

Sweden

Structure of schooling

Students in Sweden start school at the age of 7, after one or more years in kindergarten, and follow a nine-year compulsory 'basic' program, although almost all (around 95%) students choose to stay on at school for another three years to complete 'upper secondary' education. Of the students in upper secondary education, approximately half pursue vocational programs, and half opt for academic programs, and those pursuing academic programs take either a natural sciences/technical or a social sciences/arts pathway (approximately 40% of the students taking academic programs choose natural or technical science and 60% choose arts or social sciences). All students following the natural science pathway study Biology, Chemistry, Physics, and Mathematics. There are five mathematics courses offered in upper secondary school. Course A, the least demanding, is taken by all students in upper secondary school, after which students may be required to take other courses, depending on the program they are following.

National assessment

There are national assessments at the end of the 5th and 9th years of school, when the students are aged 11-12 and 15-16 respectively, and during the last three years of school. At the end of year 5, the assessments are in Swedish and Mathematics, at the end of year 9, in Swedish, English and Mathematics, and in year 11 or 12, the students are assessed in the courses they have taken in upper secondary education. The assessments at the end of year 5 are reported to parents only in qualitative terms (teachers are explicitly discouraged from attempting to report marks or grades), while those at the end of years 9 and those taken during upper secondary education are graded on a four-point scale: fail, pass, pass with distinction and pass with special distinction.

What makes the Swedish system unusual is the way that these grades are derived. For some courses (for example course A in Mathematics in upper secondary school) there are national tests for all students. The tests are compulsory in that all schools must set the tests, and students are expected to take them when set, but if a student is absent on the date of the test, there is no requirement for the student to take the test on another occasion. Based on these tests students are awarded a test-grade on the four-point scale. The intention is that the teachers should use these grades as one indicator among many to assign course-grades to their students. This means that even though national assessments in the mentioned 'core subjects' are compulsory for Swedish students, it is up to the teacher to decide how the assessment results should be used. There are also national tests arranged twice a year in courses B, C and D in mathematics. Schools must offer tests in course A and the highest mathematics course in a particular program, although it is strongly recommended that schools use the available tests as part of their assessment of the other courses as well.

Copyright © 2004 National Academy of Sciences. All rights reserved. No part of these pages, either text or image may be used for any purpose other than personal use. Reproduction, modification, storage in a retrieval system or retransmission, in any form or by any means, electronic, mechanical or otherwise, for reasons other than personal use, is strictly prohibited without prior written permission.

Opinions and statements included in the draft papers are solely those of the individual author(s), and are not necessarily adopted or endorsed or verified as accurate by the Committee on Test Design for K-12 Science Achievement or the National Academy of Sciences, including the National Academy of Engineering, Institute of Medicine, or National Research Council.

Only English, Mathematics and Swedish have regular national tests. However there are web-based test-banks in Physics, Biology, Chemistry and for Mathematics course E (the only mathematics course without a national test). Initially pre-tested assessment items were made available to teachers, classified according to subject content and expected cognitive processes needed to complete the item, accompanied by information about the expected performance on those items of students at each level (more precisely, the proportion of minimally competent students at a particular threshold that would be expected to answer the item correctly). The purpose was to support teachers so that they could assemble a test that would faithfully reflect their interpretation of the national objectives and choice of content in a particular course and also provide nationally comparable data on standards of performance. Currently, however, rather than just issue new items individually, new items are published in the form of an examination for Mathematics course E and one for each of the Physics courses A and B. These examination papers are made available through the internet to authorized teachers, each of whom decides whether to administer them to her or his students or not. In practice around 70% of the teachers teaching these courses make some use of the examinations, although teachers are allowed and encouraged to modify the examinations, for example by adding other items from the bank or deleting items that relate to topics that the teacher has not taught (approximately 20% of teachers choose to modify the examinations in some way). Teachers are provided with cut-scores on the examinations for pass, pass with distinction and pass with special distinction, but as with the national tests, the teacher does not have to award the student the grade gained on the examination. If the teacher feels that the performance on the examination is not representative of a student's achievement, the teacher can choose to award the student some other grade.

The teachers of Physics and Mathematics who choose not to use the examination, and the teachers of Chemistry and Biology (for which there are no examinations) are supported by the test bank in very much the way the bank was initially meant to work. In Biology, as well as individual items, there are also 'blocks' of items relating to particular topics (e.g. Ecology), which can be included in a test.

The final grades on the upper secondary assessment are used by universities in selecting students, especially for the programs that are in high demand, such as medicine. However there is also an aptitude test, modeled on the SAT developed by the College Board in the USA, called the SweSAT, which most students in upper secondary education take in any of their final years of school. Students do not have to choose whether to make their case on the basis of school grades—they are automatically considered on whichever basis gives them the best chance of success. Currently, around 60% of students gain admission to a university on the basis of their school grade and 40% on the basis of the SweSAT, although since the school grades of girls tend to be higher than those of boys, around 70% of girls and only 50% of boys are admitted on the basis of their school grades.

France

Structure of schooling

The school system has three levels: primary (ages 6 to 10/11), lower secondary *collèges* (ages 11 to 15/16) and upper secondary *lycées* (ages 15/16 to 17/18+). In addition, most children attend pre-primary schools from age 3, although schooling is only compulsory for ages 6 to 16. There is a national curriculum for all levels, although schools have some flexibility in their time allocation between different subjects. In the last two years of the *collège*, students study eleven subjects. About two hours per week are spent on the sciences. There are two types of *lycée*, general and vocational. All students follow a curriculum composed of a common core of about eight subjects, together with some electives, in the first year, and then choose between different courses for the subsequent years. There are three different general courses, focused respectively on literary, scientific, and economic and social studies, and there are also vocational courses. Students have a choice only between different courses as a whole for the whole year. This is a general feature of the French system; students have no possibility of choice inside a given structure, so that, for example, a flexible system of credit accumulation is not possible.

