

National Foundation for Educational Research



International Survey of Results Reporting

Jenny Bradshaw

Rebecca Wheeler

December 2009

Ofqual/10/4705



This report has been commissioned by the Office of Qualifications and Examinations Regulation.

Contents

Executive summary	3
1 Introduction	4
1.1 Aims	4
1.2 Methodology	5
1.2.1 Literature review.....	5
1.2.2 Development of taxonomy	5
1.2.3 Classification of assessments	7
1.3 Structure of the report	7
2 Review of the literature	8
2.1 Introduction.....	8
2.2 Representation of results	8
2.3 Level of detail	9
2.4 Error and uncertainty	11
3 The reporting of results	15
3.1 Introduction.....	15
3.2 Types of assessment.....	15
3.3 How results are represented	16
3.3.1 Grades	17
3.3.2 Number of grades	17
3.3.3 Scores	18
3.3.4 Profiles.....	19
3.4 The level of detail reported	20
3.5 Error and uncertainty	20
4 Conclusion	23
5 References	24
Appendix A Classification of assessments	26
Appendix B Search Strategy	34
Appendix C Example results reports	47

Executive summary

Introduction

The NFER was commissioned to carry out an International Survey of Results Reporting. The relevant literature and examples of assessments were searched to identify how results are represented, what level of detail is reported and what steps are taken to quantify and report on error internationally. From this evidence, a taxonomy has been developed to classify approaches to the reporting of assessment results. This has been used to classify a range of international assessments.

Key findings

- The way results are reported depends on the intended use of the results and who the results are to be reported to.
- Two opposing issues must be weighed up when deciding on the level of detail of results reporting. These are the increased reliability when few grades are reported and the greater information when many are reported.
- International assessments have been classified using the developed taxonomy. The classification is by three main areas. These are:
 - A description of the assessment, which includes at what stage of secondary education the assessment is used, the purpose, who makes the award, the mode and method of the assessment, and whether the assessors are external or internal.
 - How the results are represented, for instance by grades, scores or a profile and the numbers of these.
 - Whether error or uncertainty is reported.
- Few examples were located of reporting uncertainty or error in their results to students. An introduction of the reporting of error in high-stakes qualifications would need careful handling to ensure this did not result in misinterpretation and a loss of confidence in the system.

1 Introduction

1.1 Aims

Ofqual has put in place a programme of research into the reliability of results of qualifications, examinations and assessments. One of Ofqual's stated aims is to shine a light on this important issue, finding out more about the consistency of results, and what learners, teachers, employers, higher education bodies and the general public think and feel about reliability. One aspect of such a searchlight has to be to examine approaches to the nature of reporting, as this gives scope for understanding what can be said and done about reliability.

To this end NFER was commissioned to carry out an International Survey of Results Reporting. This survey aims to:

- identify international approaches to the reporting of assessment results for individual students, with a particular focus on the following three aspects:
 - how results are represented (e.g. letter grades, number grades, marks, percentiles, profiles, descriptions, etc.)
 - the level of detail reported (e.g. how many grades, what combinations of quantitative, categorical and qualitative information, etc.)
 - what steps are taken to quantify, report and explain (interpret and communicate) error and uncertainty;
- develop a taxonomy of approaches to the reporting of assessment results, which reasonably exhausts the range of approaches adopted internationally; and
- classify a substantial number of assessments, from around the world, according to this taxonomy.

The current report analyses and discusses the range of approaches to results reporting used internationally, based both on information in published literature and the information found on individual assessments. For purposes of comparison, some assessments from England are included in those classified alongside those from other countries.

1.2 Methodology

1.2.1 Literature review

As a first stage in searching for relevant literature, keywords and parameters for the search were established in consultation with the NFER library. The keywords and search strategy are in Appendix B along with details of the literature databases and websites which were included in the search.

The results of this search were somewhat disappointing although this did to some extent match our expectations, since we had predicted that websites and direct enquiries were more likely to give us the type of detail we require. However, a small amount of relevant published literature was identified. This has contributed to a general discussion of approaches to reporting of results and the investigation and reporting of error and uncertainty which is reported in Chapter 2.

A second parallel stage was a search of the proceedings of conferences of the Association for Educational Assessment-Europe and the International Association for Educational Assessment. Here again there was little of direct relevance but a small number of papers were identified.

1.2.2 Development of taxonomy

Identification of the categories within which assessments can be described was done partly by first identifying relevant descriptive categories which would apply to any assessment – for example whether reporting is by grade, score or in some other way, the stages of education at which qualifications are taken, modes of assessment and so on. These categories were further refined in the light of the actual information found, so development of the categories for the taxonomy was to some extent an iterative process. The categories included in the taxonomy are described below. The sub-categories used in the taxonomy are shown in bold. A full description of the meaning of each category is shown in the Key for Appendix A.

Stage

This describes whether the assessment is aimed at **lower secondary** (to age 16) or **upper secondary** (generally age 17-18).

Purpose

The purpose of the assessment may be for **general** reporting of achievement – for example, a school-leaving certificate or an assessment intended to report progress to parents. Or it may be designed for a particular purpose such as entry to **tertiary** education or employment (**vocational**). In some cases an assessment or qualification may serve several purposes (**mixed**).

Award

This category describes whether a final result is given by an **external** body such as a government education department or an examinations authority, or is an **internal** result awarded by individual schools.

Mode

In the taxonomy the categories used to summarise modes of assessment are: **modular** which is used to describe a qualification which is generated from separate elements being combined into a final result; **terminal** in which case the results of a (usually formal) examination or test are used to summarise achievement at the end of a course of study; or a **mixture** – for example results from teacher assessments or from work done during a course of study may be combined with results from a final formal examination.

Method

The method of assessment may be **coursework/continuous assessment** using a variety of methods of assessment throughout a course. These results may be generated from, for example, tests set by teachers or teacher judgments of achievement, tests or examinations either set by schools or externally provided, or projects or other work done during a course. Alternatively a result may be based on **formal** assessment using tests or examinations. Finally the results may be based on a **mixture** of these.

Assessors

The assessors may be **internal** which is used in the taxonomy to describe assessment by schools or teachers, possibly with some form of external moderation. Or it may be **external** in which an assessment is set and marked by an external body. The third type in the taxonomy is a **mixture** of assessors – ie some elements of both internal teacher or school assessment and externally set and marked assessment.

Results

This category summarises the ways in which results can be represented. This may be by **letter grades** (eg A, B, C etc), **numerical grades** (eg 1, 2, 3 etc), **numerical scores** (such as percentages, standardized scores, raw scores) or a **profile** which may include elements of the others – for example, a profile report may give an overall decision on whether a diploma or certificate has been awarded, but may then also give grades for some elements, scores for others, and perhaps also achievement of non-academic requirements such as completion of work experience.

Structure

The structure of an assessment result or qualifications award may be some form of **certificate** which gives an overall summary of achievement in all subjects (for example as a final school-leaving diploma). The other type of structure included in the taxonomy is one which gives **discrete** results with completely separate awards for different subjects.

Error

This category in the taxonomy describes whether a results report or certificate for an assessment includes any form of reporting of possible error or uncertainty in the results. The categories used are **no** where information has been found that results reports do not include this and **not known** where insufficient evidence has been obtained on the exact details of results reports. Where there is some indication, the categories used are **statistical** where uncertainty is shown by an indication of a band within which a particular score falls, or **text**

where a result includes some form of statement which either expresses uncertainty or recommends caution in the interpretation of the result.

1.2.3 Classification of assessments

Following development of the taxonomy, the next stage was to decide on the assessments to be classified.

As a first stage in searching for information on individual assessments, the countries and suggested assessments to be included in the search were identified. The aim was to cover the widest possible range of approaches and types of assessment used in secondary education, within well-established systems which were most likely to have publically available information, and gave a good range of international coverage. An additional criterion was that there should be at least some element of external assessment or external regulation to ensure the closest comparison with the situation in England. This information was obtained either from websites or from Eurydice or INCA databases. Information on GCSE and GCE in England is also included for purposes of comparison, and this information was obtained from the websites of QCA and of the awarding bodies.

The level of detail and the usefulness of the information obtained varied. In general it is relatively straightforward to obtain descriptive information such as the way in which results are represented (eg numbers, grades etc.), the ways in which subjects or different types of assessment are combined for award of certificates and the results needed for different purposes such as progression through school or university entry. There is however more variation in the extent to which it has been possible to find out about the level of detail which is given to various stakeholders (students, parents, schools, etc.) or included in individual result reports. The aspect on which least published information is available is (not unexpectedly) the steps which are taken to explain, and to quantify, error and uncertainty in the assessment results which are given to individuals.

Following this stage decisions were made about the assessments to be classified according to the taxonomy. In some cases only minimal descriptive information was found which was not useful to the discussion and these assessments were therefore not included.

1.3 Structure of the report

Chapter 2 of this report discusses more general issues and insights from the review of relevant literature on how results can be represented and issues regarding the ways in which error and uncertainty can be quantified and reported.

In chapter 3 the assessments which are classified according to the taxonomy categories described above and included in Appendix A are discussed.

In Chapter 4 the discussion is summarised and conclusions are drawn.

2 Review of the literature

2.1 Introduction

This chapter discusses general issues relating to the reporting of results and the quantification of error or uncertainty in results. Literature on reporting of results tends to have been produced during times of educational change, such as the introduction of GCE A Level examinations in the UK and since the implementation of No Child Left Behind in the USA. However, aspects relating to the reporting of error or uncertainty in individual results are less well discussed than aspects which relate more to the structure of assessments or the ways in which results are represented. Even in articles relating only to the reporting of student test results, aspects of reliability or reporting of error and uncertainty are often ignored (for instance Roeber, 2003).

2.2 Representation of results

Test results are represented in different ways depending on the intended use of the score and who it is to be reported to. They can be presented to show what a student can do compared to others in the same class, school or area, which is often termed 'norm referenced' or 'cohort referenced'; or they can be presented to show achievement against a pre-defined standard, often expressed as what is expected of a student at that age, which is frequently referred to as 'criterion referenced' or 'standards referenced'. Results can be presented in both ways, so that students know that they are progressing and whether they are level with their peers. Below is a collection of the ways that test results can be represented.

The raw score of a test is usually the number of points achieved in a test administration. For multiple choice tests, it can be the number of items that were answered correctly or the number answered correctly adjusting for guessing. The raw score is useful when all students take the same test, but not when different versions of a test are sat and so if it is reported, it is often given with normative scores, such as percentile ranks, stanines, and normal curve equivalents (Harris, 2003).

Percentile ranks are an indication of achievement compared to other scores by reporting the percentage of examinees who earned the same or a lower score. Stanines are integer scores converted from the raw score. They range from 1 to 9 and have a mean of 5 and a standard deviation¹ of 1. Normal curve equivalents are integers converted from the raw scores that range from 1 to 99 and have a mean of 50 and a standard deviation of 21.06.

