is being used as an indicator of their success as a teacher. However he also noted that one teacher to whom this applied, who had newly arrived from overseas, was "shifting fast".

Reflecting the complexity of change in schools, one science teacher said he had noticed changes in his teaching but he was not sure how much to attribute these to the NCEA implementation and how much to the school-wide focus on study skills, co-operative learning and so on. He noted that "these things are not new — we have had them before" but said that organised weekly professional development serves as a reminder and provides the motivation to try new ideas and report back. He did say, however, that the Jumbo Days for NCEA implementation have moved in similar directions, and are now keeping "learning as a focus".

Communicating student achievements to the wider school community

One science teacher noted that the visit of the external moderator had prompted him to begin a process whereby his students use their collected, moderated work to present portfolios that display their achievements for their parents to view.

Summary of key findings in Section Five

The weight of responsibility that teachers feel when their students are assessed for qualifications may mitigate against some changes teachers would otherwise like to make in their classroom practices. For some teachers, this tension is exacerbated by internal assessment for qualifications because an explicit focus on preparation for such assessments is now spread through the year rather than being focused on a one-off end-of-year event. On the positive side, more time and attention are now being given to the practical course components that are internally assessed.

The achievement standards have made teachers more aware of differences in levels of student achievement for a range of aspects of learning. They are more focused on teaching that could assist at least some students to develop the skills needed to demonstrate learning for merit and excellence. There are some tensions between this teaching goal and current practice including:

- learning for higher-level achievement may be restricted to more able students, with "average" students expected to concentrate on gaining achieve level passes;
- some teachers have responded to the likelihood that unfamiliar contexts will be used for excellence components of examination questions by trying to cover more content;
- formative assessment is often strongly associated with holding trial runs for summative assessment rather than with extending and deepening learning;
- teachers are reluctant to promote the use of self-regulated learning strategies except where these simply monitor students' self-management of their overall progress; and
- some strategies that could help develop students' thinking skills are being used less rather than more. Examples are the use of rich mathematical tasks, and strategies such as concept mapping in science. There may have also been a decline in the use of meaningful contexts for learning.

These contradictory changes appear to be related to the imperative that teachers feel to "cover" the curriculum, and/or the time that preparation for new and unfamiliar types of assessments is taking from the overall learning programme.

There is a strong focus on the development of literacy skills in both curriculum areas. While the secondary schools literacy initiative has contributed to this in at least some of the case study schools, teachers' awareness of literacy challenges appears to have been raised by the types of examination questions now being used in external standards-based assessments.

There has been some reshaping of the curriculum content offered to students, usually with the aim of reducing the "coverage" pressure and maximising students' chances of gaining credits for those assessment standards selected to assess their learning. For both mathematics and science, a number of schools have dropped, or are considering dropping, at least one internally assessed standard. Some schools have also selectively dropped externally assessed standards in both curriculum areas. Since there is no one pattern to these changes, the proportion of internally and externally assessed credits that students can gain in each subject varies between subjects and between-schools.

Learning programmes are typically organised around the various achievement standards, which has led to perceptions that the curriculum is "chunked". An exception is the integration of science achievement standard 1.1 within a topic that is externally assessed. (This also happens to some extent with science AS 1.2 — but only five of the study schools offer this standard.)

In most schools, preparation for standards-based assessment is now beginning at Years 9 and 10. Conversion of existing assessments to this format has increased workloads in the short-term but teachers appreciate the chance to experiment in a "low-stakes" context. Some internal achievement standards are being assessed at Year 10 to ease curriculum coverage pressures at Year 11.

Students are experiencing very different reassessment practices in different schools. Some receive considerable help and support to improve their achievement levels before reassessment. At the other extreme, some students get one shot at an internally assessed standard and no reassessment is offered.

Within-school moderation meetings and procedures have provided positive opportunities for professional learning. Some teachers have welcomed the strong focus on student learning and achievement, and say they have used this to rethink aspects of their teaching practice.

Section Six

Finding a way forward

This report has described the perceptions of changes in their teaching, in relation to the introduction of the NCEA, held by teachers in nine mathematics schools and nine science schools. These teachers were nominated by their professional colleagues as people who are coping well with NCEA implementation, and thinking about interesting ways to use standards-based assessment in their teaching and learning programmes.

These are mostly highly experienced teachers, although the voices of three science teachers who are relatively new to the profession have also been heard. Between them, they teach in a wide range of schools — single sex and co-educational, high decile and low, city, town and rural, large and small. This range of school types has doubtless contributed to the differences in some of their perspectives that were discussed in Section Five.

Some of these teachers welcomed the change to the new assessment regime. Some of them did not, saying that they were doing their best to “make it work” because they had to do so for the sake of their pupils. However, none of them would like to “go backwards” now that the changes are well underway. One mathematics teacher noted that teachers are less anxious about the system now, saying “Don’t throw it out but refine it. Schools need to think smarter, simpler, easier to manage.” Another mathematics teacher said, “It will improve through time. Teachers are now more realistic and will make it work.”

In the spirit of “thinking smarter” and refining aspects of the assessment initiative we offer some comments about the changes we have described, and we explore the implications of some important issues that have emerged during the course of the research. First, however, we revisit the initial research questions.

Answering the initial research questions

Question One

> As a result of the introduction of the internally assessed achievement standards, are there identifiable changes in the content, structure, and balance within programmes for Maths and Sciences?

The answer to this "triple-banger" question is a qualified "yes" on all counts.

- Courses in both subjects have changed in content. Skills are now emphasised more than was previously the case. The "room" to do so appears to have been created by selective dropping of some topics from the taught curriculum.
- The structure of courses has changed to become more aligned with the discipline-specific divisions created by the various achievement standards. In some cases the selective mixing of science achievement standards from the various discipline areas has created courses that focus on one particular area (e.g., "biological sciences") or on a new type of discipline area such as environmental education. There appears to be some differentiation of the structure and content of courses offered to more and less able students, especially in mathematics.
- Many of the balances have shifted, although some shifts may well have compensated for others, resulting in little overall change. An example is the increased interest in deepening levels of thinking rather than teaching for "content coverage". This shift in balance may have been countered where teachers are also using fewer rich mathematical tasks, or fewer strategies that develop students' metacognitive awareness in science.