In addition to these courses, taking three years in all at the *lycée* and leading to the *baccalauréat* examination, there are two-year courses that lead to the Certificat d'Aptitude Professionnelle. In the general scientific course, students have to choose in the third year either mathematics and technics, mathematics and biology, or mathematics and 'physics'. Physics and chemistry are never separated in school teaching: the subject is called 'physics', although in most countries, this combination would be termed 'physical science'. About five hours per week are devoted to physical science in the last two years, with about 1.5 of these spent on practical work. The total teaching time for all subjects in the second and third years is between 24 and 30 hours per week each year, depending on extra options that may be taken.

Formal summative examinations

An examination at the end of primary school existed from 1945 until it was abolished in the early 1970s. An examination at the end of the *collège*, called the *brevet des collèges*, has also ceased to be compulsory. For some students, this could be the last examination they take in compulsory schooling, but it has little value in the job market: many take it for its value as practice for the later *baccalauréat* examination.

The *baccalauréat*, taken at the end of the third year of the *lycée*, is offered in three versions, the *général* academic version, the *technologique* (technical) and the *professionnel* (vocational). The proportion of the age group taking one or other of these versions has risen from 30% in the 1970s to over 90% by the year 2000: about 75% of candidates pass this examination.

The Baccalauréat consists essentially of written tests (except for foreign languages). The nature of the papers is approximately the same in the different versions. Each examination lasts for 3

to 3.5 hours and the duration of the set of written examinations is 4 days. There is only one paper for physics and chemistry, with separate exercises for the physics and the chemistry.

Each of the science examinations (i.e. biology and physical science) consists of between four and seven questions, all of which have to be answered. The general approach is to test students in depth on a limited number of topics. Multiple-choice items are not used—the questions are mainly structured open-response questions, with most requiring extended rather than short responses, or (in biology) an essay question. By comparison with tests in other countries, these papers give relatively high weight to investigating the natural world

The *baccalauréat* examinations are set separately, within an overall national framework, by each of a set of 28 regional academies, although some collaborate to set joint papers. The preparation is done by a team of between two and four high school teachers, each of whom teaches classes at the examination level, and their work is co-ordinated by a university professor (who must have taught for a minimum of six years). The entire process is supervised by the general and local education inspectors. Three papers have to be set each year: in March of the year before the examination, the team for preparing for each kind of paper is selected—one for the June paper, one for the September paper, and one for a spare (substitution) paper. Government inspectors will coordinate the work of the different teams. Each draft is submitted to one or two teachers who must solve the problems completely under the same conditions as a candidate; they then report on the likelihood of their students being able to solve this set of questions. At the end of the preparation procedure, three reports are made for each of the three papers: one from the teacher who tested the paper, one from the team, and one from the inspector. The director of the academy responsible has to make the final choice, choosing between those papers that have obtained three positive reports. If no paper fulfils these conditions, the director can call for a meeting of all those consulted and if no agreement is obtained director will use a paper that has been accepted in a previous year and not used.

Papers are marked by teams of local teachers, some of whom may have been involved in setting the paper. A teacher may mark the paper of his own students (because of the anonymity) but is not allowed to examine any of his or her students orally. Oral examinations are used to make decisions for candidates whose marks are very close to (but below) the minimum pass mark.

A pass in the *baccalauréat* is the main requirement for entry to higher education; for many university courses there is no other criterion for selection. The capacity of the university is the only factor that can limit the number of students and thus, implicitly, introduce the possibility of selection: this is the real situation in the academies of the Parisian region.

There are two highly selective routes, however. The first concerns the *institut universitaire de technologie* (I.U.T.) or polytechnics. These institutes belong to Universities but are autonomous particularly in their choice of their students. Selection involves submission of school reports as well as the *baccalauréat* results. Their courses last two years and are sanctioned by a degree called the *diplôme universitaire de technologie* which allows the

holder to get a job or possibly, with very high marks, to study at the University or Engineering schools.

The second selective route is one for preparation to enter the *grandes écoles*: These are prestigious engineering schools with competitive entrance examinations. Special preparatory courses are provided, in institutions which are administered within the secondary school system, to prepare students for these entrance examinations over two or three years (*classes préparatoires*). As no certificate is delivered if a student fails in these special entrance examinations these students register also at the university for a “free entrance” route. In this route, they don't follow teaching but can get through the examinations. Moreover in some cases, and in some universities, according to the results at the *grandes écoles* entrance examination they can get the “*équivalence*” of the diploma that students obtain after two years in a normal university course.

The formative use of system evaluation

The evaluation of the education system is the responsibility of a specific department in the Ministry of Education called the Office for Assessment and Forecasting. It monitors all possible aspects, including facilities and resources, classroom practices, students' achievements and school effectiveness.

Achievement is measured for students at the end of the *collège* stage, in all subjects, by use of sample surveys. Non-cognitive aspects, including attitudes and values, are also surveyed by student questionnaires. Results are analyzed in relation to various background variables for the samples. These data are supplemented by sample surveys of students at the beginning of the primary stage, and by cohort studies in which samples of students are followed through several years of their schooling. Some items are also used repeatedly over many years so that long-term trends can be monitored.

The underlying policy is to encourage educational change by promoting a culture of evaluation amongst all stakeholders. As a member of the Office has expressed it:

This is based on the belief that such stakeholders, among whom teachers are prominent, will improve their professional practices only if they are shown, as in a mirror, the consequences of their actions. (Bonnet, 1997, p.299)

For this purpose, blanket testing of all students is conducted, in alternate years, for all students at ages 8 and 11, and every year in all subjects for students at age 16. The tests use multiple choice or constructed-response items according to the topic involved. Whilst a sample of the results of these tests is analyzed centrally to supplement the data from sample surveys, the main purpose is to provide teachers with diagnostic information. This pedagogical and formative purpose is made evident by having the tests set at the *beginning* of the school year, so that they inform each teacher about their new class and are to aid teaching, not to judge teachers. The scoring is done, using standardized procedures, by the teachers themselves. The same approach is emphasized by involving teachers from a number of regions in the setting of the tests, and by providing national

norms so that teachers can calibrate the results of their own classes. The centrally analyzed sample data are used to inform the design of pre-service and in-service training programs.

One example of the feedback effects of the tests started with a finding in 1989 that primary school children produced results that were much weaker in geometry than in arithmetic even although both had equal weight in the national curriculum. Publicity for this result was enough to lead to improvement in the geometry scores. It has been shown that the majority of teachers use the test results as a guide for their teaching. The Ministry has gone further by providing banks of questions, in all subjects and for all levels, for teachers to use for their own diagnostic purposes. Initial reactions to this initiative were favorable, although change in classroom practices that can make effective use of diagnostic procedures have been slow to develop.