Scale scores are scores that have been converted to have a particular mean and standard deviation. The advantage is that pupils that have not taken the same tests can then be

¹ Standard deviation is a measure of the spread of a set of values from the mean. It is calculated from the square-root of the variance.

compared. One disadvantage of scaled scores is that as the scales are arbitrary, users can have difficulty understanding the importance of differences between different scale scores until they have been used for some time. (Jaeger, 2003)

Level, category, or proficiency scores such as those used by the National Assessment of Educational Progress (NAEP), a nationally representative assessment of students in the US conducted across a range of subjects, put students into categories, such as 'basic', 'proficient' and 'advanced' (Jaeger, 2003). The scores generally have descriptors associated with them to describe what a student at a particular level is able to do. Developmental scores are used to show a student's position on a developmental continuum. One example is grade equivalents, which try to establish a score scale that ranges across multiple school grades so that students can be easily tracked over time. Harris (2003) considers that these are easily understood by teachers and parents in the US and that another advantage is that teachers and parents also tend not to over-interpret extreme scores, such as when a student's grade equivalent score is a large number of grades higher than their year group.

Some tests report multiple scores, for instance, the Iowa Tests of Basic Skills reports the raw score, developmental standard score, grade equivalent, national and local percentile ranks, stanines and normal curve equivalent score (Harris, 2003). Goodman and Hambleton (2004) reviewed practices of reporting of results from the information contained within the scoring guides from a range of tests. They collected Grade 10 student reports and the accompanying interpretive materials from the departments of education from 14 US states, three US commercial testing companies and the departments of education of two Canadian provinces. They found that the most common way of providing meaning to scores was to report overall scores and describe the skills and knowledge that each performance level represented. Another popular method was to report student scores with relevant comparison groups, such as other students from the same school, district or state.

Similarly, many published school tests in the United Kingdom give a range of scores, often incorporating measures of error. For example the Progress in Maths series, produced by NFER, reports an age standardised score, stanines and percentile ranks. The standardised scores are presented with an accompanying 90 per cent confidence interval.

2.3 Level of detail

It is important that the grade assigned to a candidate is correct as examination grades can have long lasting implications. Much of the literature involves mathematical studies with hypothetical scores and test reliabilities to calculate a suitable number of grades for an assessment. The main arguments for determining a suitable number of grades are based around:

- how reliable should the grade be (Mitchelmore, 1981), and
- the loss of information from having few grades justifies having a larger number of grades (Ebel, 1969)

There are also concerns over how many grades examiners, students and other interested parties can understand and are, therefore, meaningful. The difficulty is between being able to come to a description of a candidate's achievement and the reality that achievement must be a continuous scale. Therefore, each grade can represent a level along the continuum and within each grade will be candidates with a range of achievement. There are two possible approaches to differentiate between grades. Grade criteria usually specify the minimum a candidate must do to achieve that grade and so will be most accurate for candidates at the bottom of a grade. The second approach is grade descriptions, which usually give an overview of what candidates will be able to do at each grade and describe a typical candidate in the middle of the grade. Grade descriptions are useful to help examiners reach appropriate decisions and others to interpret what a grade means. Cresswell (1986) considers that it is not necessary for all grades to have a specific set of criteria. It can be acceptable to have criteria for, for instance, grade C and grade A and for the criteria for grade B to be that a candidate has done more than meet the criteria for grade C, but not enough to meet the criteria for grade A.

Some researchers believe that the risk of being wrongly graded should dictate the number of grades awarded. This was an idea proposed by Mitchelmore (1981). He compared examinations in the US, England and Jamaica and was concerned that many examinations use different numbers of grades, but with no clear rationale as to why. He found exams varied from simple pass-fail to 15-point scales to define achievement. Mitchelmore was in favour of using grades as a replacement for percentage scores because percentage scores can be misleading in terms of their implied accuracy. By grouping percentage scores into grades, the number of scale points is reduced and so the likelihood that repeating the test would affect the result of the assessment is also reduced. As a standard for the accuracy of grading, he proposed that 'the number of points on a grading scale be considered acceptable if it is almost certain that, were an assessment to be repeated, each student would be re-graded within one scale point of the original grade', where 'almost certain' is a mathematical probability of at least 90 per cent. Mitchelmore devised a formula to work out the acceptable grade width for a test, and in turn the number of acceptable grades in a scale, using the standard error of an assessment. The smaller the standard error, the greater the number of acceptable points on the scale and so the number of acceptable grades varies for each test and also the method of assessment, as for instance, multiple choice tests tend to have smaller standard errors than essay based tests.

One difficulty with basing the number of acceptable grades on the reliability of a test in this way is that it can mean that there are few grades and therefore the grading may not be fit for purpose, either for the selection they are to be used for or for student motivation. If too much caution is put on the reliability of grades, and the range of scores within each grade thus becomes too large, the results may not differentiate sufficiently between students. As more grades are used, grades become more like percentage scores in terms of the intervals between grades, but the theoretical advantage of grades is that they should be comparable between subjects, so that, for instance, at A level an A grade in maths and an A in English should represent a similar level of achievement, as should an A in maths from one year to the next

(Cresswell, 1986). However, this is not necessarily the case, as discussed in 'Techniques for monitoring the comparability of examination standards' (Newton et al, 2007).

The fewer the number of grades, the more likely that a candidate has been awarded their true grade. However, the consequences associated with getting the grade wrong are more acute compared with the situation in which more grades are available. Candidates whose scores are near grade boundaries are at greatest risk of being mis-graded and the probability of being mis-graded is greater the higher the number of grades, but each mis-grading will be less serious. For instance, in a pass or fail exam, the likelihood that a candidate is awarded the correct results is higher than an exam with grades A to E, but the consequence to the candidate of receiving a fail can be much larger than being awarded a D when a C was required. In the first case, there is no way to differentiate between a candidate that just failed and a candidate that failed badly (Cresswell, 1986). Sometimes it may be difficult to give a rationale for not awarding grades in this pass or fail way, for instance when grading competency-based assessments it may only be meaningful if students are described as 'competent' or 'not competent' and so it can be hard to assign grades as levels of competency (Suto and Green, 2005). However, a scale with few grades can accentuate the problems of making students with very different scores that are at the low end and high end of a grade seem similar, whereas students on either side of a grade boundary seem more different than they are (Kleven, 1979).

2.4 Error and uncertainty

There is very little evidence in the literature of steps that are taken to explain or quantify error and uncertainty in results given to individuals, especially in graded assessments. By not providing this information, test results can be perceived to be more reliable than they actually are, with the risk that errors are compounded if these grades are then combined (Cresswell, 1988).

Please (1971) demonstrated a method of estimating the proportion of candidates that are mis-graded for various reliabilities of A level exams. Due to the percentage of candidates that were estimated as being mis-graded even with examinations with high reliabilities, he suggested that candidates should be given more than one grade to show the limitations of the assessment. The problem, which Please points out, is that grade 1/2 would be regarded as the top grade and 2/3/4 as the second, and so on so that students with grade 1/2 would be picked over those with 2/3/4.

One example where the lack of information on reliability is misleading is discussed in Gardner and Cowan (2005) paper. They considered the issues of wrong grading and providing enough detail for selection in their analysis of the Northern Ireland Transfer test. This is a high stakes test of English, Mathematics, and Science and Technology that was used for the last time for the 2009 intake to Grammar schools. Over 3100 completed practice scripts were collected that had been completed by students in preparation for the Transfer Test. Over 900 students were matched as taking two tests. Although the tests had high reliabilities the difference in test score between getting an A grade and a D grade was only 18 marks out of a test that is 150 marks in

total. This 18 mark difference is split into 6 grades (A, B1, B2, C1, C2 and D). The grades are not even widths, as it is decided in advance the percentage of students that will fall into each grade, and are very narrow, for instance B1 is only 3 marks wide. The standard error of measurement was used to identify the 95 per cent confidence interval of where a student's score might lie. They found that for these tests, the 95 per cent interval was between 9 and 10 marks above or below their actual score. As 18 marks was the span of all grade boundaries, the possibility and significance to the student of misclassification is very high as the type of the rest of their schooling is determined from this score. Based on the standard error, over 70 per cent of students were at risk of misclassification.

Concerns have also been raised about the reliability of levels from national curriculum testing in England. This was discussed by Newton (2009). For students, these are not high stakes tests, but their use for school league tables means that they are of high importance to teachers. Newton discusses some estimates of the reliability of levels, which have ranged from over 30 per cent to about 15 per cent misclassification at Key Stage 2 English depending on the method used to estimate misclassification. He later suggests that a misclassification rate of 20 per cent may not be indefensible.

Skurnik and Nuttall (1968) wrote that some USA public exam bodies issued results for candidates as a band of scores based on the standard error and a single score. This was to communicate the inherent error within the assessment. These boards often also published the reliability of their assessments. They concluded that although correlation coefficients are often used to describe the reliability of an examination, they are hard for users to understand and know what the difference means, for instance, for a test with a reliability of 0.9 and another with a reliability of 0.8. They suggested overcoming this by using the standard error of measurement, expressed in terms of the test score or examination grade units. They proposed using a measurement index based on getting the correct grade plus or minus one, which would tell the examiner an appropriate number of grades that could be assigned. One problem with this idea would be the possibility of differing numbers of grades for different examinations of the same subject due to differing reliabilities of each test.

The Washington's School Improvement Strategy report on learning (Bergeson, 2000) contains examples of an individual student (parent) report. Student's scores are reported on a scale from 150 to 600 and a score of 400 indicates that the student has met the required standard. Scores are represented graphically with a dotted line at 400. There is a shaded area on either side of the 400 mark that indicates an area of one standard error of measurement for the score of 400 and so a range of scores that are not statistically different from 400.

In Goodman and Hambleton's (2004) review of student reports and interpretive materials, they found that from the US states, Canadian provinces and commercial companies that they looked at, four of the 11 states and two of the three commercial companies provided information on the precision of at least one type of overall test score. This was done graphically by two states and one commercial company. The examples shown were reporting scale scores with associated

numerical standard errors; reporting the scaled score numerically and graphically to display the score and the probable range of scores within the overall grades; and finally reporting the scaled score numerically and then graphically showing the percentile range and associated stanine error. There were no examples of reporting errors of grades. In their opinion, displaying errors in a graphical way rather than as a numerical standard error is an easier way to understand the information. Although they had some concerns over whether reporting of precision of test results would lead to confusion for parents and students, not reporting error was seen as a far more serious problem as it makes the results seem far more accurate than they are. They concluded that error should be reported and that any statistical wording used to describe it should be clearly defined.