Question Two

> Are there identifiable changes in teaching and learning styles used within Maths and Science programmes that support the development of practical skills, or that allow teachers to address students' attitudes and values relevant to this subject area?

Evidence for these types of changes is much less clear cut. Rather than changing actual teaching styles for developing students' practical skills, most teachers in both subjects appear to be "doing much the same but doing more of it". That is, more time is being spent on developing these skills, but not through the use of necessarily different approaches.

Changes that would allow teachers to address students' attitudes and values were not evident. Descriptors of practice that focused on working with students' interests, in contexts relevant to their lives, or that allowed them to take some responsibility for their own assessment decisions, were neither prioritised nor practised. This situation appears to be unchanged from pre-NCEA practice. On the other hand, some students are taking unilateral action to decide which assessments they value and will attempt. Teachers are as yet unsure how to cope with this change.

Question Three

> What case study/best practice lessons can be drawn from 2002/2003 practice in Maths and Science programmes for NCEA level 1?

Because changes in teaching and learning styles were not particularly evident, this question cannot be answered in the manner that was initially intended. However, many of the tensions we identified in Section Five could be addressed as the NCEA continues to evolve. This seems to be a much more productive avenue for exploring best practice lessons. We do so next.

Promoting higher-level thinking

In both mathematics and science classes, teachers are paying more attention to the development of students' thinking skills. The need to produce evidence of learning that requires more than the recall of information, or the straightforward application of algorithms, has prompted this change. Where students aspire to merit or excellence, teachers recognise that they cannot neglect higher-order thinking. The balance between recall and thinking skills such as application, critique, and/or synthesis has already shifted and potentially could shift further. Expectations are powerful. At this stage it seems this shift is most likely for students identified as able, raising the question of how to encourage teachers — and students — to aspire to higher-levels of learning for all students.

While all the teachers we interviewed are conscious of this shift to valuing critical thinking, they are at different stages in the development of new or refined teaching strategies. One science teacher described routinely using the POE — a strategy developed and refined during the 1990s as a response to the challenges posed by findings from constructivist research. Another mentioned using strategies modelled in the recently developed Royal Society resource *Entering the debate on genetic modification by developing a critical thinking response.*[18] Two science teachers have designed their own strategies, including the "bridge" diagram described in Section Five.

---

[18] The resource models critical thinking strategies that are more often used in English, showing students how to detect bias in newspaper articles and so on. It has been sent to all secondary schools in New Zealand.

It was very apparent that teachers who have not yet developed such strategies are keenly interested in doing so. Should a resource that describes good strategies, and preferably includes a collection of trialled exemplars, be made available, it seems very likely to us that teachers would welcome it and use it immediately. This could be an appropriate focus for further professional development related to the NCEA implementation.

We offer a small cautionary note about this shift — one related to the much discussed "knowledge economy". One science HOD in the *Learning Curves* project observed that critical thinking is being developed at the expense of creative thinking and/or practical skills[19] at which different sorts of students could take their opportunity to shine (Hipkins et al., forthcoming.) Several of the mathematics teachers in this project lamented what they saw as the loss of opportunities for fun in their teaching programmes, with the freewheeling explorations of rich mathematical activities displaced by the need to prepare students for the new demands of NCEA assessments. Recent international commentary sets these types of concerns in a wider context. A critical analysis of policy documents in the UK showed a disjuncture between the emphasis placed on creative thinking by policy makers in the innovative science/technology area and the emphasis placed on critical thinking by policy makers in science education (Kind, 2003).

Practical and theoretical knowledge and skills, and critical and creative thinking, are *all* important in the development of any new knowledge or innovation. Finding new balances between these quite different types of emphases within mathematics and science could be a subject for exploration and debate. Since teachers appear to shape their level 1 programmes around assessment for qualifications, it may be that some different *types* of standards will ultimately need to be developed.

Balancing validity and manageability of internal assessments

We have noted differences in the balance between internally and externally assessed course components, in different schools, and for students perceived to have different types of learning needs within the same school. Typically, students who are seen to be low achievers are likely to have all or most of their learning internally assessed whereas "able" students may mainly experience external assessments.

However, while teachers may express the view that certain types of assessment "suit" different groups of students, they are also concerned about workload, practicality, and validity issues associated with the internal assessment of achievement standards. The same degree of concern does not seem to attach to the unit standards. This may be in part because they are more familiar

[19] The "intellectualisation" of practical subjects is a related challenge, also discussed in the *Learning Curves* report.

to some teachers. Another reason could be that the achieved/not achieved judgment is more clear cut than the four levels of judgment required for achievement standards.

We found a clear tendency to impose examination conditions on the internal assessment process even when, for example, students are working outdoors carrying out mathematical measuring tasks. This practice impacts substantially on teachers' workloads and can make assessment of practical tasks all but unmanageable. Some science teachers are taking a more pragmatic approach to the actual investigative tasks, although they typically require these to be "written up" under examination conditions, creating difficulties if students are not present for the consecutive number of periods required.

Some teachers appear to believe that the valid demonstration of evidence of learning requires situations in which students cannot "cheat" by profiting from the ideas of their peers, or otherwise gaining clues and prompts from their surroundings. At present these views appear to be reinforced by external moderators' advice and feedback. These beliefs and practices contrast with what assessment expert Paul Black identified as one of the most significant principles on which the NCEA initiative is founded. This principle is:

> That all approaches, from external formal tests, through external assessment of work produced in less formal contexts, to internal assessments conducted by teachers *in the context of normal classroom work*, are to be used as appropriate (Black, 2001, p. 2, emphasis added).