The system of school inspectors is mainly devoted to assessment of individual teachers rather than to schools as such. The Ministry has developed a set of indicators and of procedures to help schools to evaluate their own performance against regional and national norms; the procedures include attention to the 'value-added' by a school through analysis of changes in students' attainments as they proceed through the school. Whilst newspapers publish 'league tables' based on schools' results in the *baccalauréat* examinations, the Ministry policy is to help each school to analyze its own data. Thus, for example, results of the mass tests of its students are not used centrally to calculate a school's 'value added', for this would "most certainly antagonize the teaching profession – it would be seen as a way to evaluate the teachers themselves – and alter beyond recognition the formative nature of mass assessment" (Bonnet 1997, p.303).

References

Black, P. (ed.) (1993) *Physics Examinations for University Entrance*. Science and Technology Education Document Series No.45. Paris: UNESCO.

Bonnet, G. (1997) Country Profile from France. *Assessment in Education* 4(2), 295-306.

Britton, E.D.& Raizen, S.A. (eds.) (1996) *Examining the Examinations*. Norwell, MA: Kluwer.

Ministère de l'Éducation Nationale (2003) Les évaluations: les outils nationaux – année 2002-2003. Bulletin officiel de l'éducation nationale no.19.
<http://www.education.gouv.fr/botexte/bo20409/MENK0201038N.htm>

Germany

Structure of schooling

The main responsibility for education belongs to the sixteen regional states (the *Länder*) although the federal Ministry of Education has a limited role in guaranteeing certain freedoms and rights, a strong role in the administration of part-time vocational education, in pay and conditions for teachers, and in the running of the university system. Each *Land* has its own

Minister of Education. The sixteen ministers do collaborate with one another to align many aspects of their school systems, and with the federal ministry to agree the pay and conditions for teachers and the organization of the university system.

There are four main levels in the school system. For the pre-school, up to age 6, schooling is not compulsory although over 80% of children attend pre-school classes, mainly in private schools for which most parents have to pay fees (education being free from age 6 onwards). Children spend four years, from ages 6 to 9, in primary schools, and a further five or six years in the 'lower secondary' schools. After these nine or ten years in school, pupils have to complete compulsory schooling with a further three years either in full-time or in part-time school study.

Whilst all are taught together in primary schools, the secondary system is strongly differentiated between practical and liberal (i.e. professional and academic) curricula. At the lower-secondary level, about 29% of pupils enter schools which offer only practical, non-academic, courses (the *Hauptschule*); 28% attend schools which offer both practical and liberal courses (the *Realschule*) and 35% attend the liberal (*Gymnasium*) schools which offer academic courses with three options – classical, modern languages, and mathematics and sciences. The remaining 9% attend comprehensive schools (*Gesamtschule*), which offer either integrated practical/liberal courses, or separate liberal and practical tracks within the one school. The choice of school depends on school reports from the primary level; in some *Länder* choices can be reconsidered in the first year or two of lower secondary.

At upper secondary level, most of those attending a *Gymnasium* and some in a *Gesamtschule* (about 21% of the age group in all) carry on for a further three years of study for the *Abitur*—an academic examination which is required for university entrance. The rest move on either to full-time study in a variety of vocational schools (16% of the age group), or to three years of part-time study in programs arranged between private employers and vocational schools (62% of the age group).

Curricula and examinations

There is a variety of practice between the different *Länder*, both in curriculum and assessment. For example, the number of hours per week given to science in the *Gymnasium* can vary between 4 and 6 hours, and the assessment for the *Abitur* examination is by externally-set state tests in seven of the *Länder* but by tests set by each school and approved by the state ministry in the other nine. However, the Ministers of the *Länder* have established mutual agreements about national curriculum and assessment standards, so that qualifications from any one are recognized in the others (such mutual agreements have to be confirmed by the legislature of each region). Overall, there is a delicate balance between national uniformity and regional control. Much of the co-ordination effort is focused on mathematics and science subjects, and the National Council for Science is influential in such initiatives.

Each regional ministry prescribes the subjects to be taught, the textbooks to be used, and the total numbers of periods per week—yet this latter number can differ by as much as 20% between different *Länder*. In the primary *Grundschule* science is included along with the basics of German and mathematics, and social sciences and language arts. At the end of grade 4, teachers

produce an assessment of each student and recommend the appropriate lower secondary school; one state has a central examination for the end of grade 4. The decision about the lower secondary school has to be ratified by parents, who can appeal against the school recommendation.

In lower secondary schools, science is one of seven subject areas that have to be studied. At the end of lower secondary, i.e. the end of grade 10, there is summative assessment for a school-leaving certificate. In seven of the *Länder*, this certificate is partly determined by a central state examination, in the others it is assessed by the schools. Those not in a *Gymnasium* or a *Gesamtschule* may transfer to one if they do well in this certificate. For grades 12 and 13 in courses oriented towards the *Abitur* assessment, there are rules governing a balance between basic and advanced courses. About one-third of the time has to be spent on six basic courses and two thirds on twenty-two basic courses. The advanced courses must be taken in at least two different subject fields. A basic course will require about three hours study per week for a semester, an advanced course five to six hours per week. The courses must usually include study of language and literature, modern foreign languages, social sciences, mathematics science and technology, religion, and civics.

For the *Abitur* examination, students must offer four subjects, of which three (two at advanced level and one basic) are tested by written examinations, and the fourth by an oral examination. The four must include at least one from each of the following areas: language literature and the arts, social sciences, and mathematics, science and technology. The assessments, whether school based or regional, must conform to national standards that are prescribed for all subjects. For the overall *Abitur* assessment, 36% of the marks come from these four assessments, 39% from coursework assessments in the basic subjects and 25% from coursework assessments in the advanced subjects. Discrepancies between coursework and test scores may be resolved by further oral examinations. A pass requires a total score of at least 33% and 95% of candidates achieve a pass.

