

NCEA IN CRISIS

An Analysis of the Externally Assessed Achievement Standards prepared by the Macleans College Senior Management Team

Introduction

This paper looks at the wide variability of grade distribution in NCEA Levels 1, 2 and 3.

The variability of pass rates and grade distribution from subject to subject and from year to year has greater consequences than scaling¹ ever did in University Bursary. As a result the year a student sits NCEA and the selection of subjects can determine the student's chances of success, including university entrance.

Our Major Concerns

- 1 Large variability in the distribution of grades between Achievement Standards and between subjects within any one year. As a consequence, at Level 3, the choice of subjects becomes critical to the student's chance of gaining UE.
- 2 Large variability in failure rates and in the distribution of grades from year to year in the same standards and subjects. As a consequence, at Level 3, the year in which a student attempts UE will become critical to the student's chance of success.
- 3 Large variations such as this never occurred in either Bursary or School Certificate despite NZQA officials stating to the contrary.

An analysis of 18 School Certificate² subject medians from 1997 to 2001 showed a variation from year to year of less than 7% except in two instances. In 14 of the 18 subjects the variation was 6% or less between these years.

Within each year the medians in 13 of the 18 subjects were within 8% of each other. With the exception of Physics, each subject had a high degree of consistency in the median from year to year.

- 4 Poor management and systems failure by NZQA. A number of unacceptable and inconsistent management practices point to a failure by NZQA to provide a robust quality management system for the NCEA.

¹ University Bursary was scaled for a number of reasons, two of which were: to ensure comparability between subjects for UE as UE was an aggregate of subject marks and as University Bursary carried a financial reward, the number of students who gained an A or B Bursary was rationed.

² The scaling of School Certificate for the purposes of normalising the distribution of marks within and between subjects was abolished in the early 1990s.

Levels 1 and 2

After three years of NCEA Level 1 and two years of Level 2 the variations are still unacceptably large. The qualification should have stabilised by now if this SBA model is a viable one.

There is little consistency in the distribution of *Excellence* and *Not Achieved*³ grades from 2002 to 2004 for the same Achievement Standards across a range of subjects.

- In 2004, Level 1 failure rates for the External Achievement Standards ranged from 74.5% to 3.1%.
- In 2004, Level 2 failure rates for the External Achievement Standards ranged from 74.6% to 17.9%.
- When the failure rates of the same 151 External Achievement Standards in Levels 1 and 2 were compared for 2003 and 2004, the failure rate increased by as much as 37% for approximately half of the standards, and decreased by up to 31% for the other half.
- From 2003 to 2004, of the External Standards sampled at Level 1 and 2, only one third fell within a plus or minus 5% variation in pass rates (which the PPTA has stated is the maximum variation they consider to be acceptable).⁴

Note. A plus or minus 5% variation in a large subject cohort will result in thousands of students either passing or failing depending on the year they sat. For example, in 2004, in English 1.6, over 9500 additional students failed compared to 2003 as the percentage variation from year to year was +22.7% for the Not Achieved grades. (see below)

The following are just a few examples of extreme variation from year to year at Levels 1 and 2 for the Not Achieved and Excellence grades. The variations are unacceptable both in terms of percentage change and the number of candidates affected.

English 1.6 (e)

2002: of 41,670 candidates, 4667 got *Not Achieved*

2003: of 42,922 candidates, 10,032 got *Not Achieved*

2004: of 42,381 candidates, 19,537 got *Not Achieved*

Biology 1.6 (e)

2002: 40.4% got *Not Achieved*

2003: 24.9% got *Not Achieved*

2004: 48.5% got *Not Achieved*

³ The distribution of merit grades shows a similar inconsistency

⁴ Education Review February 17-23 2005, Page 5.