While it is possible to define practical examinations as part of "normal classroom work" we doubt very much that Black would do so, given his comments elsewhere in his report on the NCEA to New Zealand's Ministry of Education. Indeed NZQA's communications to teachers encourage the diversification of internal assessment practices. For example, the "key questions" page on the website states that:

> Skills and knowledge that can be assessed by examination will be included in normal end of year examinations run by NZQA. Most internal assessment involves skills such as giving a speech, making a product, carrying out research or laboratory work (www.nzqa.govt.nz).

As early as 2000, the NZQA Update circular sent to schools gave advice for addressing authenticity concerns when students work in "normal classroom" settings to produce work for internal assessments.

> Authenticity concerns can be addressed in a range of ways and vary according to the nature of the evidence being collected. Strategies used by schools include:

- changing the context of the assessment from year to year
- supervising the research process by including regular checkpoints
- requiring plans, resource material and draft work to be submitted with the final product
- keeping ongoing work on site
- oral questioning to confirm a student's understanding or requiring a repeat performance where there is doubt
- being familiar with or controlling the resources available
- controlling group work by breaking the task into group and individual components
- requiring a signature on an authenticity statement to highlight the issue for both parents and students (NZQA, Update 4, November 2000).

The possibility of using normal classroom tasks was reiterated in Update 8 when the manageability of internal assessment was at issue:

> Teachers can collect evidence of a student's achievement during the teaching and learning process, or through a one-off performance such as an assignment or test (NZQA, Update 8, August 2001).

It seems that this is message is yet to be heard by a number of teachers. Models of good practice may not suffice as a means of addressing the issue. Rather, teachers' assumptions about what can constitute valid evidence of learning may need to be further explored, and addressed via appropriate professional development.

Balancing competing professional development demands

Amongst the many competing priorities for the ongoing evolution of the NCEA, we suggest four reasons to prioritise strategies to address the issue of "authentic" internal assessment in mathematics and science:

1. Teacher workloads are high. The use of actual classroom work for assessment purposes would make internal assessment in mathematics and science, and teacher workloads, more manageable.

2. Casas (2003) critiques the disjuncture between the constructivist approaches encouraged in contemporary mathematics and science teaching[20] and the use of standardised examinations based on behaviourist models of learning. This gives both teachers and students mixed messages about what is valued and can considerably increase stress for all concerned.

3. Boaler[21] (2003) documents a case in which students at a low-income urban American school, with the help of very dedicated teachers, demonstrated huge learning gains in mathematics, and outperformed students at more privileged schools when assessed with authentic tasks. However they failed to demonstrate their learning in a standardised testing regime because they could not adjust to the formal context and/or the unfamiliar language of the tasks. In this context, learning mathematics "to pass the test" was demonstrably rewarded over learning to develop rich mathematical thinking, and the students who had developed such thinking skills subsequently developed a negative view of themselves as successful learners of mathematics. Such findings are challenging, given New Zealand's policy focus of raising achievement for all students, and of preparing students to be lifelong learners.

4. Gilbert (2003) addresses the different types of challenges posed for education by the development of the idea of a "knowledge economy". Amongst other aspects, she notes that innovators work in teams, so that between them the necessary skills for the task will be available. She suggests that we need to find ways to assess the learning of teams of students as they work together on authentic tasks, creating knowledge that is genuinely new for them.

Developing teachers' formative assessment skills

In his report on the NCEA initiative, Black cautions that teachers may not be able to develop sound internal assessment practices if their formative assessment skills and knowledge are weak (Black, 2001). Accordingly, we turn now to some insights into teachers' thinking about formative assessment that emerged during this research.

Section Four describes the practices that the teachers in this study typically referred to as "formative assessment". Whereas experts like Paul Black think of formative assessment as "assessment for learning", these teachers typically described processes that sound more like "assessment for assessment". That is, the feedback teachers give to their students is primarily

---

[20] Such approaches are modelled in both *Mathematics in the New Zealand Curriculum* and *Science in the New Zealand Curriculum.*

[21] An internationally recognised expert in mathematics education.

focused on the summative assessment to come (and hence on their *past* learning) rather than on their learning either "in the moment" or in the immediate future. Teachers do use insights from moderation discussions to highlight learning issues that can be addressed the next time a topic is taught. However, such insights benefit the next student cohort rather than the current group.

Section Three noted that the science teachers in this study value strategies that reveal student thinking (descriptor 6 in Table 2, Section Two) somewhat more than do the mathematics teachers. Yet the mathematics teachers perceive that they use such teaching strategies more often now, and the science teachers think that they use them less.[22] Such strategies can be very effective for developing higher-order thinking skills, but we were shown few examples in the context of our conversations.

A national survey undertaken in relation to the curriculum stocktake investigated the range of strategies that teachers say they use to assess learning in science and mathematics (McGee, Jones, Bishop, Cowie, Hill, Miller, Harlow, Oliver, Tiakiwai, and MacKenzie, 2002; McGee, Jones, Cowie, Hill, Miller, Harlow, and MacKenzie, 2003). Figure 16 (mathematics) and Figure 17 (science) show the tabulted results from this survey for secondary school teachers (Years 9–13) in graphical form. Responses have been ordered from largest to smallest "often" response.[23]

While not specifically directed to *formative* assessment, the responses demonstrate that teachers are aware of, and to varying degrees already use, strategies that can be just as easily used for formative or summative assessment purposes.