Samples of the tests set in the sciences by two of the states (see Britton and Raizen 1996, AFT 1994, 1995) show that in each of biology, chemistry and physics, there is a 4 to 4.5 hour test usually composed of between four and six questions, each sub-divided into several subsections. There is no use of multiple-choice items, and the ratio of long to short free response items (2 to 1) is one of the highest amongst the country tests reviewed. Some of the questions in biology were on unusually advanced topics, but calling mainly for recall. In chemistry and physics, the percentage of marks assigned to theorizing, analyzing and solving problems was higher than for almost all of the tests of other countries.

A pass in the *Abitur* used to guarantee university entrance, but as the numbers taking it, and the general demand for higher education, have grown, universities have had to find ways to restrict admissions, particularly in the most popular fields (notably medicine, dentistry and veterinary science). Many now set their own entrance tests and combine results of these with the *Abitur* results in making admission decisions.

Part of teachers' initial training is concerned with creating and scoring *Abitur* examinations, and there are procedures within schools for cross-checking grading standards between teachers.

However, whilst there is strong reliance on the assessments made by teachers and schools, it must be recognized that these are high-stakes assessments in a system with a strong orientation towards summative assessment. This is primarily because the German system has been based historically on a belief in the pre-determined ability of its pupils. This justifies making decisions about choice of secondary school, which strongly affect future career prospects, at age 10. However, this also puts great pressure on primary schools to get it right. A great deal of attention is given to providing personally differentiated work within mixed ability primary classes. Primary schools start to give grades in grade three, and grades are then assigned, by formally established procedures, twice each year. Results are based on written and oral tests and on written exercises, balanced as appropriate for the different subjects. For each test which counts towards a grade, pupils must be given a few days prior notice, and parents have to sign a report informing them of each test result. Tracking (grouping by ability) is common in secondary schools, although in some *Länder*, the first two years in lower secondary are organized as orientation years tracking being postponed until grade 7.

Across all levels, pupils are not automatically promoted to the next year of study. They must achieve a mean grade of at least 4 on a scale from 1 (high) to 6 (low). Those achieving only grades 5 and 6 will have to repeat the whole year of study; however, they can attempt a written examination at the start of the next school year to try to redeem their failure. Estimates of the proportion of pupils who repeat a year at some stage in their school career vary, from an official figure of 4% to unofficial estimates of 10 to 20%. The continual pressure on schools' assessments may help explain why much of the teaching, particularly in the *Gymnasium* classes, is formal, teacher-directed and based mainly on lecturing.

As might be expected, the early selection system leads to strong social bias across the different types of secondary school, so that social inequity is enhanced rather than reduced by the school system. This is matter of political controversy between right-wing parties, who look to maintain the system, and left-wingers who seek for ways to offset the inequities. Such disagreement partly explains the differences between the different *Länder*—indeed one main point of tension between them is that some worry that the standards for the *Abitur* are too low in some *Länder* compared with their own.

References

AFT (1994) *What college bound students abroad are expected to know about biology: Exams from England and Wales, France, Germany and Japan*. Washington DC: American Federation of Teachers.

AFT (1996) *What college bound students abroad are expected to know about chemistry and physics: Exams from England and Wales, France, Germany and Japan*. Washington DC: American Federation of Teachers.

Britton, E.D.& Raizen, S.A. (eds.) (1996) *Examining the Examinations*. Nowell MA: Kluwer.

Duit, R. (2004). Private communication.

Eckstein, M.A. and Noah, H.J. (1993) *Secondary School Examinations : International Perspectives on Policies and Practice*, New Haven: Yale University Press.

Stevenson, H. W.; Lee, S.-Y. & Nerison-Low, R. (Eds.). (1999). *Contemporary research in the United States, Germany, and Japan on five education issues: structure of the education system, standards in education, the role of school in adolescents' lives, individual differences among students, and teachers' lives*. Washington, DC: ED Online, U.S. Department of Education.

Japan

Structure of schooling

Japanese schools follow the old US model, with elementary schools (*shogakko*) covering grades 1 through 6, junior high schools (*chugakko*) covering grades 7, 8 and 9, and high schools (*kotogakko*) covering grades 10, 11 and 12. Education in Japan is compulsory for children between the ages of 6 and 15 (i.e. for the period covered by elementary and middle school). Almost all Japanese children attend either kindergarten or day-care center before beginning their formal schooling. Promotion from one grade to the next is based primarily on attendance, with almost no use of grade retention.

The Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) specifies 'National Courses of Study' which teachers must take into account in planning course content, and in deciding how much time is allotted to each topic. The regulations for some subjects (e.g. learning to write the Japanese scripts) are very specific, but in others, such as science, only broad outcomes are specified.

Curriculum and assessment

In elementary and junior high schools, the curriculum is divided into three major categories. In elementary schools, the *standard* curriculum includes Japanese language, social studies, mathematics (mainly arithmetic), science, life environment studies, music, arts and handicrafts, homemaking, and physical education. In addition there is a program of *moral education*, including attention to health and safety, self-discipline, public manners and courtesy, and environmental awareness. There are also programs of *special activities*, which include ceremonies, cultural performances and sports.

In elementary schools, science is introduced in grade 3, and teaching is organized into a number of topics that typically cover 10 to 14 lessons of 45 to 50 minutes each. For example, a unit on 'Levers' might have the objectives that students learn:

- the names of the parts of a lever
- what happens when the load, the fulcrum or the effort is moved
- what needs to happen for the lever to balance

The content of a topic is typically designed to last less than the time allocated, so that the timetable allows time for students to take a short test, and for teachers to take any necessary

remedial action with the whole class before moving on. These tests typically combine open and closed items and focus on the 'big ideas' in the topic. For the 'Levers' unit above, for example, students might be shown four pictorial descriptions and asked which would be the easiest to lift. The marks for these tests, and for home work assignments, may feed into the term grade, which is generally reported on a three-point scale, but this is a decision that is made by the school, rather than the prefecture or the ministry of education.

Since most students progress automatically to the nearest junior high school, there is relatively little pressure on students in elementary schools, except for the small number of students who apply for places at one of the national or elite junior high schools.