Technology 1.6 (e)

2002: 51.8% got *Not Achieved*
2003: 36.4% got *Not Achieved*
2004: 51.2% got *Not Achieved*

English 1.3 (e)

2002: of 41,642 candidates, 13,408 got *Not Achieved*
2003: of 42,545 candidates, 19,889 got *Not Achieved*
2004: of 41,757 candidates, 15,784 got *Not Achieved*

Human Biology 1.4 (e)

2002: 1.3% gained *Excellence*
2003: 20.9% gained *Excellence*
2004: 4.6% gained *Excellence*

Economics 1.5 (e)

2002: 5.3% gained *Excellence*
2003: 17.6% gained *Excellence*
2004: 5.3% gained *Excellence*

Mathematics 1.8 (e)

2002: of 39,973 candidates, 5116 gained *Excellence*
2003: of 40,509 candidates, 96 gained *Excellence*
2004: of 40,475 candidates, 647 gained *Excellence*

Biology 2.7 (e)

2003: 37.3% got *Not Achieved*
2004: 52.1% got *Not Achieved*

Biology 2.8 (e)

2003: 33.6% got *Not Achieved*
2004: 62.8% got *Not Achieved*

Graphics 2.1 (e)

2003: 20.1% got *Not Achieved*
2004: 47.0% got *Not Achieved*

Physics 2.4 (e)

2003: of 9162 candidates, 13.6% got *Not Achieved*
2004: of 10,302 candidates, 46.9% got *Not Achieved*

Science 2.6 (e)

2003: of 1517 candidates, 62% got *Not Achieved*
2004: of 1831 candidates, 33% got *Not Achieved*

Geography 2.2 (e)

2003: of 5278 candidates, 50.1% got *Not Achieved*

2004: of 5341 candidates, 64.4% got *Not Achieved*

Accounting 2.6 (e)

2003: of 3561 candidates, 49.72% got *Not Achieved*

2004: of 4393 candidates, 29.7% got *Not Achieved*

Mathematics 2.7 (e)

2003: of 20,046 candidates, 33.3% got *Not Achieved*

2004: of 20,981 candidates, 19.3 % got *Not Achieved*

Some of the percentage variations were very large, meaning that for many of the standards thousands of students were either greatly advantaged or greatly disadvantaged, depending on which year they sat a particular standard or group of standards.

The variation occurs in nearly every subject with little pattern between subject or year to account for it.

Research indicates that national cohorts of New Zealand students do not vary greatly in ability across a range of learning areas from year to year. Elley⁵ notes that over the past 20 years there is practically no difference in the ability of national cohorts of students in a range of subjects from year to year.

In instances where one would have thought the examiner should have tried to maintain what was an acceptable pass rate from the previous year, the pass rate has actually declined substantially.

There are many examples of this, however in Level 2 English such variations are critical because this subject contains literacy credits for UE.

English 2.5 (e)

2003: of 26,937 candidates, 10,845 got *Not Achieved*

2004: of 31,956 candidates, 18,822 got *Not Achieved*

English 2.6 (e)

2003: of 26,768 candidates, 28.2% got *Not Achieved*

2004: of 30,605 candidates, 57.4% got *Not Achieved*

English 2.4 (e)

2003: of 26,550 candidates, 57.9% got *Not Achieved*

2004: of 30,954 candidates, 66% got *Not Achieved*

A near two-third failure rate is unacceptably high as is the variation from year to year.

In 2004, thousands of English students had a far less chance of gaining UE literacy than in the previous year.

⁵ Elley. Facts and Fallacies of Standards Based Assessment unpublished address October 2004

Level 3

An analysis of the first year of Level 3⁶ also exhibits a large variation in failure rates:

- The failure rate ranged from 86.6% to 10.6%.
- In 19 of the 85 standards over 60% of the students failed, but in 16 other standards the failure rate was less than 30%.
- Four of the 11 easiest standards were in the Visual Arts.
- Four of the 15 hardest standards were in Technology.

These vast differences in failure rates are critical for UE when students have to gain a minimum aggregate credit total. Naturally, high failure rates greatly disadvantage students in gaining UE, but what is even more unacceptable is the variance between subjects.

Level 3 Excellence grades (and Merit grades) are critical for students wishing to gain a place in select entry courses at university as they carry a greater weighting than Achieved grades when the universities calculate grade point totals. Again, there was a large variation in Excellence grades.