The most frequently used assessment strategy is the pretest/post-test combination. The *Shifting Balances* teachers say they use such tests for formative assessment purposes *(see* Section Five) which seems likely to keep the focus on evaluation of the extent of learning that has been completed. Professional development that explores ways to use some of the less popular strategies could help shift the focus to the next learning steps as these are taking place. The surveyed groups made least use of interviews with students. Similarly, self- and peer- assessment were unlikely to be used "always" or "often" *(see* Figures 16 and 17). This pattern of responses is congruent with responses from the *Shifting Balances* teachers, who rated students' input into assessment decisions as a low priority, and something that they almost never did. Yet self-assessment is one of the essential components of self-regulated learning (Zimmerman, 2001) and this has been demonstrated to result in more effective and committed learners (Black and Wiliam, 1998). Since

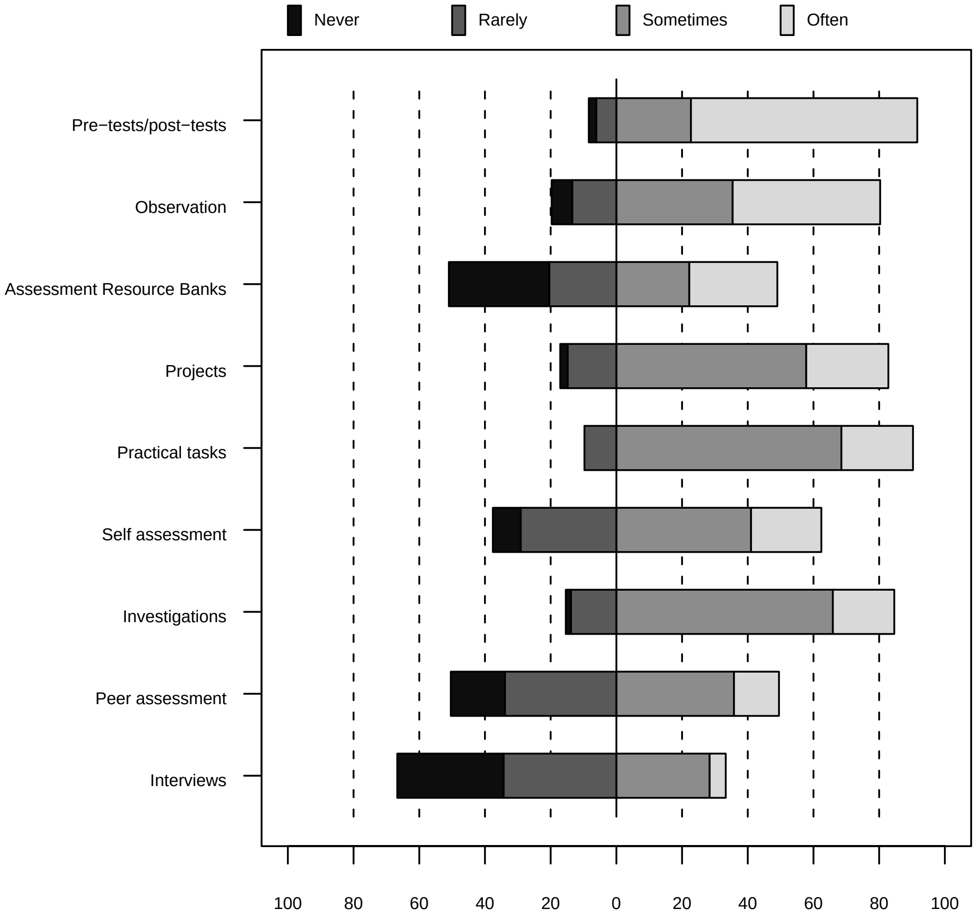
---

[22] However, the differences are marginally statistically significant and would probably be so with a bigger sample.

[23] Science teachers were not given an "always" option. For comparative purposes we collapsed the "always" and "often" categories together on the mathematics graph. There was some change in the wording of the list of options provided from which teachers could select. Results for NEMP tasks have been omitted from both graphs and the numeracy project diagnostic interview has been omitted from the mathematics graph because these instruments were designed for primary school use. (Nevertheless a small number of secondary teachers do use these sometimes or rarely.)

teachers are worried about the impact of the NCEA on student motivation, a professional development focus on self-regulated learning, as well as on other types of formative assessment, could address several of the issues we have raised in this report.

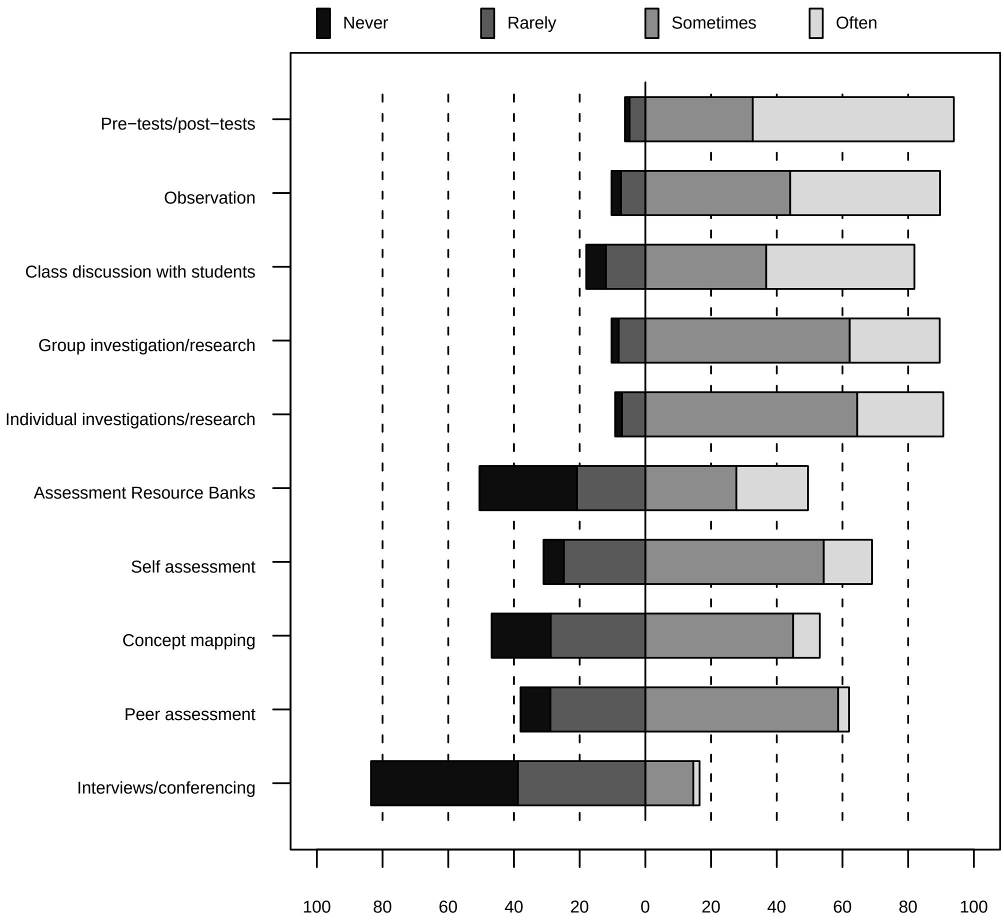
Figure 16 **Secondary school mathematics teachers' self-reported use of assessment strategies**



Note: Response numbers range between 102 and 132 for individual categories.