However, in junior high schools, there is a marked change in the demands made of students, since one of the primary objectives is to prepare students for the high-school entrance examinations (*koko nyugaku shiken*), which will determine which high-school a student will attend. In grades 7 and 8, students are given short tests on each topic which serve a formative role, as in elementary school, and longer, more synoptic, tests in the middle and at the end of each term, which usually include multiple-choice, short-answer and extended-answer items. These assessments are used principally to rank the students in each class. At the end of grade 9, students take a common high-school entrance examination. All students in the same prefecture will take the same examination, which is constructed by a committee that is dominated by high-school teachers. Individual items from these tests are released periodically, although the tests are re-used from year to year. However, it is also possible to buy booklets of 'unofficial' practice questions, which are widely available.

In addition to the high-school entrance examination set by the prefecture, individual elite high schools set their own entrance examinations (typically written by teachers at the high school). Since these examinations are generally scheduled for the same two or three days in February, it is possible for a student to apply only to one or two elite high schools, and so one of the key functions that the results of mid-term and end of term tests taken in the ninth grade serve is to help the student, teachers and parents to determine the high schools to which the student should apply.

Each set of high-school entrance examinations is held over a single day in February or March, and consists of four 50-minute examination papers, in Japanese language (*kokugo*), mathematics (*sugaku*), social studies (*shakai*), science (*rika*), and a foreign language (typically English). The results are generally announced a week later. On the basis of these results, students are offered places at elite academic high schools, ordinary academic high schools, or vocational high schools.

All high-school students applying for entrance to university must take the University Entrance Qualification Test (*Daiken*, or *Daigaku Nyugaku Shikaku Kentei*) which assesses the high school portions of the national course of study, and those applying to national or public universities are also required to take the Center Examination (*Senta Shiken*), administered by the National Center for University Entrance Examination (*Daigaku Nyushi Senta*) in January. Applicants may also have to take additional examinations set by the university they wish to attend, which are set by members of the university faculty.

The prestige of elite academic high schools is primarily determined by the proportion of their students gaining places at the most prestigious universities and so a considerable proportion of curriculum time in these schools is devoted to preparing students for the examinations set by the most prestigious universities. While these examinations are nominally based on the National Course of Study, they generally require a synthesis of ideas across different topics that is well beyond most students, and so, in a very real sense, the curriculum of the elite academic high schools is determined by faculty at the prestigious universities, rather than by the Ministry of Education. As well as focused coaching on exam technique, students also take a large number of 'mock' university entrance examinations (*mogi shiken*) to practice their examination technique, and to inform the choice of university to which to apply.

The ordinary academic high schools send few students to the prestigious universities, and so, while many of the students will go on to university, there is much less pressure to prepare students for university entrance examinations, and the curriculum therefore tends to follow the National Course of Study closely. The curriculum at vocational high schools is based on the minimum vocational training requirements laid down by the Ministry of Education.

References

Britton, E.D.& Raizen, S.A. (eds.) (1996) *Examining the Examinations*. Nowell MA : Kluwer.

Eckstein, M.A. and Noah, H.J. (1993) *Secondary School Examinations : International Perspectives on Policies and Practice*, New Haven: Yale University Press.

Stevenson, H. W.; Lee, S.-Y. & Nerison-Low, R. (Eds.). (1999). *Contemporary research in the United States, Germany, and Japan on five education issues: structure of the education system, standards in education, the role of school in adolescents' lives, individual differences among students, and teachers' lives*. Washington, DC: ED Online, U.S. Department of Education.

New Zealand

Structure of schooling

Schooling is compulsory from ages 6 to 16. Early childhood education is provided privately, but over 90% of 3-year olds are catered for by such provision. Over 98% of children attend state-funded schools. Primary schooling extends for about 8 years (years 1 to 8) and at ages 12 or 13 students transfer to secondary schools where they attend for at least 3 and up to 5 years (years 9 to 13). There are a few intermediate schools that cater for years 7 and 8. After the last year of compulsory schooling about 80% stay on for at least one more year (year 12), about 55% stay for a further year, and a small number stay on for year 14 to upgrade their qualifications. Many schools are small—for example, there are less than twelve year 4 students in over a third of the primary schools.

Curriculum frameworks

In 1991, following the example of the UK, national curriculum and assessment frameworks were introduced for all years of compulsory schooling. Curriculum plans were introduced quickly for English, mathematics, science and technology, then more slowly for languages, social sciences and health. To take the example of the curriculum for science, this is specified in terms of six strands; four in the main content areas—biology, chemistry, physics, and earth and space science—and two entitled respectively “Developing Scientific Skills and Attitudes” and “Making Sense of the Nature of Science and its Relationship to Technology”. These last two are to be addressed within the learning of the four content strands rather than by separate teaching. For each strand there is an overall set of achievement aims, and then a set of objectives set out in eight levels to specify the learning progression: each of the first five levels is planned to cover two years of normal progress, whilst the next three are each to cover one year. The objectives for each level are supplemented by suggestions for possible learning experiences, by samples of suitable learning contexts, and by examples of suitable assessment exercises: all of these are intended only as general and helpful guidance. It is recognized that different students will move from one level to the next at different ages: thus, for example, most students at level 2 will be distributed across years 3 to 5, whilst those at level 6 may be found across years 10 to 12.

School based assessments

A system for nationally monitoring achievement in *primary and intermediate schools* was first instituted in 1995. The aim is to establish criterion-referenced assessment based on the eight levels and the assessment is to be conducted by each student’s own teachers. Teachers have been struggling with this responsibility, for it has presented them with many new problems. There are issues of reliability, in that students’ performance on tasks can fluctuate from day to day, and between tasks which are intended to exemplify the same level. It has also become clear that different teachers have different interpretations of the standards. The work involved in regular and individual monitoring of a whole class is a burdensome for many, yet the gap of two years between levels is too large to give a satisfying sense of progress. Finally, students’ performances can differ widely between the different strands in a subject, so that it is hard to see how to aggregate performance to report a single subject level. Whilst some ways to help with these have been developed (e.g. sub-dividing levels to give a finer-grained scale of progress) the difficulties have not been fully resolved. Several national tools and exemplars of assessment have been provided as guides for teachers to use as they wish. There are tools for entry-level assessment and a further set of Assessment Resource Banks for levels 2 to 5, in mathematics and in the sciences (over 1200 items in each) and in English (over 600 items), freely available on the internet, and each with a marking guide. A further set of Assessment Tools for Teaching and Learning is being developed for the Ministry by academics in universities in Auckland and in New South Wales. The intention is that these assessment tools will allow teachers quickly to assemble summative tests of numeracy and literacy, tailored to the preferences of the teacher and the school, in both English and Maori. The resources are provided free on a CD-ROM, and include software, both to help assemble a suitable test from items that have been nationally normed and calibrated (by one-dimensional IRT modeling), and for printing reports for individual students and for groups.