Gill and Bramley (2008) studied the implications of marker correlations on grade classifications consistency. They used simulated data to estimate the extent to which examinees might get a different grade with two variables: the level of correlation between markers, and the grade bandwidths. Their analysis was based on hypothetical markers that marked with the same severity rather than markers that marked with different levels of severity. They simulated sets of test scores with markers with different correlations and calculated the proportion that received the same grade, one grade different, two grades different, and so on. They then investigated the effect of grade bandwidth on these proportions. The score distributions in different subjects were simulated using 'reasonable values for mean and standard deviation and plausible inter-marker correlations based on previous research.' They varied the inter-marker correlation to simulate the effects of marking for different subjects as it has been shown that papers containing many highly structured questions, such as maths, generate higher correlations than those with a few longer essay type questions, such as English (Newton, 1996). They also studied how aggregating grades, such as when different modules at A level are combined, impacted on the overall reliability of the final grade. They found the impact at aggregate level was much less than at unit level. They concluded that the grade could be reported with an estimate of the proportion of candidates with that grade who might have received a higher or lower grade if marked by another. However, the proportions cannot be interpreted as probabilities for individual candidates as the probabilities for individual candidates depends on how close their true score is to the grade boundary.

2.5 Conclusion

This chapter has identified discussions in the published literature of more theoretical aspects of results reporting. This contributed to the interpretation of the assessments which have been classified in the taxonomy developed for this report.

The way test results are reported depends on the intended use of the results and who they are to be reported to. The total number of grades in an assessment influences the reliability of the grades. By decreasing the number of grades in an assessment, the reliability of the grades is increased, but there is also a loss of information, with two main effects. These are that a misclassification has a more serious implication for a student, and that students and their

teachers learn less from the results of an assessment. Although there has been research into the reporting of error and uncertainty, there have been no clear answers as to how the reliability of grades could be reported in a meaningful way.

3 The reporting of results

3.1 Introduction

This chapter describes the range of approaches used to report results– ie the type of information which is included on the results notices or certificates which are given to students. The assessments which have been classified according to the taxonomy in Appendix A and the sample reports in Appendix C are used to illustrate the points made.

The tables in Appendix A show the classification according to three main areas:

- **Description of assessment**, which includes the stage of secondary education at which the assessment is used and the purpose of assessment; whether the award is made by the school or by an external body; the mode and method of the assessment; and whether assessors are internal or external.
- **Representation of results**, first shows whether they are represented by grades, scores or a profile which records various aspects of achievement which may be represented in different ways. If grades are used the number of these is shown. The structure of the results in terms of whether they are discrete subject awards, a certificate which combines several elements into a single result is then shown.
- **Individual results reports** shows whether reports of results given to students indicate error or uncertainty in the result, and also shows what details are included on results reports (although full information is not available for all the assessments which have been classified).

In this chapter, first the various types of assessment which have been classified are described. The representation of results given to students is then discussed in terms of whether this is given as grades, scores or a profile of achievement. The level of detail included in these reports of results is then described. The chapter ends with a description of the steps taken to quantify and explain error and uncertainty in results, and the extent to which this is made available either on results notices or in some other way.

3.2 Types of assessment

The majority of assessments which have been identified and included in the tables in Appendix 1 are for the purpose of either certification at the end of secondary education, generally at the age of 17 or 18, or for entry to higher education. In some cases, as with GCE examinations in England, the results of an assessment serve both purposes. In others, for example in the case of the SAT and other tests in the United States, an assessment may be intended for the specific purpose of certification of achievement for entry to post-secondary education.

Le Metais (2002) describes the types of certification used internationally for achievement at the end of secondary education. She identifies three main types:

- Diplomas based on achievement in a range of subjects including a core of common subjects, a block specific to an area of study, and range of elective subjects.
- Individual subject qualifications within an overall framework.
- Credit-based models with accumulation of points towards high school graduation.

The assessments described in Appendix A mainly fall into one of these three types and most of those which are for upper secondary have the function of providing evidence for entry to higher education as well as providing a more general record of school achievement. However the tables also include a fourth type of assessment, which is tests which are intended specifically for entry to higher education rather than for summarising school attainment. These are found mainly in countries where high school graduation is based on school or state-based standards, such as in the United States, Korea or Sweden. In such a case results in an external test may be required for college or university entry, although in the case of the Swesat in Sweden the main purpose is for those whose attainment at the end of secondary school is not high enough for university entry. The results of such tests are commonly reported as scores which may be based on the results of one test paper which is not subject-based (as with the SAT reasoning test in the USA) or may be based on separate scores from more than one subject (as with the CSAT in Korea).

Although the purposes and the overall structure of the assessments reported generally fall into one of the four categories identified above, there is much more variation in the detail of how results are represented and reported to students.

3.3 How results are represented

This section describes how results are represented – ie the nature of the result which is given to students, described on results notices and certificates, and so on. The detail of how this result is arrived at and the underlying information which makes up this result are also discussed.

As discussed in the previous chapter there are many ways of representing results which to some extent depend on the intended purposes and audiences. The taxonomy divides ways of doing this into four groups. The first three groups are: grades (either letter or number grades); scores of various types (eg raw scores, percentages, ranks, scaled scores etc.); or profile reporting which may include an associated overall ‘pass/fail’ decision or a statement of a learner’s overall general level of achievement.

In practice many assessments use a combination of methods to represent results, which is the fourth category in the taxonomy. In this case a results report or certificate may include various component parts which are all represented in different ways. These may be reported as a profile or the component parts may be aggregated in some way.

3.3.1 Grades

Some assessments use grades as the main way of reporting results, although some of these also give other information in results notices or certificates. These grades may be represented either by letters, with grade A as the highest, or by numbers.

In England the traditional practice in school examinations has been to have individual subject qualifications and to use letter grades as the main method of reporting overall results. This is the practice followed with GCSE and GCE A/S level qualifications. The Diploma qualifications which are currently being piloted move away from this to some extent in that the different elements which will contribute to achievement of the Diploma cover a range of types of reporting, but nevertheless the final Diploma result will be expressed as a familiar letter grade.

Letter grades are also used in other countries which had historical influences from British educational traditions and which have individual subject qualifications. They are, for example, used in most subjects of the Hong Kong School Certificate examinations, in the Junior and Leaving Certificates in Ireland and in the Graduation program of the Canadian province of British Columbia, although in the latter case the grade is for a diploma which is based on a range of subjects rather than for individual subjects. Letter grading is also used in many Commonwealth countries which are not included in this report. An example of a results report for the Hong Kong School Certificate Examination can be seen in Appendix C.

Other countries which once had influences from the British examination system have now moved away from this approach, generally towards a more standards-based or level-based approach. Some states of Australia, for example, now have profile-based diplomas based on a range of subjects and also include other aspects of achievement. In Hong Kong the approach is also changing. Both English and Chinese Language at HKCE now use numerical grades with associated descriptions of levels of achievement.

Other qualifications use number grades rather than letters. This is the case in Denmark and the Netherlands and also for the Advanced Placement Test in the USA. With number grades there may be variation in the order of grades. In the Hong Kong School Certificate examinations for English and Chinese for example the highest grade is 1, as is the case with the APT in the USA, while in the examinations in the Netherlands 1 represents the lowest grade.

3.3.2 Number of grades

Qualifications which use grading vary in the number of grades into which the results are classified. The previous chapter discussed the issues to be taken into account when deciding the number of grades for a qualification and the compromises needed between having a large number of grades which may increase the risk of misclassification or a smaller number which may make the qualification less meaningful and reduce the amount of discrimination possible between students. The number of grades in the assessments included in Appendix A ranges from 13 for the Leaving Certificate in Ireland to 5 for the Graduation Program Diploma in British

Colombia and the APT. In the case of the British Columbia Diploma, however, grades are given only for scores above 50 and other information is also included on results reports.

The 13 grades for the Leaving Certificate in Ireland have subdivisions of grades A-F and each represents a fixed percentage score. For example the lowest grade which is F represents a score between 10-24. The highest grade which is A1 represents a score of 90 per cent or over (The total score is 100). Between F and A1 each grade represents a score range of only 5 marks. This large number of grades and correspondingly small range of marks between grade boundaries would have the advantage of increasing the amount of discrimination between candidates but would be sensitive to misclassification.

A more common approach is to have somewhere in the range of five or six grades. For example, the Ontario Graduation Program and the APT have five grades. The GCE examinations in England have grades A-E with the addition of an A* grade for the highest performers, so in effect have six grades. The Junior Certificate in Ireland and the Hong Kong School Certificate examination also have six.

The decision on the number of grades should of course take into account the spread of attainment in the cohort, the minimum standards expected and the purpose and intended uses of the assessment, as discussed earlier in section 2.3; so it is in fact not straightforward to compare one qualification with another as far as grading is concerned.

3.3.3 Scores

A second common practice internationally is to represent results numerically in some way. This may be, for example, a raw score, standardised score or percentage. In some cases a combination of types of numerical reporting is used.

In Ireland the Leaving Certificate Applied Programme which was introduced in 1995 reports percentage scores but also three levels of achievement (Pass, Merit, Distinction). These are linked to specific percentage scores. The Baccalaureate in France has a similar approach, with three levels of achievement linked to specific score bands (*AB : Assez bien, B: Bien, TB: Tres bien*). Individual subjects are reported on a numerical scale.

Some statements of results show scores for the candidate and also compare these scores with those of others by means of rank orders or mean scores. This is the case, for example, with the Higher School Certificate in New South Wales, the Ontario Secondary School Diploma and the CSAT in Korea.

Reporting of scores is not restricted to qualifications which use this as the main method of representing results. In fact, most qualifications which give grades also report scores of some type. This is the case, for example, with the Junior and Leaving Certificates in Ireland which report percentage scores and the British Columbia Graduation Program which report grades, percentage scores and course credits. In England 'uniform marks' are given for GCE and Modular GCSEs to represent the scores given for different components.

3.3.4 Profiles

The final type of reporting is those qualifications which report a profile of achievement. This type of profile generally includes individual subject results which may be expressed as grades or scores. Award of the diploma is usually a simple statement of whether all requirements have been met rather than being based on aggregation of results or composite scores. The profile may also record of achievement of other requirements which may not be the results of study, such as attendance.

The Queensland Certificate of Education (QCE) is based on an individualised learning plan in which credits can be gained from various sources. These may be either school-based or external. There are also compulsory literacy and numeracy requirements which must be met. There are no grades of achievement and no composite scores. However, the student is also given a Tertiary Entrance Statement which gives their ranking against others in the cohort. This is used for application for higher education.

The South Australian Certificate of Education (SACE) has a similar approach but has less flexibility in the mode of assessment, which must be 30% external. Grades of A-E are given for individual subjects and there are requirements to reach a C grade in certain subjects to meet the requirements of the certificate. As with the QCE, the student is given a Tertiary Entrance Rank.

The Victorian Certificate of Education (VCE) also includes credits from various internal and external sources. The statement of results includes a 'Study Score' based on percentiles which serves the same function as the tertiary ranks of the QCE and the SACE.