Data source: (McGee et al., 2002, pp. 94–96). "Often' category shows collapsed data from "always" and "often" categories of originating tables.

Figure 17 **Secondary school science teachers' self-reported use of assessment strategies**



Note: Response numbers range between 101 and 141 for individual categories.

Data source: (McGee et al., 2003, pp. 213–215).

We have already noted that teachers are keenly interested in tasks that will develop the skills their students need to demonstrate merit and excellence in external examinations. It could be that any professional development strategies that address these "thinking skills" issues could also address the internal assessment issues outlined above *and* help to develop teachers' formative assessment skills. By moving the focus to what students know and can do now, and their next learning steps, teachers may more readily recognise when students reach the specified standard. They would also then be able to marshal more evidence to justify this judgment without needing to rely on a specific one-off assessment event. Such strategies might simultaneously address several important issues whilst avoiding the danger that coaching students to demonstrate higher-level thinking in examinations could become, in the words of one science teacher "low-level demonstration of higher-level thinking skills".

Curriculum "coverage" issues

> While the literature in the curriculum field recognises the difficulty in creating meaningful change within current school structures, the majority of innovations and analyses are blind to the bigger and more significant questions surrounding change: Who are the young people in schools? And what, where and how do they learn? If curriculum reform continues to focus upon subjects, teachers, school-based lessons, and the modernist structures of schools that obfuscate difference, meaningful learning and the impact of technology, the reform movement will become more irrelevant in the lives of young people (MacDonald, 2003, p. 147).

At present, the majority of achievement standards developed for level 1 mathematics and science focus on the content associated with these subjects as discrete traditional curriculum areas. There have been advantages of this in that teachers recognise the NCEA initiative as "business pretty much as usual" when considering the curriculum they deliver. This has been reassuring in the face of such substantial assessment change. However a disadvantage has been that teachers worry about the "chunking" of the traditional curriculum. When discrete topics are discretely assessed, teachers may struggle to find ways to build links between the separate units in their teaching programmes. Professional development could help teachers to explore strategies to reduce the compartmentalisation of mathematics, in particular, and to encourage linkages and connections between the standards.

Paradoxically, an advantage of assessing separate topics as separate mini-examinations is that teachers are free, in theory, to build quite new types of curriculum subject mixes. However, in this study, the teachers who have taken advantage of this new flexibility have mostly done so to narrow the scope of the assessed curriculum (and in most cases, seemingly, the taught curriculum). Some science teachers no longer address the Planet Earth components of the curriculum. Some mathematics teachers no longer explore geometry as fully as they might once have done. We did find some evidence of new mixes of discipline areas in science.[24] Some of our schools are planning such innovative courses in the immediate future and one is already offering a wide range of semester-long courses that allow students to "pick and mix" the discipline areas in which they are interested.

During the *Shifting Balances* project, and during this year's *Learning Curves* research, we have noted some anxiety about the specific subject content of pathways that will lead students successfully from school to tertiary study, especially into the universities. Ensuring their students can show an achievement record that allows them to make such transitions is one of the "high-stakes" aspects of assessment for qualifications. At the time of data-gathering for both projects, teachers were feeling an uncertainty about the actual subject components that may be required. This uncertainty will doubtless settle once the NCEA has been implemented at all three levels of the senior secondary school and pathways to tertiary study become clearer. It seems to us that the more widespread creation of new and innovative types of courses such as those mentioned above will not happen until this source of uncertainty is removed.

Research – done to death or centrally important?

MacDonald's critique of traditional curriculum reform in a post-modern world, from which the quote at the start of this sub-section was taken, draws attention to the power of electronic forms of communication in the lives of young people in the developed world. Others describe the power of the multi-modal nature of such learning, in which meaning may be conveyed through visual elements as much if not more than through the more traditional verbal/written components (Johnson and Kress, 2003). Much of what today's students will learn, in the immediate future and throughout their lives, will not be regulated within traditional learning institutions where teachers are the arbiters of authoritative knowledge.

An Australian study recently reported on the attributes of school programmes that successfully support students to develop the attitudes of lifelong learners. Central to the features they identified is the support given to independent research, and the provision of ICT facilities and staff support for students as they develop their information literacy skills (Bryce and Withers, 2003).

Despite the attention researchers are paying to these significant developments, the teachers in this study did not prioritise the use of ICT in their programmes, nor did they say they used such technologies very often *(see* Section Three). Some science teachers are not teaching or assessing

[24] We had sought to work in one school that has developed innovative courses in environmental education that mix social studies and science unit standards, but the principal declined to participate.

students' independent research skills, a trend that we have found in several curriculum areas, including science, in the *Learning Curves* study (Hipkins et al., forthcoming).

In some schools teachers appear to assume that others will cover research, with the danger that no-one will. In at least one instance a school is co-ordinating assessment across the curriculum and research at Year 11 has been assigned to a curriculum area other than science. While obviously preferable to a "hit or miss" approach, this may not be the sound solution that it seems at first glance. No teachers in our sample (or in the *Learning Curves* sample) talked about attributes of research that are unique to the science disciplines and that give a distinct flavour to the type of second-order research students might wish to carry out, for example, in the investigation of a socio-scientific issue.