- The percentage of Excellence grades awarded by Achievement Standard ranged from 31.9% to 0%.
- Eight standards had over 10% of the candidates gaining Excellence.
- Seven standards had 0% of the candidates gaining Excellence and six of these were in Technology.
- The percentage of Excellence grades ranged from 14.8% to 0.4% by subject when aggregated.

Very few students in a large number of standards are achieving at the highest level.

- 19 of the 85 standards had less than 1% of the students gaining Excellence.
- 60 of the 85 standards had less than 5% of the students gaining Excellence.

In the past, there was some variation in mark distribution between subjects but not as great as is evident under the NCEA. Students choosing particular “courses of study” such as the Technologies and Design, and Commerce will have their chances of UE seriously affected by the subjects they choose, whether they are approved subjects or not.

This includes the internal components where it is expected the achievement rate will be higher.⁷

⁶ From 85 externally assessed Achievement Standards in the traditional Year 13 subjects. From the provisional national results as at February 2005.

⁷ An analysis of 148 Level 1 and 99 Level 2 externally and internally assessed Achievement Standards from 2002 and 2003 showed that the internal standards have much higher pass rates. At Level 2, on average, it was three times as difficult to pass an external standard as an internal. A similar analysis cannot be done for the internally assessed NCEA results in 2004 because the Not Achieved rates are unavailable.

- In the practical based subjects students who studied Technology/Design had a far less chance of gaining UE than those who studied the Visual Arts.
eg. In the externally assessed Achievement Standards 69.5% of the students failed Graphics but in Practical Art only 24.6% failed.
- In Commerce, students who chose Economics were disadvantaged over those who chose Accounting.
eg. In the externally assessed Achievement Standards 56.5% failed Economics but in Accounting just 31.2% failed.
- Students studying Media Studies were at a disadvantage, compared to those studying English.
eg. In the externally assessed Achievement Standards 59.22% failed Media Studies but in English, just 44% failed.
- In the Humanities, students studying Social Studies and English were disadvantaged over those who studied History, Geography and Classical Studies if they required Excellence grades to boost their grade point total for limited entry courses to university.

In the externally assessed Achievement Standards:

- approximately one in 15 and one in 20 students achieved an Excellence in History and Geography respectively but in Social Studies only one in 250 achieved an Excellence.
- one in every 14 students gained an Excellence in Classical Studies but in English only one in every 40 achieved an Excellence.
- one in every seven students gained an Excellence grade in Accounting, but in Economics, only one in 100 gained Excellence.

There are many other examples where subject combinations and courses of study can affect a student's chance of gaining university entrance.

Poor Management and Systems Failure by NZQA

A number of unacceptable, unprofessional and inconsistent management practices point to a failure by NZQA to provide a robust quality management system for the NCEA.

These include:

- Poorly written and formatted examination questions resulting in high failure rates and highly variable grade distributions.
- Examination questions that did not assess to the published achievement standard criteria.
- The use of non-statistical methods to adjust marking schedules, such as the pre-marking of guinea pig scripts in some subjects, but not in others.
- The alteration of marking schedules during the marking process, again in some subjects but not in others.
- Some markers reporting that no official marking schedules were made available to them by the examiner at the initial markers' meeting.
- Hundreds of candidates' papers across a range of subjects being devoid of any marker notations whatsoever.
- Schools reporting inaccurate marking, where students had failed a standard yet had given clear evidence of having achieved the requirements of the standard in their answers.
- Students with incorrect grades on their interim results notices and other students being marked absent for examinations they sat and completed, with the marked scripts returned to the candidates by NZQA as conclusive evidence.
- Many other students have reported having their scripts mislaid or having received scripts and grades of other candidates.

Conclusion

- A student's chance of success in the NCEA can be determined more by the year in which they sat and the subjects they took than by their ability.
- Wide variability between Achievement Standards and subjects has a critical effect at Level 3 where University Entrance is calculated by aggregating subject credit totals.
- The variability is greater than in past national examinations.
- Change in the ability of a student cohort from one year to another does not explain the variability seen in the NCEA. (Elley⁵)
- Poor management practices and systems failure have exacerbated the failings of the NCEA.

For a national qualification, such matters place the NCEA's credibility in question.