The fourth example is the National Certificate of Educational Achievement (NCEA) in New Zealand. This has similarities with the Australian examples in that achievement from a variety of internal and external earns credits towards the award of the diploma. However, there are 3 levels of the diploma and in each of the levels good achievement is endorsed with what in effect amounts to a grade: A (achieved), M (achieved with merit) and E (achieved with excellence). (See example results report in Appendix C.)

3.4 The level of detail reported

The previous section outlined three main approaches to representing results. However within these main methods of representing results by grades, scores or profile reporting there are variations in the amount of detail included in the results reports which are given to students.

The most extensive reporting of scores among those in Appendix A is found with the SAT and ACT tests which are used for entry to Higher Education in the USA. The SAT statement of results gives raw scores, scaled scores, percentiles and means. The raw score report includes a breakdown by question types and the level of difficulty of test items. The statement of results for the ACT tests includes composite scores, subscores and comparison with national and state averages. Benchmark scores are included to enable the student to see college readiness. The report also includes the student's Grade Point Average from their High School Studies. It thus fulfils the function of being a record of achievement and a guide to the student on where results might need to improve for entry to the college of choice.

Qualifications which use profile reporting have the most detail, these have been described in section 3.3.3 above so will not be repeated here in full. They generally give a summary of a range of aspects of a student's educational experience, some of which contribute to the award of the qualification and some which do not. As described in the previous section, they may include grades or scores for components or for individual subject achievement, records of attendance, fulfilment of requirements such as participation in work experience or community service, and a statement of achievement of the relevant overall diploma or certificate. They may give an overall score, grade or level which summarises achievement. In the case of the Australian qualifications described they include a percentile rank for use in applications to higher education.

An example of reporting of grades can be seen in the sample Hong Kong School Certificate results report in Appendix C. This gives a grade for each subject and for sub-components plus composite scores.

3.5 Error and uncertainty

The previous chapter described possible methods of measuring and quantifying measurement error and reliability of scores. It also discussed the implications of different methods of representing results, for example decisions on grading, and the possible implications of these for the meaning of the results given to students. This section discusses these issues further in the light of the assessments which have been classified in Appendix A.

The section reports on ways in which error and uncertainty is conveyed on the actual statements of results which are given to students. In some cases specific examples of reports or full descriptions of the details included have not been obtained and so it is not possible to state with any certainty whether this information is included. Where such information has been found, however, the majority of these assessments do not report error or uncertainty. The exceptions are the SAT and the ACT in the united States. In the case of the ACT this is by means

of a general statement of caution on the results report rather than an individualized estimate of the accuracy of the result:

Your test results are only estimates, not precise measures, of your educational development.

<http://www.actstudent.org/scores/understand/studentreport.html>

The SAT score reports include the range within which the score falls. This is explained as follows on the College Board website:

Your performance is best represented by a score range. The score range is an estimate of how your scores might vary if you were tested many times.

To consider one score better than another, there must be a difference of 60 points between your critical reading and mathematics scores, 80 points between your critical reading and writing scores, and 80 points between your mathematics and writing scores.

Most of the time, your score would fall in a range about 30 to 40 points above or below your true ability. Colleges know this and they receive the score ranges along with your scores.

<http://www.collegeboard.com/student/testing/sat/scores/understanding/scorerange.html>

In this approach to reporting, as discussed in Chapter 2, the score is given along with a band which is based on the standard error of measurement. This may be presented as a graph or as a range of scores, or in some cases both. As described by Goodman and Hambleton (2004), this is a practice in some US state assessments and is also used by some commercial companies in the United States and in some Canadian provincial school assessments. Hambleton and Zenisky (2008) also give an example of this type of reporting in the Literacy and Numeracy National Assessment (LANNA) in Australia.

There are some examples of this type of report in Appendix C and these are included as an illustration of ways in which uncertainty could be reported.

The first example is from a Grade 9 provincial assessment of mathematics in Ontario. In this case the score is given on a band shown graphically and is also shown according to how it matches provincial standards. There is no explanation of the meaning of the band on the report itself. The second example is from a grade 3 reading and literature assessment in the US state of Oregon. In this case the score is given both graphically and as a score range. The meaning of the scale is explained as follows:

The horizontal bars illustrate the typical range of scores for students with a score like yours (also called the standard error of measurement).

This approach can of course only be used with a report which gives a numerical score, although as discussed earlier many systems which represent results by grades also give some form of numerical score which underlies a grade, and it should in most cases be possible to report this as

a band score. An approach which could be used with grades was discussed in Chapter 2, which is to encompass uncertainty by giving more than one grade – for example, by reporting a result as a grade A/B. However, no examples of this have been identified at the time of reporting.

4 Conclusion

This report has discussed a variety of international approaches to representing and reporting results and explaining them to students. This is an area in which it is very difficult to identify best practice and to make recommendations about the lessons which can be learned. Qualifications for certifying the achievement of school leavers and for entry to employment and higher education need to meet standards of good assessment practice and need to be reliable. To this extent there are standards which do not vary internationally. Beyond this, however, they also need to meet the needs of the educational system and of society and here the extent to which international comparison can lead to standards of 'best practice' is more questionable. Assessment practices and the award of qualifications are the result of traditions and the expectations of the educational community and society as well as the more technical considerations of the assessment community. For example, a school-leaving certificate which requires fulfilment of an attendance requirement or of a period of community service, which is the case with some mentioned in this report, may be unacceptable in other countries in which meeting a high standard of academic achievement is seen as more important.

There will always be tensions between the need for reliable measurement, the requirements for learning, the number of places available in higher education, the purposes for which results will be used and public perceptions of what is acceptable. There is also likely to be variation in the amount of information which those who award qualifications are willing to give about the amount of error or uncertainty in their results. As Newton (2005) points out:

*Improving the public understanding of assessment error presents a major challenge. Any campaign designed to increase understanding of the weaknesses of an assessment system runs the risk of simultaneously decreasing confidence in it. In many ways, assessment agencies are trapped between the horns of a dilemma: do they take active steps to improve the **public understanding** of assessment, including assessment error, or do they focus instead on improving **public confidence** in the system?*

(Newton 2005:458)

It is perhaps not surprising that this report has found few examples of reporting of error or uncertainty in high-stakes assessments. Introducing such reporting would risk public loss of confidence and would need to be handled carefully to ensure that it does not lead to misinterpretation by the media and a loss of faith in the assessment system. It would need to be accompanied by education of the media and the public so that they appreciate the existence of measurement error in the assessment process and are better able to interpret assessment results.

5 References

Bergeson, T. (2000). *Reporting on Learning: Washington's School Improvement Strategy*. Olympia, WA: Washington Office of the State Superintendent of Public Instructions.

Clausen-May, T., Vappula, H. and Ruddock, G. (2004). *Progress in Maths 4-14*. London: GL Assessment.

Cresswell, M.J. (1986). 'Examination grades: how many should there be?' *British Educational Research Journal*, **12**, 1, 37–54.

Cresswell, M.J. (1988). 'Combining grades from different assessments: how reliable are the results?' *Educational Review*, **40**, 3, 361–382.

Ebel, R.L. (1969). 'The relation of scale fineness to grade accuracy', *Journal of Educational Measurement*, **6**, 4, 217–221.

Eurydice/Eurybase (2007) *The education system in France*. Online: available: http://eacea.ec.europa.eu/education/eurydice/eurybase_en.php/FR_EN.pdf

Gardner, J. and Cowan, P. (2005). 'The fallibility of high stakes '11-plus' testing in Northern Ireland', *Assessment in Education*, **12**, 2, 145–165.

Gill, T. and Bramley, T. (2008). 'Using simulated data to model the effect of inter-marker correlation on classification consistency', *Research Matter*, **5**, 29–36.

Goodman, D.P. and Hambleton, R.K. (2004). 'Student test score reports and interpretive guides: review of current practices and suggestions for future research', *Applied Measurement in Education*, **17**, 2, 145–220.

Hambleton, R.K and Zenisky, A.L. (2008) 'Reporting Test Scores in More Understandable and Meaningful Ways', ITC Conference Invited Talk, Liverpool, July 14, 2008

Harris, D.J. (2003). 'Reporting and interpreting test results.' In: Wall, J.E. and Walz, G.R. (Eds) *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators*. Austin, TX: Pro-Ed.

Ipsos MORI (2009). *Public Perceptions of Reliability in Examinations*. OfQual. Online: available: http://www.ofqual.gov.uk/files/2009-05-14_public_perceptions_of_reliability.pdf

Jaeger, R.M. (2003). *NAEP Validity Studies: Reporting the Results of the National Assessment of Educational Progress* (Working Paper No.2003-11). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Joint Council for Qualifications (2009) *Post-Results Services Booklet 2008-9* Online: available: <http://www.jcq.org.uk>

Le Métails, J. (2002) *International developments in upper secondary education: Context, provision and issues: INCA Thematic Study No. 8*. Online: available: www.inca.org.uk/pdf/cav_final_report.pdf.

Kleven, T.A. (1979). 'The relation of the scale coarseness to the dependability of marks', *Scandinavian Journal of Educational Research*, **23**, 3, 109–130.

Mitchelmore, M.C. (1981). 'Reporting student achievement: how many grades?' *British Journal of Educational Psychology*, **51**, 218–227.

Newton, P.E. (1996). 'The reliability of marking of General Certificate of Secondary Education scripts: Mathematics and English', *British Educational Research Journal*, **22**, 1, 405.

Newton, P. (2009). 'The reliability of results from national curriculum testing in England', *Educational Research*, **51**, 2, 181–212.

Newton, P., Baird, J., Goldstein, H., Patrick, H. and Tymms, P. (Eds). (2007). *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.

Please, N.W. (1971). 'Estimation of the proportion of examination candidates who are wrongly graded', *British Journal of Mathematical & Statistical Psychology*, **24**, 2, 230–238.

Roeber, E. (2003). 'Steps in the right direction: reporting assessment results to students, parents, school board members, and the media.' In: Wall, J.E. and Walz, G.R. (Eds) *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators*. Austin, TX: Pro-Ed.

Skurnik, L.S. and Nuttal, D.L. (1968). 'Describing the reliability of examinations', *The Statistician*, **18**, 2, 119–128.

Suto, I. and Green, S. (2005). *International Perspective on Vocational Education: What Can We Learn from Each Other?* Cambridge: Cambridge Assessment.

Ward, C. (1980). *Designing a Scheme of Assessment*. Cheltenham: Stanley Thornes. Cited in: Cresswell, M.J. (1988). 'Combining grades from different assessments: how reliable is the results?' *Educational Review*, **40**, 3, 361–382.