Considerable scholarly writing on this topic exists. Some researchers address discipline-specific differences in what it means to be "scientific" *(see* for example Shavelson and Towne, 2002). Others have analysed the cognitive tools lay people need to bring to the evaluation of others' scientific research where social issues are involved (for example, Ryder, 2001). Such work could be used to devise research strategies that help develop the higher-order thinking skills that teachers are learning to value. However the ideas are currently in forms and places that are not easily accessible to busy classroom teachers. The shaping of professional development materials and teaching strategies that draw on these rich resources could help teachers to rethink the importance of second-order research for their students' immediate and future learning needs. However the research standard itself would likely need revision so that it assesses more than the "information retrieval" that is currently interpreted by teachers as a cross-curricular skill.

Literacy skills – valued and growing

Professional development initiatives that are currently assisting secondary teachers to address literacy in all areas of the curriculum do appear to be shifting classroom priorities. A number of teachers in this study perceive that they are spending more time on such skills than they did pre-NCEA. While this change clearly relates in part to the wider contexts of school-wide professional development initiatives, it is NCEA-related in that teachers recognise their students need to be able to write more clearly with the new, more open question styles of the externally assessed achievement standards. In this case, issues triggered by the NCEA implementation have amplified interest in other professional development programmes that were already underway. A similar process of amplification was noted in cases where there was a school-wide focus on aspects of professional development related to enhancing classroom practice and improving student achievement.

Building on this research

> Rapid pedagogical change may be easy to secure within one curriculum unit, one classroom, one syndicate or one school but, beyond these, experience suggests that there is rarely sufficient common commitment to educational change for it to take place consistently across a system as a whole. This means that it is very difficult for policy developers to envisage or predict system-wide changes to the nature of teaching and learning (Codd, Brown, Clark, McPherson, O'Neill, O'Neill, Waitere-Ang, and Zepke, 2002).

This small-scale research project provides an in-depth analysis of the dynamics of change in some mathematics and science classrooms as the NCEA implementation at level 1 beds in. We have found a series of inter-related changes. While some of these seem to have readjusted existing balances in classroom practices, with little real change, others have intersected with different types of professional development initiatives underway in schools and positive changes have occurred.

Teachers are currently very interested in strategies that could help students develop the deeper thinking skills to address questions at merit and excellence levels, and this interest could be built on with timely professional development. Although they say they have made changes in their classroom practice to emphasise higher-levels of achievement, these positive changes have the potential to be more effective if other, seemingly compensatory changes, are rethought.

Our research supports Paul Black’s advice that secondary teachers need support to develop stronger and more varied formative assessment practices if the potential benefits of the internally assessed components of the NCEA are to be fully realised. The introduction of practices that support self-regulated learning could also assist teachers to deal constructively with the currently vexed issue of increasing student autonomy in decision making about assessment of various course components.

This research can be used to inform the Ministry's ongoing work in a number of ways. These include:

- potential to monitor ongoing NCEA-related changes if the research is repeated in several years' time, and to use the teacher self-reflection instrument developed in this research for a larger-scale survey of teachers in these or other subjects;
- using the insights into the nature and range of interacting factors that impact on teachers' classroom practices when their students face high-stakes assessment for qualifications;
- informing the focus of any ongoing professional development initiatives that explicitly support the NCEA implementation, including the development of strategies that encourage teachers to revise their expectations of students perceived to be low- or under-achievers;
- auditing the work being carried out in other professional development to identify opportunities to create synergies that will enhance the likelihood of changes in classroom practice taking place;
- informing principals about such opportunities so that they can also make matches to any school-wide professional development that may be underway or planned;
- informing the revision of the suites of achievement standards already available for level 1 mathematics and science, and providing a basis for the discussion about the possible creation of new achievement standards; and
- aligning these findings with ongoing curriculum stocktake work, to encourage professional dialogue about the range of possible purposes for learning mathematics and science, and using these insights to develop new conversations — in addition to achieving success in assessment for qualifications — for teachers to draw on when motivating students to learn.

References

Black, P. (2001). *Report to the Qualifications Development Group Ministry of Education, New Zealand on the proposals for development of the National Certificate of Educational Achievement.* London: School of Education, King's College. www.minedu.govt.nz/web/downloadable/dl5591_v1/Paul%20Black%20report.pdf

Black, P., and Wiliam, D. (1998). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Boaler, J. (2003). When learning no longer matters: standardized testing and the creation of inequality. *Phi Delta Kappan, 84*(7), 502–506.

Bryce, J., and Withers, G. (2003). *Engaging secondary school students in lifelong learning.* Research report. Melbourne: Australian Council for Educational Research.

Casas, M. (2003). The use of standardized tests in assessing authentic learning – a contradiction indeed, Teachers' College Record. www.tcrecord.org

Codd, J., Brown, M., Clark, J., McPherson, J., O'Neill, H., O'Neill, J., Waitere-Ang, H., and Zepke, N. (2002). *Review of future-focused research on teaching and learning.* Wellington: Ministry of Education.

Gale, T., and Densmore, K. (2003). *Engaging teachers: towards a radical democratic agenda for schooling*. Maidenhead, Philadelphia: Open University Press.

Gilbert, J. (2003). "New" knowledges and "new" ways of knowing: implications and opportunities. In B. Webber (Ed.), *Educating for the 21st. century: rethinking the educational outcomes we want for young New Zealanders – NZCER Conference proceedings 2003*. Wellington: New Zealand Council for Educational Research, pp. 15–31.

Harlen, W. (1998). Classroom assessment: a dimension of purposes and procedures. *New Zealand Association for Research in Education Conference*. Dunedin, December 3–6.

Hipkins, R. (1997). Should contexts be used in science examinations? *New Zealand Science Teacher,* (86), 31–35.

Hipkins, R., and Arcus, C. (1997). Teaching science in context: Challenges and choices. In B. Bell and R. Baker (Eds.), *Developing the Science Curriculum in Aotearoa New Zealand*. Auckland: Addison Wesly Longman.