Overall however, this is low-stakes testing and is the current compromise emerging from political disagreements on whether or not to introduce national blanket testing. The focus is on

encouraging and helping teachers to improve learning through feedback arising from good quality assessments. Written or oral reports of results are made to parents to complement other aspects of school reports. The assessment can be used for monitoring within the school and examined for the school monitoring by the national Education Review Office which oversees the implementation of the national curriculum as a part of its remit to review the work of schools. For *secondary schools*, assessment of years 9 and 10 is internal to the schools, whilst for years 11 to 13 it is dominated by the pressures of national testing. For the first two years, assessment is more summative in character than in the earlier years, and is based on scores and grades given to written work and to test papers. Attempts have been made to replace the old norm-based practices with a standards-based approach, and have led to use, for each school year, of a five-level scale for assessing each of a set of about five learning goals for each subject over the year. This approach has been found helpful by teachers; strangely, however, few secondary schools use the strands and levels scheme of the national curricula.

National assessments

The country is still in the midst of a process of implementing a completely new system that will provide certification of achievements, using a combination of externally- and internally-based assessments for each student. At the same time the new system will have a profound effect on the structure of the curriculum for students in years 11 to 13 since it is replacing two examinations. One was for the School Certificate taken at the end of year 11 and the other was the University Entrance Examination, usually taken at the end of year 12. In these older systems, all included a proportion of the total score—typically between 20 and 40%—assigned by teachers, on the grounds that some of the learning aims could not be assessed by written tests.

Work on the new “Qualifications Framework” started in 1990. The framework was to serve two purposes. First, it was to allow for greater flexibility in students’ choice, and secondly it was to close the status gap between academic and vocational qualifications by including all within a single system. The framework was originally designed by dividing the learning outcomes into quite small packages, called *unit standards*, which could be attempted at any time, and on multiple occasions. The initial intent was to adopt a strict mastery learning model, so that the only outcomes for a unit standard assessment would be pass or fail. Whilst this initial plan was acceptable in the vocational areas, those teaching in the academic subjects complained of the fragmentation of the curriculum and about the availability of only one level of pass.

What is now being implemented is a revised scheme retaining *unit standards* in some subject areas, but supplementing them, for most secondary school subjects, with rather larger packages called *achievement standards*. Each such standard will have four grades of result. Separate achievement standards will be provided at three levels. A collection of credits in a set of standards will qualify a student for a National Certificate of Educational Achievement (NCEA). A Level 1 NCEA will require 80 credits at level 1, a Level 2 will require 80 with at least 60 of these above level 1, and a Level 3 will require 80 with at least 60 above level 2. These will generally be aligned to work in years 11, 12, and 13 respectively. Credits gained for a certificate at any one level may be counted towards a certificate at a higher level. In any one subject, between 5 and 8 standards at the appropriate level will be available in each of the three years, so that up to 24 credits may be available altogether in one subject across the three years. Examples

of achievement standards in Biology are “Carry out a biological investigation with supervision” (level 2), “Describe patterns of evolution” (level 3) and “Describe gene expression” (level 3). Any one achievement standard may be worth several credits (typically 2, 3 or 4). In most subjects, half or more of the achievement standards will be assessed through national examinations. A typical test for two credits in a physics achievement standard will contain two structured questions, each with about seven sub-sections calling for open response in a few lines of writing or with a calculation. Other achievement standards are to be assessed by schools but with checks via inter-school moderation (i.e. calibration) procedures. In Biology, school-based assessment can account for 40% of an NCEA, while in Physics the proportion is 20%. The use of school assessments will enhance opportunities to assess learning aims for which external formal tests cannot be valid. The highest level of NCEA will be important for gaining university scholarships: the government intends to introduce a further level (level 4) on which to grant such scholarships. The new NCEA scheme is being introduced, level by level, over the years 2002 to 2004, so it is too soon to evaluate its success.

National monitoring

This is an additional and separate feature of assessment in New Zealand. Surveys of samples of students at ages 8-9 and 12-13 are conducted each year, covering thirteen areas of school achievement on a four-year cycle so that three or four are covered each year. Matrix sampling is used so that each of three sets of tasks is attempted by 500 students, giving a sample of 1500 in all. Tasks are presented to students using video and/or audio on laptop computers, and the computers may also be used to record student responses. Tasks are presented in one of four ways; first, one-to-one interview on a set of 15-20 tasks; secondly, video records of work by groups of four students on several longer tasks; third, independent work of students as they rotate around four task stations; and fourth, independent student work on paper and pencil tasks, or on creating works of art, or on demonstrating physical skills. For younger students, many tasks are administered orally. The tasks are administered by teams of approximately 100 experienced teachers who are released from schools over a six-week period, with a further group of about 200 teachers and students being employed to score the students’ responses.

These exercises produce very rich and vivid accounts of students’ attainments. Full reports are published and sent to all schools. About one-third of the tasks are kept constant from one cycle to the next so that trends over time can be explored. However, the monitoring is designed to give detailed and valid reports at the level of single tasks, and not to provide for generalization over broad domains of learning.

References

Crooks, T.J. (2002) Educational Assessment in New Zealand Schools. *Assessment in Education*, 9(2) 237-253.

Ministry of Education (2004) *What is the new Zealand curriculum framework?* See <http://www.minedu.govt.nz> for this and other curriculum information.

New Zealand Qualifications Authority (2004). See <http://www.tki.org.nz/r/assessment> for a range of information and of assessment materials for schools.