Appendix A Classification of assessments

Key

Stage	LS	Lower secondary (15-16)	
	US	Upper secondary (17/18)	
Purpose	GEN	General school assessment	Including general school-leaving certificates or diplomas
	VOC	Vocational qualification	
	TERT	Entry to tertiary education	
	MIX	School-leaving and tertiary entry	
Award	EXT	External body	
	INT	Internal school or teacher assessment	
Mode	MOD	Modular	Modular or unit structure - credit accumulation for final grade
	TERM	Terminal	Final assessment or examinations
	MIX	Mixture of modular and terminal	
Method	CW	Coursework/continuous assessment	May be either internally or externally assessed
	FORM	Formal assessment	May be either internally or externally assessed
	MIX	Mixture of types of assessment	Formal assessment plus coursework or continuous teacher assessment
Assessors	INT	Teacher/school assessment	All assessment is done at school level (possibly with external moderation)
	EXT	External assessment	eg exams set and marked by an external body
	MIX	Mixture of internal/external assessors	
Results	LET	Letter grades	eg grade A, B, C etc.
	NUM	Numerical grades	eg grade 1, 2, 3 etc
	SCORE	Numerical score	Raw score, percentage, standardised score, etc.
	PROF	Profile	May be combination of grades and scores, may include non-learning elements
Structure	CERT	Certificate covering all learning	eg High School diplomas
	DISC	Separate awards for individual subjects	Subject results not combined into overall result
Error reported?	NO	Error/uncertainty not reported	Error/uncertainty not included on statement of results or certificate
	NK	Not known	Information not available at time of reporting
	STAT	Statistical method of representing error	eg reporting score as a band
	TEXT	Caution about interpretation of results	

International Survey of Results Reporting

Country	Qualification	Issuing/ regulating body	Description of assessment						Type of result			Individual results reports		Notes
			stage	purpose	award	mode	method	assessors	results	no of grades	structure	error reported?	Details included	
Australia (New South Wales)	Higher School Certificate (HSC)	Board of Studies New South Wales	US	MIX	EXT	MOD	MIX	MIX	SCORE	-	CERT	NO	Record of Achievement, Results Summary showing eligibility for HSC, University Admissions Index (rank order). HSC Course Reports showing level of achievement and standards reached by other students in course	Courses must be completed over 5-year period. Internal results adjusted according to external exams. Award of certificate requires things other than study - eg successful completion of work experience, course attendance, etc. Score out of 100, based on average of all subjects. Minimum standard at least 50. Performance bands with descriptions for marks over 50. *N = non-completion, decided by principal. 50% internal, 50% external.
Australia (Queensland)	Queensland Certificate of Education (QCE)	Queensland Studies Authority	US	MIX	EXT	MOD	MIX	MIX	PROF	-	CERT	NO	Senior Education Profile including: QCE; senior statement; statement of results; tertiary entrance statement gives rankings	Based on individualised learning plan. Credits from various sources - SB and external. Literacy and numeracy requirements
Australia (South Australia)	South Australian Certificate of Education (SACE)	SACE Board of South Australia	US	MIX	EXT	MOD	MIX	MIX	PROF	-	CERT	NO	Certificate showing achievement of SACE; Record of Learning Achievements with subjects and results; Tertiary Entrance Rank for university entry	Introduced 2009 - first awards 2011. Grades A-E in individual subjects. C grades in some needed to complete SACE. 200 credits needed. 70% internal, 30% external.
Australia (Victoria)	Victorian Certificate of Education (VCE)	Victorian Curriculum and Assessment Authority	US	MIX	EXT	MOD	MIX	MIX	PROF	-	CERT	NO	VCE Data Service (VCEDS) gives schools access to their results and those of other schools. Statement of	Credits from various sources, SB and external including graded VCE units

International Survey of Results Reporting

Country	Qualification	Issuing/ regulating body	Description of assessment						Type of result			Individual results reports		Notes
			stage	purpose	award	mode	method	assessors	results	no of grades	structure	error reported?	Details included	
													results to students. Includes 'study score' based on percentiles, used for tertiary admission.	
Canada (British Colombia)	Graduation Program	Ministry of Education, British Colombia	US	MIX	SB	MIX	MIX		LET*	5	CERT	NK	Transcript of grades - also shows course credits.	Credits can be included from various types of internal or external assessment. For individual courses, 5 passing grades based on scores, minimum score 50. Course credits given for completion with passing grade. Grade Point Average based on course credits weighted by grade. GPA higher than 3 for Honours.
Canada (Ontario)	Ontario Secondary School Diploma	Ministry of Education, Ontario	US	MIX	SB	MIX	MIX	SB	SCORE	-	CERT	NO	Provincial Report Card has transcript of grades, credits, course median, average percentage grade, attendance, learning skills.	Percentage grades, 70% evaluations throughout course, 30% final. Minimum score for credit 50%. 40 hours community involvement also required.
Denmark	Higher Preparatory Examination	Indervisningsministeriet (Danish Ministry of Education)	US	GEN	EXT	TERM	MIX	MIX	NUM	7	DISC	NK	Score	7-point scale for term work and final exams (-3 to 12). Final score is average of term and exam marks.
England	General Certificate of Secondary Education (GCSE)	EdExcel, OCR, AQA*	LS	GEN	EXT	MIX*	MIX	MIX	LET	8	DISC	NO	Grades for each subject. For uninitised GCSEs also given uniform marks (UMS) for components.	*Central specifications from QCA. New format from 2009. Most are terminal assessment, some are uninitised with re-take possibilities. Grades A*-G.

International Survey of Results Reporting

Country	Qualification	Issuing/ regulating body	Description of assessment						Type of result			Individual results reports		Notes
			stage	purpose	award	mode	method	assessors	results	no of grades	structure	error reported?	Details included	
England	General Certificate of Education (A/AS levels)	EdExcel, OCR, AQA*	US	GEN	EXT	MOD	FOR	EXT	LET	5	DISC	NO	Grades for each subject plus uniform marks (UMS) for components.	*Central specifications from QCA. Grades A*- E . A* over 90% of marks
France	Baccalaureat	Ministre de l'Éducation nationale	US	MIX	EXT	TERM*	FOR	EXT	LET	-	CERT	NK		*Some practical subjects have continuous assessment. Average of 10 out of 20 needed to pass. Honours grades: AB 12-13, B 14-15, TB 16+
France	Diplôme National du Brevet des Collèges	Ministre de l'Éducation nationale	LS	GEN	EXT	MIX	MIX	MIX	SCORE	-	CERT	NK		Continuous assessment by schools, external tests in French, maths, history-geography-civics. Need general average of 10 out of 20 to pass
Germany	Abitur	<i>Regional</i>	US	MIX	EXT	TERM	FOR	MIX*	SCORE	-	CERT	NK		*Varies by Lander. Comparability of exam procedures and requirements overseen centrally. Total score 300 points
Hong Kong	Hong Kong Certificate of Education Examination	Hong Kong Examinations and Assessment Authority	LS	GEN	EXT	MIX	MIX	MIX*	LET*	6	DISC	NO	Levels for subject and for components	Internal assessment in growing number of subjects, to be introduced in new HKDSE from 2012. English & Chinese have standards-referenced assessment; Grades 1-5*, U (unclassified). Written descriptors for each level

International Survey of Results Reporting

Country	Qualification	Issuing/ regulating body	Description of assessment						Type of result			Individual results reports		Notes
			stage	purpose	award	mode	method	assessors	results	no of grades	structure	error reported?	Details included	
Hungary	Higher level school-leaving certificate <i>(erettségi vizsga)</i>	National Examination Centre for Public Education and Evaluation (OKEV)	US	MIX	EXT	TERM	FOR	EXT	SCORE	-	CERT	NO	Percentages and description for each subject	
Ireland	Junior Certificate	State Examinations Commission	LS	GEN	EXT	TERM	FOR	EXT	LET	6	DISC	NK	Grade and percentage score	3 levels of exams: Higher, Ordinary, Foundation (last is for Irish, English & maths only). Grades A-F. Set percentage ranges for each grade. Eg 85 or over, grade A; 10-39, grade F
Ireland	Leaving Certificate	State Examinations Commission	US	MIX	EXT	TERM	FOR	EXT	LET	13	DISC	NK	Grade and percentage score	Grades from A1-F. Set percentage ranges for each grade. Eg 90 or over, grade A; 10-24, grade F
Ireland	Leaving Certificate Applied Programme	State Examinations Commission	US	GEN	EXT	MOD	MIX	EXT*	SCORE	-	CERT	NK		Oral examinations of coursework by external examiners. Pass (60-69%, 120-139 credits), Merit (70-84%, 140-169 credits), Distinction (85-100%, 170-200 credits). Record of Credits if not completed or if below 120 credits

International Survey of Results Reporting

Country	Qualification	Issuing/ regulating body	Description of assessment						Type of result			Individual results reports		Notes
			stage	purpose	award	mode	method	assessors	results	no of grades	structure	error reported?	Details included	
Korea	College Scholastic Ability Test (CSAT)	Korea Institute of Curriculum and Examination	US	TERT	EXT	TERM	FOR	EXT	SCORE	-	DISC	NO	Standard score, percentile rank, stanine for each subject.	
Netherlands	Leaving Certificate (VMBO)	Ministry of Education	LS	VOC	EXT	TERM	FOR	INT	NUM	10	CERT	NK		Exams are a mixture of internal and external. Some subjects only school-based. School exams are approved by inspectorate and results submitted before external exams. External exams marked by teachers. Grades 6-10 are passing grades
Netherlands	Leaving Certificate (HAVO)	Ministry of Education	US	GEN	EXT	MIX	MIX	MIX	NUM	10	CERT	NO	Certificate gives overall grade and transcript of attainment in various elements.	Exams are a mixture of internal and external. Some subjects only school-based. School exams are approved by inspectorate and results submitted before external exams. Internal element is portfolio-based. Grades 6-10 are passing grades
Netherlands	Leaving Certificate (VWO)	Ministry of Education	US	TERT	EXT	MIX	MIX	MIX	NUM	10	CERT	NO	Certificate gives overall grade and transcript of attainment in various elements.	Exams are a mixture of internal and external. Some subjects only school-based. School exams are approved by inspectorate and results submitted before external exams. School-based element is portfolio-based. Grades 6-10 are

International Survey of Results Reporting

Country	Qualification	Issuing/ regulating body	Description of assessment						Type of result			Individual results reports		Notes
			stage	purpose	award	mode	method	assessors	results	no of grades	structure	error reported?	Details included	
														passing grades
New Zealand	National Certificate of Educational Achievement	New Zealand Qualifications Authority	US	GEN	EXT	MIX	MIX	MIX	PROF	-	CERT	NO	NQF results notices sent annually with internal and external results. Record of Achievement gives overall results plus transcript of results in all subjects and sub-components.	Standards-based - 3 levels. At each level, achievement of sub-components reported. A (achieved), M (achieved with merit) E (achieved with excellence).
Sweden	Swedish Scholastic Assessment Test	Department of Educational Measurement, Umea University	US	TERT	EXT	TERM	FOR	EXT	SCORE	-	CERT	NK	Score	Multiple choice, OMR. Seen as 'second chance' for those whose average marks from school do not give them enough for university. Standardised score between 0.0 and 2.0.
USA	Advanced Placement Test (AP)	College Board	US	TERT	EXT	TERM	FOR	EXT	NUM	5	DISC	NO	Grade	Grade 3 or above considered enough to give course credits or placement on advanced programmes although colleges set their own requirements
USA	Scholastic Assessment Test (SAT)	College Board	US	TERT	EXT	TERM	FOR		SCORE	-	CERT	YES	Raw scores, scaled scores, percentiles, means, score range	Scale 200-800 for reasoning test, 20-80 for subject tests. Raw score report includes breakdown by question types, level of difficulty

International Survey of Results Reporting

Country	Qualification	Issuing/ regulating body	Description of assessment						Type of result			Individual results reports		Notes
			stage	purpose	award	mode	method	assessors	results	no of grades	structure	error reported?	Details included	
USA	ACT test	ACT	US	TERT	EXT	TERM	FOR	EXT	SCORE	-	CERT	TEXT	Composite scores and subscores, comparison with national and state percentages at or below score.	Composite scores are average of 4 components, range 1-36. Benchmark scores given to show college readiness. Report includes GPA based on High School grades

Appendix B Search Strategy

Databases

Search strategies for the bibliographic databases were developed using terms from the relevant thesauri and/or free-text searching (denoted by ft). The keywords used in the searches, together with a brief description of each of the databases searched, are outlined below.