Hipkins, R., Bolstad, R., Baker, R., Jones, A., Barker, M., Bell, B., Coll, R., Cooper, B., Forret, M., France, B., Haigh, M., Harlow, A., and Taylor, I. (2002). *Curriculum, learning and effective pedagogy: a literature review in science education.* Wellington: Ministry of Education.

Hipkins, R., Vaughan, K., Beals, F., and Ferral, H. (forthcoming). *Learning Curves, one year on...* Interim research report – second year. Wellington: New Zealand Council for Educational Research.

Johnson, D., and Kress, G. (2003). Globalisation, literacy and society; redesigning pedagogy and assessment. *Assessment in Education, 10*(1), 5–14.

Kind, P. (2003). Creativity or rationality: what should be the main target for 21st. century science education? *ASERA 34th Annual Conference*. Melbourne, July 8–12.

MacDonald, D. (2003). Curriculum change and the post-modern world: is the school curriculum-reform movement an anachronism? *Journal of Curriculum Studies, 35*(2), 139–149.

Mallard, T. (2002). *Update on NCEA workload reporting.* Media Statement released on 9 August.

McGee, C., Jones, A., Bishop, R., Cowie, B., Hill, M., Miller, T., Harlow, A., Oliver, D., Tiakiwai, S., and MacKenzie, K. (2002). *Curriculum Stocktake, National School Sampling Study. Teachers' experiences in curriculum implementation: General Curriculum, Mathematics and Technology*. Report to Ministry of Education for Curriculum Stocktake: National School Sampling study: Wellington: Ministry of Education.

McGee, C., Jones, A., Cowie, B., Hill, M., Miller, T., Harlow, A., and MacKenzie, K. (2003). *Curriculum Stocktake, National School Sampling Study. Teachers' experiences in curriculum implementation: English, Languages, Science and Social Studies.* Wellington: Ministry of Education.

Minister of Education. (2002). *Update on NCEA workload reporting.* Media Statement released on 9 August.

Ninnes, P., and Burnett, G. (2001). Postcolonial theory and science education: Textbooks, curriculum and cultural diversity in Aotearoa New Zealand. *New Zealand Journal of Educational Studies, Te Hautaka Mätai Mätauranga,, 36*(1), 25–39.

Osborne, R., and Freyberg, P. (1985). *Learning in Science: The implications of children's science*. Auckland: Heinemann.

Ryder, J. (2001). Identifying science understanding for functional scientific literacy. *Studies in Science Education, 36*.

Shavelson, R., and Towne, L., (Eds.). (2002). *Scientific research in education*. Washington DC: National Academy Press.

Stigler, J., and Hiebert, J. (1997). Understanding and improving classroom mathematics instruction. An overview of the TIMSS video study. *Phi Delta Kappan*, *79*(1), 14–21.

Tytler, R. (2003a). School innovation in science: how best to support school and teacher change? *Australasian Science Education Research Association (ASERA) 34th Annual Conference*. Melbourne, July 8–12.

Tytler, R. (2003b). A window for a purpose: Developing a framework for describing effective science teaching and learning. *Research in Science Education, 30*(3), 273–298.

Wackerly, D., Mendenhall, M., and Schaeffer, R. (1996). *Mathematical statistics with applications*. Belmont, USA: Duxbury Press.

Welch, D. (2003). Done too much, much too young. *New Zealand Listener*, 190, pp. 16–21.

Zimmerman, B. (2001). Theories of self-regulated learning and academic achievement: an overview and analysis. In B. Zimmerman and D. Schunk (Eds.), *Self-regulated learning and academic achievement: theoretical perspectives.* New Jersey: Lawrence Erlbaum Associates Inc.

Appendix One: The Teacher Self-reflection Instrument

Shifting the Balance: Preliminary Teacher Survey (to be completed before the visit)

The statements down the left hand side of this table have been distilled from recent research on shifts in a teaching and learning focus in science/mathematics that could better align students' learning with their likely post-school needs in this twenty-first century. They are intended as a *beginning* point for our conversation. Please tick ONE box in each of the three columns (Q1, Q2, Q3). We will discuss these responses during the school visit. The various statements might also help you when thinking about interesting teaching and learning materials to bring to our discussions.

	Q1. What priority do you think should be given to each of these practices?					**Q2. PreNCEA how often did these practices happen in your Year 11 classes?**				**Q3. How often do these practices happen in your Year 11 classes now?**			
	Very high	High	Moderate	Low	Very low	All/most of the time	Often	Occasionally	Hardly ever/never	All/most of the time	Often	Occasionally	Hardly ever/never
1. Providing stimulus materials that challenge students' ideas and that encourage discussion, speculation, and ongoing exploration by groups of students working together.													
2. Moving away from a strong focus on content "coverage". Moving towards a focus on ensuring understanding and meaningful learning of a reduced amount of content.													

	Q1. What priority do you think should be given to each of these practices?					Q2. PreNCEA how often did these practices happen in your Year 11 classes?				Q3. How often do these practices happen in your Year 11 classes now?			
	Very high	High	Moderate	Low	Very low	All/most of the time	Often	Occasionally	Hardly ever/never	All/most of the time	Often	Occasionally	Hardly ever/never
3. Encouraging students to make their own decisions in practical investigations concerning hypotheses to be explored, experimental design, measurement and recording techniques, analysis and interpretation.													
4. Including frequent open-ended investigations or short-term open explorations.													
5. Ensuring higher-order tasks involving the generation, application, analysis and synthesis of ideas, are well represented.													