Australia (Queensland)

Structure of schooling

Schooling in Australia is compulsory from age 6 to 15 (16 in Tasmania), although many students attend state-funded pre-school beforehand, which is compulsory in some states. Primary schools cover the first seven years of compulsory schooling, after which students transfer to secondary or high schools. Students can leave school when they are 15, although almost all stay on until the end of year 10 to gain their school certificate. In most states, the school certificate grades are entirely school-based, but in New South Wales, there are externally-set examinations in a number of subjects. Students move from year to year with their peers—there is no ‘grade retention’ in Australia. At the end of year 10, students can leave school or continue to upper secondary school (years 11 and 12).

In most states, assessment at the end of upper secondary schooling combines both external examinations and school-based assessment, but in two (the Australian Capital Territory and Queensland) the final assessment is based entirely on peer-moderated school-based assessments.

The Queensland Senior Certificate Examination System

The Queensland Senior Certificate Examination System was designed specifically to provide formative feedback to teachers and 16-17 year-old college-bound students with the goal of improving teaching and learning. The examination system was conceived from the very start to focus on the formative function of assessment and build a summative examination from the elements of a formative system. The system was developed in response to an externally mandated summative system similar to the British A-Level examinations set by the University of Queensland and later, in transition, by the Board of Senior Secondary Studies.

In 1971, responding to the Radford Committee report (see below), the State of Queensland, Australia, abolished external examinations and replaced them with a teacher-judgment, criteria-based assessment that students had to pass to earn a Senior Certificate at the end of upper secondary schooling. The abolition of the A-level-like external examination was in response to: (a) “... the recurring tendency for the examinations to be beyond the expectations of teachers and the capability of students” (Butler, 1995, p. 137), (b) teachers learning all they could about the senior examiner to anticipate test questions, (c) the absence of formative feedback to teachers and students, and (d) public dissatisfaction with the process and narrowness of the curriculum syllabi (Butler, 1995).

In 1970 the Radford Committee recommended abolishing the external examination and replacing it with a school-based assessment. The Board of Senior Secondary Studies (BSSS, hereafter “Board”) was to have responsibility for setting the content of the 2-year syllabus in each content area and the methods of assessing accomplishment. The system included moderation to establish comparability of achievement ratings through a Moderation Committee and system of Chief Moderators. The “Queensland Experiment” has evolved over the past 30 years and will no doubt continue to do so.

At present it contains two major components. The first component is a school-based assessment system in each subject area (e.g., science). The school-based assessment system includes, in each subject area, a locally adapted “course of work” that specifies content and teaching consistent with the subject syllabus set by the Board and accredited by an external, government sponsored subject review panel. Teachers evaluate their students’ performance on a 5-point standards-referenced scale: very high achievement, high achievement, sound achievement, limited achievement, very limited achievement and teachers’ ratings are reviewed through a process of moderation (see below).

The second component of the system is a general test called the Queensland Core Skills (QCS) test. The test is used as a statistical bridge for equating ratings *across subjects* for the purposes of comparison. If there is a large discrepancy between teacher ratings based on two-years of work in a subject and the QCS, the former carries the weight, not the latter.

The assessment system works as follows. A subject syllabus (e.g., in physics) is set by the Board and covers objectives in four areas: (1) affective (attitudes and values), (2) content (factual knowledge), (3) process (cognitive abilities), and (4) skills (practical skills). Students are assessed in the last 3 areas only. The syllabus and assessments are instantiated locally, taking account of local contexts to link learning to everyday experience, in the form of “school work programs.”¹

Comparability of curriculum and standards are assured through a program accreditation process overseen by State and District Review panels comprised largely of teachers (about 20% of teachers are engaged in this manner in the State in each subject). These same panels assure comparability of ratings across the state by reviewing and certifying each school’s results in each subject area.

The assessment of achievement is formative as well as summative and continuous throughout the two years of study; students are told which assessments are formative and which summative (for the purpose of rating their performance on the Levels of Achievement). Each assessment must clearly specify which criteria and standard it measures and against which a student’s performance will be judged; students are to receive explicit feedback on each and every assessment exercise, formative or summative, along with recommended steps for improving their performance.

The Queensland Studies Authority (QSA) is responsible for oversight of the assessment system, devolving specific authority to the State and District (subject-matter) Review Panels. The Review Panels have responsibility for: (a) accreditation—examining a school work program and verifying that it corresponds to the Board’s syllabus guidelines; (b) monitoring—examining samples of Year 11 student work with tentative Levels of Achievement assigned to check

¹ “The science teachers in each school are encouraged to write unique work programs guided by the broad framework in the syllabus document but using all the resources available within the community and environment surrounding the individual school and taking account of the unique characteristics of students in the school’ (Butler, 1995, p. 141).

compatibility with criteria and standards; and (c) certification—of Year 12 Exit Levels of Achievement by insuring compatibility.

The various phases of the moderation process are (Cumming & Maxwell 2004):

1. approval of work programs
2. monitoring of year 11 standards
3. verification of year 12 standards
4. confirmation of year 12 results
5. random sampling (post hoc by the QSA for quality assurance).

Important features of the system are (Maxwell, 2004):

- The contents of student portfolios can differ substantially but can be judged against common standards – what matters is not whether two portfolios are identical but whether they are equivalent in terms of the specified standards: this involves judgment – and interpretation of quality.
- Teachers can provide assistance to students as they work through the assessments: the extent to which support has been provided changes the nature of the inference concerning what the student actually knows and can do.
- The validity of the assessment is improved by assembling the portfolio from a variety of assessment types and contexts, giving students many opportunities: formal tests can be part of this.
- Assessment is progressive, with assessment distributed over time; later assessments allow for improvement on knowledge and skills assessed in earlier assessments.
- Thus evidence is being collected throughout the course and so is diversified over time.
- Weighting of assessments collected at various times is for judgment of the final (exit) portfolio against the exit standards; the two principles are *fullest* (information must be there on all mandatory aspects and on all criteria) and *latest* (later work normally supersedes earlier work covering the same criteria). So poor starts and atypical performances can be ignored.
- The allowance for advice and for selective up-dating, and the need to cover all criteria, call for feedback advice so that assessment for learning matches naturally with assessment of learning.
- Where teacher feedback has been so close that it is impossible to disentangle the student's input, it may be necessary to ask the student to undertake another assessment of a similar nature.