Australian Education Index (AEI) (searched via Dialog Datastar 13/03/09)

AEI is Australia's largest source of education information covering reports, books, journal articles, online resources, conference papers and book chapters.

SET 1

- #1 educational testing
- #2 educational assessment
- #3 testing
- #4 examinations
- #5 tests
- #6 high stakes tests (ft)
- #7 college entrance examinations
- #8 examination results
- #9 test results
- #10 certification
- #11 grades (ft)
- #12 grades scholastic (ft)
- #13 letter grades (ft)
- #14 number grades (ft)
- #15 percentiles (ft)
- #16 scores

International Survey of Results Reporting

- #17 scoring (ft)
- #18 test scores (ft)
- #19 test scoring (ft)
- #20 examination scores (ft)
- #21 examination scoring (ft)
- #22 marking (ft)
- #23 test marking (ft)
- #24 examination marking (ft)
- #25 profiles
- #26 descriptions (ft)
- #27 narrative evaluations (ft)
- #28 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27

SET 2

- #29 reports
- #30 reporting (ft)
- #31 #29 or #30
- #32 #28 and #31
- #33 #32 and measurement error (ft)
- #34 reporting student achievement
- #35 SAT (ft) or scholastic achievement test (ft)
- #36 #35 and #31
- #37 tertiary entrance rank (ft) or TER (ft)
- #38 #37 and #31
- #39 equivalent national tertiary entrance rank (ft) or ENTER (ft)

International Survey of Results Reporting

- #40 #39 and #31
- #41 universities admission index (ft) or UAI (ft)
- #42 #41 and #31
- #43 #31 and overall position (ft)
- #44 #31 and swesat (ft)
- #45 #31 and baccalareat (ft)

British Education Index (BEI) (searched via Dialog Datastar 13/03/09)

BEI provides information on research, policy and practice in education and training in the UK. Sources include over 300 journals, mostly published in the UK, plus other material including reports, series and conference papers.

SET 1

- #1 educational testing
- #2 educational assessment
- #3 testing
- #4 examinations
- #5 tests
- #6 high ADJ stakes ADJ tests (ft)
- #7 college ADJ entrance ADJ examinations (ft)
- #8 examination results
- #9 test results
- #10 certification
- #11 grades (ft)
- #12 grades ADJ scholastic (ft)
- #13 letter ADJ grades (ft)
- #14 number ADJ grades (ft)

International Survey of Results Reporting

- #15 percentiles
- #16 scores
- #17 scoring
- #18 test ADJ scores
- #19 test ADJ scoring
- #20 examination ADJ scores
- #21 examination ADJ scoring
- #22 marking
- #23 test marking (ft)
- #24 examination marking (ft)
- #25 profiles
- #26 descriptions
- #27 narrative ADJ evaluations
- #28 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27

SET 2

- #29 reports
- #30 reporting (ft)
- #31 #29 or #30
- #32 #28 and #31
- #33 #32 and measurement error (ft)
- #34 SAT (ft) or scholastic achievement test (ft)
- #35 #31 and #34
- #36 tertiary entrance rank (ft) or TER (ft)
- #37 #31 and #36
- #38 equivalent national tertiary entrance rank (ft) or ENTER (ft)

- #39 #31 and #38
- #40 #31 and (universities admission index (ft) or UAI (ft))
- #41 #31 and Overall position (ft)
- #42 #31 and Swesat (ft)
- #43 #31 and baccalaureat (ft)

British Education Index (Free Collection) (searched via the WWW 18/03/09)

The free collections search interface of the British Education Index (BEI) (formerly the British Education Internet Resource Catalogue) includes access to a range of freely available internet resources as well as records for the most recently indexed journal articles not yet included in the full BEI subscription database.

SET 1

- #1 achievement tests or certification or educational assessment or educational testing or examination results or examinations or marking (scholastic) or profiles or pupil evaluation or scores or scoring or scottish certificate of education or sixteen plus examinations or student evaluation or summative evaluation or test results or testing or tests
- #2 high stake tests (ft)
- #3 college entrance examinations (ft)
- #4 grades (ft)
- #5 profiles (ft)
- #6 #1 or #2 or #3 or #4 or #5

SET 2

- #7 report (ft)
- #8 reporting (ft)
- #9 #7 or #8
- #10 #6 and #9

CERUK Plus (searched online 25/03/09)

The CERUK Plus database provides access to information about current and recently completed research, PhD level work and practitioner research in the field of education and children's services. Searching was limited to completed works only.

SET 1

#1 Tests or Testing or Assessment or Examinations or Test results or Test scores or Certification or Grading or Grades or Profiles or Marking or high stakes tests (ft) college entrance examinations (ft) examination results (ft) or test results (ft) or letter grades (ft) or number grades (ft) or scores (ft) or scoring (ft) or descriptions (ft) or narrative evaluations (ft)

SET 2

#2 reports or reporting (ft)

#3 #1 and #2

Education Resources Information Center (ERIC) (searched via the WWW on 25/03/09)

The ERIC database is sponsored by the US Department of Education to provide extensive access to education-related literature.

SET 1

#1 educational testing

#2 educational assessment

#3 testing

#4 college entrance examinations

#5 exit examinations

#6 tests

#7 high stakes tests

#8 test results

#9 certification

#10 grades (scholastic)

#11 grade equivalent scores

International Survey of Results Reporting

- #12 scores
- #13 scoring
- #14 grading
- #15 profiles
- #16 letter grades (ft)
- #17 number grades (ft)
- #18 percentiles (ft)
- #19 test scoring (ft)
- #20 examination scores (ft)
- #21 examination scoring (ft)
- #22 test marking (ft)
- #23 descriptions (ft)
- #24 narrative evaluations (ft)
- #25 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24

SET 2

- #26 report card or reporting (ft)
- #27 #25 and #26
- #28 #27 and measurement error (ft)
- #28 SAT (ft) or scholastic achievement test (ft)
- #29 #26 and #28
- #29 tertiary entrance rank (ft) or TER (ft)
- #30 #26 and #29
- #31 equivalent national tertiary entrance rank (ft) or ENTER (ft)
- #32 #26 and #31

- #33 universities admission index (ft) or UAI (ft)
- #34 #26 and #33
- #35 #26 and overall position
- #36 #26 and Swesat
- #37 #26 and baccalaureat

PsycINFO (searched via Ovid 13/03/09)

PsycINFO contains references to the psychological literature including articles from over 1,300 journals in psychology and related fields, chapters and books, dissertations and technical reports. NB. Free text searching was employed because the thesaurus vocabulary in this database did not include the specialist keywords relating to educational testing and assessment.

- #1 educational testing (ft)
- #2 educational assessment (ft)
- #3 testing (ft)
- #4 examinations (ft)
- #5 tests (ft)
- #6 high stakes tests (ft)
- #7 college entrance examinations (ft)
- #8 examination results (ft)
- #9 test results (ft)
- #10 certification (ft)
- #11 grades (ft)
- #12 grades scholastic (ft)
- #13 letter grades (ft)
- #14 number grades (ft)
- #15 percentiles (ft)
- #16 scores (ft)

International Survey of Results Reporting

- #17 scoring (ft)
- #18 test scores (ft)
- #19 test scoring (ft)
- #20 examination scores (ft)
- #21 examination scoring (ft)
- #22 marking (ft)
- #23 test marking (ft)
- #24 examination marking (ft)
- #25 profiles (ft)
- #26 descriptions (ft)
- #27 narrative evaluations (ft)
- #28 #1 or #2 or #3 or #4 or #5 or # 6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27
- #29 reporting (ft)
- #30 report (ft)
- #31 #29 or #30
- #32 #28 and #31
- #33 scholastic achievement test
- #34 tertiary entrance rank
- #35 universities admission index
- #36 overall position
- #37 swesat
- #38 baccalaureat

Websites (Searched March 2009)

The following international websites – predominantly education ministries or national curriculum and assessment organisations – were browsed for appropriate literature including qualification manuals and guides for students and parents. The INCA and Eurybase databases were also consulted.