	Q1. What priority do you think should be given to each of these practices?					Q2. PreNCEA how often did these practices happen in your Year 11 classes?				Q3. How often do these practices happen in your Year 11 classes now?			
	Very high	High	Moderate	Low	Very low	All/most of the time	Often	Occasionally	Hardly ever/never	All/most of the time	Often	Occasionally	Hardly ever/never
6. Encouraging students to actively clarify their own ideas, and to think about their learning processes. (e.g., by using concept mapping, model making, learning journals, exploration of alternative strategies etc.)													
7. Using students' personal interests (sports, hobbies) and social/ethical concerns as the context of mathematics or science topics and involving them in making choices about their learning													

	Q1. What priority do you think should be given to each of these practices?					Q2. PreNCEA how often did these practices happen in your Year 11 classes?				Q3. How often do these practices happen in your Year 11 classes now?			
	Very high	High	Moderate	Low	Very low	All/most of the time	Often	Occasionally	Hardly ever/never	All/most of the time	Often	Occasionally	Hardly ever/never
8. Setting a variety of types of tasks during each unit.													
9. Using a variety of methods to assess student understandings, at various points in a unit, (e.g., open-ended questioning, checklists, project work, problems, practical reports, role plays).													
10. Involving students in decision making about what should be assessed, and when and how assessment should be carried out.													

	Q1. What priority do you think should be given to each of these practices?					Q2. PreNCEA how often did these practices happen in your Year 11 classes?				Q3. How often do these practices happen in your Year 11 classes now?			
	Very high	High	Moderate	Low	Very low	All/most of the time	Often	Occasionally	Hardly ever/never	All/most of the time	Often	Occasionally	Hardly ever/never
11. Ensuring assessment incorporates a range of levels and/or types of thinking.													
12. Probing student understandings and perspectives early in a learning sequence to help plan subsequent lessons.													
13. Ensuring students have ongoing feedback which indicates their strengths and weaknesses and their next learning steps.													

	Q1. What priority do you think should be given to each of these practices?					Q2. PreNCEA how often did these practices happen in your Year 11 classes?				Q3. How often do these practices happen in your Year 11 classes now?			
	Very high	High	Moderate	Low	Very low	All/most of the time	Often	Occasionally	Hardly ever/never	All/most of the time	Often	Occasionally	Hardly ever/never
14. Using a variety of types of experiment to exemplify scientific/mathematical methods and principles, including measurement techniques, variable control, survey work, modelling, and open exploratory designs.													
15. Including discussion of mathematical/scientific evidence contributing to contemporary science/mathematics-related public issues that are of interest/importance to students.													

	Q1. What priority do you think should be given to each of these practices?					Q2. PreNCEA how often did these practices happen in your Year 11 classes?				Q3. How often do these practices happen in your Year 11 classes now?			
	Very high	High	Moderate	Low	Very low	All/most of the time	Often	Occasionally	Hardly ever/never	All/most of the time	Often	Occasionally	Hardly ever/never
16. Discussing and developing understanding of language conventions of science/mathematics.													
17. Basing sequences of work around local community projects, such as environmental maintenance or studies of local industries.													
18. Using learning technologies to support quality learning behaviours such as exploration, conjecture, or collaboration (e.g., spreadsheets, Internet, data loggers, graphics calculators).													

	Q1. What priority do you think should be given to each of these practices?					Q2. PreNCEA how often did these practices happen in your Year 11 classes?				Q3. How often do these practices happen in your Year 11 classes now?			
	Very high	High	Moderate	Low	Very low	All/most of the time	Often	Occasionally	Hardly ever/never	All/most of the time	Often	Occasionally	Hardly ever/never
19. Exploring different values and perspectives that students bring to their science/mathematics learning.													

Appendix Two: Teacher Interview Questions

1. Prior to the introduction of the NCEA, did you use unit standards to assess this subject?

   Yes No (circle one)

   If yes, ask for details of experience with unit standards:

   - how many years used
   - satisfaction with them
   - any concerns or reservations

2. Can you explain to me how you help students to prepare for carrying out their NCEA assessments:

   - for the formal assessment of internally assessed standards.
   - for sitting their examinations for externally assessed standards.

3. Do you think these actions that you have just described have changed the focus of teaching and/or methods of formative assessment in your Year 11 classes?

   Yes No (circle one)

   Why/why not?

4. Do you think these actions that you have just described are helpful for students' learning?

   Yes No (circle one)

   Why/why not?

5. Thinking about the whole Year 11 course that you teach:

   - How do you balance out the time you spend teaching for internally assessed course components and teaching for externally assessed components?
   - Do you think this balance has changed the amount of time you now spend on teaching the various curriculum topics?

   Yes No (circle one)

   If so, can you explain how?

   - Has overall curriculum coverage changed?

   Yes No (circle one)

If so, can you explain how?

6. In your opinion, do you now spend more or less class time working on assessment for qualifications than you did in previous years? (i.e. NCEA vs. School Certificate preparation)

   much more **more** **same** **less** **much less** (circle one and comment)

7. Do you offer re-assessment with respect to the internally assessed standards that you offer?

   always **sometimes** **never** (circle one)

   If yes, how many reassessment chances do you give and what procedures do you follow?

8. Are you happy with these procedures? Why/why not?

9. What other organisational issues have arisen with respect to assessment using achievement standards? How have you resolved these issues (if you have)?

10. Now could you tell me about the assessment moderation procedures you use within the science/maths team at this school.

11. Do you find the moderation discussions helpful for your teaching?

    Yes **No (circle one)**

    Why/why not?

12. Have your moderation discussions led you to make changes in your teaching compared to previous years and/or other classes you teach?

    Yes **No (circle one)**

    If so, what sort of changes?

13. I am interested in your opinion about the impact of the NCEA on the learning of various groups of students. Can we consider five of these groups and discuss whether or not you think that the NCEA has allowed you to better identify and meet their learning needs, and if so, how?

    Non-achievers

    Mäori and Pasifika

    High achievers

    Average achievers

    Male/female students

14. What other impact(s) has the introduction of the NCEA had on the students who study in your subject area? What strategies has the school developed to manage these?

15. What impact has the introduction of the NCEA had on the teaching staff in your subject area? What strategies has the school developed to manage these?

16. Is there anything else that you would like to comment on, that has been raised by our conversation today?