Country/state`	Organisation and URL
Australia (Queensland)	Queensland Studies Authority (QSA) www.qsa.qld.edu.au/
Australia (Victoria)	Victorian Curriculum and Assessment Authority (VCAA) www.vcaa.vic.edu.au
Australia (South Australia)	SACE Board of South Australia www.decs.sa.gov.au/curric/pages/NumeracySciences/SACE/
Australia (New South Wales)	Board of Studies New South Wales www.boardofstudies.nsw.edu.au/
Canada (British Columbia)	BC Ministry of Education www.bced.gov.bc.ca
Canada (Ontario)	Ministry of Education Ontario www.edu.gov.on.ca
Denmark	Danish Ministry of Education www.eng.uvm.dk/
Finland	Finland’s Ministry of Education website www.minedu.fi/OPM National Board of Education www.oph.fi
France	Ministère de l’Education Nationale

	www.education.gouv.fr
Germany	Federal Ministry of Education and Research www.bmbf.de German Education Server www.bildungserver.de
Greece	Ministry of Education and Religious Affairs www.ypepth.gr/en_ec_home.htm
Hong Kong	Hong Kong Examinations and Assessment Authority www.hkeaa.edu.hk
Hungary	Ministry of Education and Culture www.okm.gov.hu/main.php?folderID=137 National Institute for Public Education www.oki.hu/english.php National Centre for Educational Evaluation and Examination www.oh.gov.hu
Ireland	State Examinations Commission www.examinations.ie National Council for Curriculum and Assessment www.ncca.ie
Italy	Ministry of Public Education www.pubblica.istruzione.it
Japan	Ministry of Education, Culture, Sports, Science and Technology www.mext.go.jp/english -
Korea	Ministry of Education, Science and Technology http://english.mest.go.kr/ Korea Institute for Curriculum and Evaluation (KICE) www.kice.re.kr
Malaysia	Malaysian Examinations Council www.mpm.edu.my Ministry of Education Malaysia www.moe.gov.my Blog on Malaysian examinations www.malaysia-students.com/

International Survey of Results Reporting

Netherlands	Ministry of Education, Science and Culture www.minocw.nl/english National Institute for Educational Measurement CITO www.cito.com National Institute for Curriculum Development (SLO) www.slo.nl
Northern Ireland	Northern Ireland Council for the Curriculum, Examinations and Assessment (CCEA) www.ccea.org.uk
Norway	Ministry of Education and Research www.regjeringen.no/en Norwegian Directorate for Education and Training www.udir.no/
Poland	Central Examination Commission www.cke.edu.pl/ Ministry of Education www.men.gov.pl/
Scotland	Scottish Qualifications Authority (SQA) www.sqa.org.uk
Spain	Ministry of Education www.mec.es
Sweden	The Swedish National Agency for Education /www.skolverket.se
USA	For SAT and AP – College Board www.collegeboard.com/ For ACT http://www.act.org/
Wales	National Assembly for Wales www.wales.gov.uk Welsh Bac website www.wbq.org.uk
Eurybase	The database on education systems in Europe

International Survey of Results Reporting

<http://eacea.ec.europa.eu/portal/page/portal/Eurydice/EuryPresentation>

INCA

International Review of Curriculum and Assessment Frameworks Internet Archive (INCA)

www.inca.org.uk

Appendix C Example results reports

The following sample reports are included to illustrate some of the methods of results reporting that have been discussed. We would like to thank each of the following organisations for allowing their material to be reproduced.

Hong Kong Examinations and Assessment Authority. This sample copy of the Hong Kong Certificate of Education Examination Results Notice is reproduced by permission of the Hong Kong Examinations and Assessment Authority.

EQAO. The Ontario Grade 9 assessment of mathematics individual student report has been reproduced by permission of the Education Quality and Accountability Office, Ontario.

Oregon Department of Education. The Oregon statewide assessment of Reading and Literature sample report has been reproduced by permission of the Oregon Department of Education.

New Zealand Qualifications Authority. The example Record of Achievement is reproduced by permission of the New Zealand Qualifications Authority (www.nzqa.govt.nz).



香港考試及評核局
HONG KONG EXAMINATIONS AND ASSESSMENT AUTHORITY
香港中學會考成績通知書
HONG KONG CERTIFICATE OF EDUCATION EXAMINATION RESULTS NOTICE

9000538

本成績通知書於會考證書印發前發出，香港考試及評核局有權修正其內容。本文件並非「會考證書」，不能作「會考證書」使用。
 This results notice is issued before the certificate is awarded and the Hong Kong Examinations and Assessment Authority reserves the right to correct it. This document is not a certificate and must not be regarded as such.

科目 Subject	科目等級 Subject Level / Grade	分部等級 Component Level / Grade
CHINESE LANGUAGE • READING • WRITING • LISTENING • SPEAKING • INTEGRATED SKILLS	5 (Five)	5* (Five*) 4 (Four) 3 (Three) 5* (Five*) 4 (Four)
ENGLISH LANGUAGE • READING • WRITING • LISTENING AND INTEGRATED SKILLS • SPEAKING	4 (Four)	4 (Four) 5 (Five) 3 (Three) 5 (Five)
PUTONGHUA • LISTENING • TRANSCRIPTION • ORAL • LINGUISTIC KNOWLEDGE AND LANGUAGE USAGE	E (e)	E (e) F (f) D (d) E (e)
CHINESE HISTORY	C (c)	
GEOGRAPHY	D (d)	
ECONOMICS	E (e)	
MATHEMATICS	D (d)	
COMPUTER AND INFORMATION TECHNOLOGY • THEORY • COURSEWORK	D (d)	D (d) E (e)
SCIENCE AND TECHNOLOGY • THEORY • INDEPENDENT STUDY	E (e)	E (e) B (b)

學校
School
1234 ABC COLLEGE

年份
Year
2009

考生編號
Candidate No.
1234567

註冊編號
Reference No.
1234 0101

性別
Sex
M

香港身分證號碼/旅遊證件號碼
Hong Kong Identity Card No./
Travel Document No.
A123456(7)

與考科目總數
No. of Subjects Sat
09

考生姓名及自修生地址
Name of Candidate and
Address of Private Candidate
CHAN TAI MAN

SAMPLE

註釋見背頁 Explanatory notes are printed overleaf

香港特別行政區政府教育局
The Government of the Hong Kong Special Administrative Region
Education Bureau

中六入學取錄紙 - 學校專用
S6 Admission Slip - for School Use Only

9000538

	年份 Year	考生編號 Candidate No.	考生姓名 Name of Candidate	香港身分證號碼 H.K. Identity Card No.	性別 Sex
	2009	1234567	CHAN TAI MAN	A123456(7)	M
成績最佳六科的總積點 Total grade points of the best six subjects	16		成績最佳四科的總積點 (中國語文及英國語文除外) Total grade points of the best four subjects (excluding Chinese Language and English Language)	09	成績最佳三科的總積點 (中國語文及英國語文除外) Total grade points of the best three subjects (excluding Chinese Language and English Language)
考獲 E 級或以上的科目總數 Number of subjects at grade E or above	07		備註：在用作中六入學時，中國語文科及英國語文科的第一至五*級會相應化為 0 至 5 積點 Note: Level 1 to 5 * for Chinese Language and English Language will be converted into points 0 to 5 accordingly for S6 admission purpose		07

註釋
中國語文科及英國語文科
水平參照成績匯報

自2007年開始，香港中學會考中國語文科及英國語文科採用「水平參照」模式匯報成績，即參照一套水平標準來匯報考生表現等級。

考生的成績會以五個表現等級（1 至 5）標示，第 5 等級為最高，第 1 等級為最低。成績低於第 1 等級會標示為「不予評級」（U）。在第 5 等級之內，成績最佳的考生可獲「5*」等級成績。

香港中學會考中國語文科及英國語文科的成績，等同於英國劍橋大學國際普通中學教育文憑試（IGCSE）成績：

香港中學會考等級	5*	5	4	3
IGCSE等級	A*	A	B	C

各能力等級附有等級描述，以說明有關等級的典型學生的能力所及，讓持份者清晰了解各等級考生所達致的知識和能力水平。有關各等級描述，請瀏覽考評局網頁 (<http://www.hkeaa.edu.hk>)。



香港考試及評核局

2009年8月

Explanatory Note
Chinese Language and English Language
Standards-referenced Reporting

With effect from 2007, standards-referenced reporting has been adopted for the two language subjects in the HKCEE: Chinese Language and English Language. Candidates' levels of performance are reported using levels with reference to a set of standards for the subject.

There are five levels of performance (1-5), with Level 5 being the highest and Level 1 being the lowest. Performance below Level 1 is designated as "Unclassified" (U). Level 5 candidates with the best performance will have their results annotated with the symbol *.

The HKCEE results in Chinese Language and English Language are recognized as equivalent to the International General Certificate of Secondary Education (IGCSE) results as follows:

Levels attained in HKCEE	5*	5	4	3
Grades attained in IGCSE	A*	A	B	C

A set of written descriptors has been developed for each of the five levels to describe the typical performance of candidates at that level. The descriptors enable stakeholders to understand explicitly the various standards achieved by candidates in terms of knowledge and skills. Details of the level descriptors are available from the Authority's website (<http://www.hkeaa.edu.hk>).



HONG KONG EXAMINATIONS & ASSESSMENT AUTHORITY

August 2009

INDIVIDUAL STUDENT REPORT

Grade 9 Assessment of Mathematics, 2004–2005

STUDENT INFORMATION

Student Name: SAMPLE, SAMPLE
 Student Number: 000 00000 00 000000 00 00 00
 Ontario Education Number (OEN): 123-456-789
 School Name: SAMPLE SS (000000)
 School Board: SAMPLE District School Board (00000)

Program Information:
 Academic
 First semester

ACHIEVEMENT SUMMARY

ONTARIO CURRICULUM ACHIEVEMENT LEVELS

Below Level 1	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4
Insufficient achievement of curriculum expectations	Much below provincial standard 50–59%	Approaches provincial standard 60–69%	Provincial standard 70–79%	Surpasses provincial standard 80–100%

OVERALL ACHIEVEMENT LEVEL

answered questions: 47 of 48



For participating students, the ■ symbol represents the student's achievement level on this assessment.

“Answered questions” refers to the number of questions the student answered in relation to the total number of questions on the assessment. Unanswered questions are treated as incorrect in the calculation of the student's results.

- LEVEL 4:** The student has demonstrated a very high to outstanding level of achievement. Achievement *surpasses* the provincial standard.
- LEVEL 3:** The student has demonstrated a high level of achievement. Achievement is *at* the provincial standard.
- LEVEL 2:** The student has demonstrated a moderate level of achievement. Achievement is *below, but approaching*, the provincial standard.
- LEVEL 1:** The student has demonstrated a passable level of achievement. Achievement is *much below* the provincial standard.
- BELOW L1:** The student has not demonstrated sufficient achievement of curriculum expectations.
- NO DATA:** EQAO did not receive completed assessment booklets for this student.
- EXEMPT:** The student was exempted from the assessment.

The result reported here is a snapshot of this student's achievement on the days of the assessment and is only one indication of how this student is learning the mathematics curriculum.

This report contains personal information that is protected under the *Freedom of Information and Protection of Privacy Act*.

INFORMATION FOR STUDENTS AND PARENTS

This Individual Student Report contains results from the province-wide Grade 9 Assessment of Mathematics, 2004–2005.

Students in first-semester mathematics courses wrote the assessment in January; students in second-semester or full-year mathematics courses wrote the assessment in May/June. Students in applied and academic programs wrote different versions of the assessment.

UNDERSTANDING THE RESULTS

EQAO assessments objectively measure how well students are meeting the provincial curriculum standards. The assessments complement the information teachers collect throughout the year from classroom assessments, daily observations and conversations with parents.

HOW TO READ THE REPORT

Achievement Levels

The Ontario Curriculum identifies four levels of achievement that indicate how well students are meeting the provincial expectations in mathematics up to the end of Grade 9. The four achievement levels used in this report are aligned with those levels.

Provincial Standard

The Ontario Curriculum establishes Level 3 as the provincial standard.

OVERALL ACHIEVEMENT

The overall achievement level provides a general picture of a student's achievement relative to the curriculum expectations in mathematics. The overall level is represented by the ■ symbol. The shaded line extending from this symbol shows the approximate range of the student's overall achievement in mathematics.

EQAO ensures greater accountability in Ontario's public education system by providing students, parents, educators and the public with information on students' achievement, measured against curriculum standards. Ontario's large-scale assessment program is one important element of a culture of capacity building and learning improvement.

Record of Achievement



Andrew Steven Learner
Address 1
Address 2
Address 3
Address 4 Post Code

SAMPLE

NSN: 101010101
Issued: 15th April 2007

Qualification and Achievement Summary

National Qualifications Framework Registered Qualifications

	Date
National Certificate of Educational Achievement (Level 3) achieved with merit	31/12/2006
National Certificate of Educational Achievement (Level 2)	08/04/2006
National Certificate of Educational Achievement (Level 1) achieved with excellence	08/04/2005

Scholarship 2006

Scholarship in Classical Studies (2006)	20/12/2006
---	------------

New Zealand University Entrance

University Entrance	9/04/2007
---------------------	-----------

Standards Achieved

Each standard can have a range of results including "A" for Achieved, "M" for Achieved with Merit, or "E" for Achieved with Excellence. As some standards can not be awarded all of these results the result code is shown in bold where a candidate has reached the maximum possible result for that standard.

English

Level 3		Credits	Result	Date
90721	Respond critically to written text(s) studied	3	M	18/12/2006
90723	Respond critically to oral or visual text studied	3	A	18/12/2006
90724	Read and respond critically to unfamiliar prose and poetry texts	3	E	17/12/2006
90722	Respond critically to Shakespearean drama studied	3	A	15/12/2006
90725	Construct and deliver an oral presentation	4	E	27/10/2006
90720	Produce an extended piece of writing in a selected style	4	E	15/09/2006
90726	Complete independent research on a language or literature topic and present conclusions in writing	4	E	15/09/2006
Level 2		Credits	Result	Date
90377	Analyse extended written text(s)	3	M	16/12/2005
90379	Analyse a visual or oral text	3	E	16/12/2005
90378	Analyse short written texts	3	M	12/12/2005
90380	Read unfamiliar texts and analyse the ideas and language features	3	M	10/12/2005
12905	Read an inclusive variety of written texts and record the reading experience	4	A	04/11/2005
90374	Deliver a presentation using oral and visual language techniques	3	M	04/11/2005
90375	Produce crafted and developed creative writing	3	E	04/11/2005
90376	Produce crafted and developed formal transactional writing	3	M	04/11/2005
Level 1		Credits	Result	Date
90054	Read, study and show understanding of extended written text(s)	2	E	15/12/2004
90056	View/listen to, study and show understanding of a visual or oral text	2	E	14/12/2004
90055	Read, study and show understanding of a number of short written texts	2	E	13/12/2004
90057	Read and show understanding of unfamiliar texts	3	E	13/12/2004
90053	Produce formal writing	3	E	11/12/2004
8808	Read an inclusive range of written texts and record the reading experience	3	A	30/10/2004
90052	Produce creative writing	3	E	30/10/2004

Record of Achievement

Andrew Steven Learner

NSN: 101010101

English

Level 1		Credits	Result	Date
90058	Deliver a speech in a formal situation	3	M	30/10/2004
90059	Produce a media or dramatic presentation	3	E	30/10/2004
90060	Research, organise and present information	3	M	30/10/2004

Mathematics

Level 2		Credits	Result	Date
90285	Draw straightforward non-linear graphs	3	M	19/12/2005
90286	Find and use straightforward derivatives and integrals	4	A	16/12/2005
90287	Use coordinate geometry methods	2	M	10/12/2005
5251	Choose and apply trigonometric methods to solve problems involving lengths and angles	3	A	08/11/2005
5253	Sketch and describe graphs	3	A	08/11/2005

Level 1		Credits	Result	Date
90148	Sketch and interpret graphs	3	A	15/12/2004
90151	Solve straightforward number problems in context	3	M	13/12/2004
90152	Solve right-angled triangle problems	2	M	13/12/2004
90147	Use straightforward algebraic methods and solve equations	4	M	11/12/2004
90153	Use geometric reasoning to solve problems	2	A	11/12/2004
5238	Draw and describe graphs involving two variables	4	A	27/10/2004
5235	Use strategies to solve number problems	2	A	15/09/2004
5236	Use Pythagoras' Theorem and trigonometry to solve problems involving right-angled triangles	2	A	27/08/2004
8489	Solve problems which require calculation with whole numbers	2	A	27/08/2004
8490	Solve problems using calculations with numbers expressed in different forms	2	A	27/08/2004
8491	Read and interpret information presented in tables and graphs	2	A	27/08/2004
90149	Solve problems involving measurement of everyday objects	3	M	23/06/2004
90150	Use geometric techniques to produce a pattern or object	2	E	23/06/2004

Music

Level 1		Credits	Result	Date
90016	Identify and describe fundamental materials of music	2	E	20/12/2004
90014	Compose pieces of music	6	E	11/10/2004
90017	Demonstrate knowledge of music works	4	E	11/10/2004
90012	Perform contrasting music as a featured soloist	6	E	10/09/2004
90013	Perform music as a member of a group	3	M	10/09/2004

Science

Level 2		Credits	Result	Date
90308	Describe the nature of structure and bonding in different substances	4	A	19/12/2005
90310	Describe thermochemical and equilibrium principles	5	A	18/12/2005
90306	Carry out an acid-base volumetric analysis	3	E	07/11/2005

Level 1		Credits	Result	Date
90185	Demonstrate understanding of electricity and magnetism	5	E	22/12/2004
90172	Describe atomic structure and bonding	3	E	16/12/2004
90183	Demonstrate understanding of mechanics in one dimension	5	E	15/12/2004
90648	Describe properties and reactions of carbon and its compounds	3	E	14/12/2004

Record of Achievement

Andrew Steven Learner

NSN: 101010101

Science

Level 1		Credits	Result	Date
90171	Describe chemical reactions	4	E	13/12/2004
90182	Demonstrate understanding of wave and light behaviour	5	E	12/12/2004
90163	Describe the transfer of genetic information	3	E	30/11/2003
90168	Describe how humans use and are affected by micro-organisms	2	E	30/11/2003
90189	Describe properties and reactions of groups of related substances	5	E	30/11/2003
90190	Describe rocks and minerals	3	E	30/11/2003
90191	Demonstrate an understanding of physical systems	5	E	30/11/2003
90192	Describe spatial relationships in astronomy and their effects on space exploration	2	E	30/11/2003
90187	Research, with direction, how science and technology are related	2	M	25/07/2003
90186	Carry out a practical science investigation with direction	4	E	17/06/2003

Social Science Studies

Level 3		Credits	Result	Date
90658	Examine a significant historical situation in the context of change, in an essay	5	E	20/12/2006
90656	Analyse and evaluate evidence in historical sources	5	M	19/12/2006
90657	Examine a significant decision made by people in history, in an essay	5	M	19/12/2006
90511	Explain a passage or passages from a work of classical literature in translation	6	E	18/12/2006
90513	Explain in essay format an aspect of the classical world	6	E	17/12/2006
90602	Explain the relationship between a media genre and society	4	M	16/12/2006
90512	Explain a work or works of classical art	6	E	14/12/2006
90600	Explain how meaning is created in media texts	4	E	02/11/2006
90605	Develop a production schedule and prepare to produce a media product	2	E	02/11/2006
90606	Create a media product using appropriate media technology	6	E	02/11/2006
90779	Investigate an aspect of media and explain its significance for New Zealand	4	M	02/11/2006
90514	Complete independent research on an area of the classical world	6	E	18/09/2006
90654	Plan and carry out independent historical research	4	E	18/09/2006
90655	Communicate and present historical ideas clearly to show understanding of an historical context	5	E	18/09/2006
7474	Investigate how news stories in the media are selected and packaged	4	A	21/05/2006
90604	Complete and justify a concept and treatment for a media product	2	E	21/05/2006
Level 2		Credits	Result	Date
90247	Examine a passage from a work of classical literature in translation	5	M	19/12/2005
90248	Examine a work of classical art and/or architecture	5	E	19/12/2005
90467	Examine evidence in historical sources	4	M	19/12/2005
90249	Examine in essay format an aspect of the classical world	5	M	18/12/2005
90469	Examine how a force or movement in an historical setting influenced peoples lives, in an essay	4	A	18/12/2005
90470	Examine individual or group identity in an historical setting, in an essay	4	E	18/12/2005
90468	Examine perspectives and responses of, and demonstrate empathy for, people in an historical setting	4	M	11/11/2005
90250	Complete an independent examination of an area of classical studies	5	M	03/11/2005
90251	Communicate knowledge of an aspect of the classical world	4	A	03/11/2005
90465	Plan and carry out an historical inquiry	4	M	03/11/2005

Record of Achievement

Andrew Steven Learner

NSN: 101010101

Social Science Studies

Level 2 Credits Result Date

90466	Communicate historical ideas to demonstrate understanding of an historical context	4	A	03/11/2005
-------	--	---	---	------------

Level 1 Credits Result Date

90214	Describe experiences that led to the shaping of the identity of New Zealanders	4	A	22/12/2004
90213	Describe an historical development, in an essay	4	A	15/12/2004
90211	Interpret historical sources	4	M	13/12/2004
90210	Communicate historical ideas	4	E	13/06/2004
90209	Carry out an historical investigation	4	M	06/05/2004

Statistics and Probability

Level 2 Credits Result Date

90288	Select a sample and use this to make an inference about the population	3	M	08/11/2005
90289	Simulate probability situations, and apply the normal distribution	2	M	08/11/2005

Level 1 Credits Result Date

90194	Determine probabilities	2	M	12/12/2004
5242	Determine probabilities	2	A	02/11/2004
90193	Use statistical methods and information	3	M	23/06/2004

Visual Arts

Level 3 Credits Result Date

90667	Produce original work within painting to show extensive knowledge of art-making methods and ideas	14	A	13/12/2006
90659	Research and analyse approaches within established painting practice	4	E	15/08/2006

Level 2 Credits Result Date

90478	Generate and develop ideas using drawing processes and procedures in printmaking practice	6	E	04/11/2005
90473	Research and document methods and ideas in the context of a drawing study in Sculpture	6	E	25/10/2005
90233	Research and document methods and ideas in the context of a drawing study in painting	6	A	27/05/2005

Level 1

		Credits	Result	Date
90020	Generate and develop ideas in making artworks	12	E	18/11/2004
90021	Extend ideas in media and techniques to produce new work	4	M	11/11/2004
90019	Use drawing processes and procedures	5	E	26/05/2004

Signature
Chief Executive

First published by The Office of Qualifications and Examinations Regulation in 2010

© Qualifications and Curriculum Authority (2010)

Ofqual is part of the Qualifications and Curriculum Authority (QCA). QCA is an exempt charity under Schedule 2 of the Charities Act 1993.