Thus the system depends heavily on the expertise of teachers, and was built on the belief that teachers can acquire the necessary expertise and that they will act both professionally and ethically.

The experience has been that the need for teachers to acquire skills in assessment and in judging quality against standards creates its own impetus. Two features of this issue are highlighted by Graham Maxwell (2004), the deputy director of student performance for the QSA:

“Teachers take up the challenge when they are given the responsibility”, and

“The most powerful means of developing professional competence in assessment is the establishment of regular professional conversations among teachers about student performance (moderation conversations). This is best focused on examples of student portfolios. Looking at actual examples and discussing the conclusions that can be drawn about the student’s performance when compared to explicit standards sharpens teacher judgment and builds knowledge and expertise about assessment more successfully than any other process.”

The Queensland system provides a number of lessons in assessment-system design. Built in response to an externally mandated system that was widely recognized as too difficult for teachers and students, too narrow, and as having undesirable consequences, the present system focuses on formative assessment with immediate feedback to teachers and students that is linked to a summative function—the senior certificate. The assessment is continuous over two years providing feedback to students on how to improve their performance from both the formative and summative examinations. The curriculum (“course of work”) and assessments are developed locally to external specifications for the content domain (a syllabus) and assessment techniques. An external, governmental governance structure is in place to assure the public of the credibility of the course of study, the assessment, and the scores derived from the assessment. And the assessment focuses on both knowledge and cognitive abilities acquired during the course of study.

References

Cumming, J.J. & Maxwell, G.S. (2004) Assessment in Australian schools: current practice and trends. *Assessment in Education* 11(1) 94-108.

Maxwell, G.S. (2004) *Progressive assessment for learning and certification: Some lessons from school-based assessment in Queensland*. Paper presented at the 3rd. conference of the Association of Commonwealth Examination and Assessment Boards, March 2004, Nadi, Fiji.

England

Structure of schooling

In England, school attendance is compulsory from the start of the term after a child’s fifth birthday until the end of the school year of her or his sixteenth birthday, although in practice most children are in schools or play-groups from the age of 3. The first full year of compulsory schooling is called Year 1, although students born between September 1st and May 1st will have been in school the previous year, called ‘Reception’. In most local education authorities (LEAs), primary schools provide education from Reception to Year 6 and secondary schools provide education from Year 7 to Year 13, although primary schooling is occasionally provided in two schools: infant schools, covering Reception to Year 2 and junior schools, covering Years 3 to 6. In a minority of LEAs, primary schooling ends with Year 4 or 5, after which students move to middle schools for three or four years, before moving on to high schools for the remainder of their compulsory schooling. At the end of Year 11 (i.e. at age 16), students may leave school, although most choose to continue their education, either by staying on at school, or transferring to a tertiary institution.

Curriculum and assessment

In England, there is a national curriculum for all students of compulsory school age (ages 5 to 16) comprising ten subjects: English, Mathematics, Science (these three form the ‘core’ subjects), Technology, Modern Foreign Languages, Geography, History, Art, Music, and Physical Education. All subjects must be taught up to the age of 14, but there is limited scope for students to ‘drop’ non-core subjects for the last two years of compulsory schooling, or to embark on a range of vocationally-oriented courses (see below). The curriculum is defined in terms of Programmes of Study, Attainment Targets, and Assessment Arrangements for each subject. The Programmes of Study lay out what students must be taught during each of the four ‘key stages’ of compulsory schooling (years 1-2, years 3-6, years 7-9 and years 10-11). The Attainment Targets specify what is to be assessed at the end of each of the key stages and exemplify the standards of achievement expected. The Assessment Arrangements detail how the assessments are to be carried out.

Although students may leave full-time education at 16, fewer and fewer students do so, and there has been increasing emphasis in recent years on planning for a coherent 14-19 curriculum. There are two main qualifications systems for students in this age-group. The traditional route, and the one taken by most students aspiring to go on to higher education, has been to take the General Certificate of Secondary Education or GCSE at age 16 and the General Certificate of Education Advanced Level (usually abbreviated to ‘A-level’) at age 18.

It is up to each school to decide the subjects for which a student will be entered, but most students are entered for 8 to 10 subjects. The assessment typically takes the form of a combination of timed written examination papers (usually two papers for each subject, lasting up to two-and-a-half hours each) and assessment by teachers of work completed by the student in school and at home. In GCSE, the weight of the school-based component ranges from 20% (in Mathematics) up to 50% (in Technology), and the results are graded on a nine-point scale: U, G, F, E, D, C, B, A, A* (highest).

For the final two years of secondary education, students following the academic route specialize in a small number of subjects (typically four or five) and after a year’s study, most specialize further, by concentrating on just three subjects for the final year of schooling, at the end of which they take the A-levels examinations. The grades on these examinations are the main source of information used by universities in selecting students. There are five passing grades for A-levels: E, D, C, B, A (highest), although there is pressure to discriminate more finely within the higher grades, since approximately 20% of all students taking A-levels gain the highest grades.

Alongside these academic qualifications, and within the same integrated qualifications framework, is a range of vocational qualifications. Level 1 vocational qualifications are deemed equivalent to grades G to D at GCSE, while those at level 2 are deemed equivalent to grades C to A*. Level 3 vocational qualifications are deemed equivalent to A-level. Some of these vocational qualifications are closely related to a specific vocation (e.g. the level 2 Certificate in Air Cabin Crew Skills) while others are not. In particular the General National Vocational Qualifications (GNVQs) are available in a range of subjects, including Science, but tend to have a more

vocational orientation than the GCSE. Many of the GNVQ courses count as the equivalent of four subjects at GCSE, and so, driven by the need to improve scores in order to improve the school's position in the 'league tables' of results published in local and national newspapers, many schools have adopted GNVQs, especially in subjects like Information and Communications Technology, where standards are perceived as being easier to attain than in GCSE subjects.

In key stages 1, 2, and 3 assessment in the core subjects (English, Mathematics and Science) involves both teacher judgment and external tests and tasks. Other subjects are not assessed in key stages 1 and 2, and at key stage 3, assessment is entirely via teacher judgment. In addition, there are 'optional tests' that can be used during a key stage to evaluate a student's progress in English and Mathematics. There is little or no use of 'grade retention' in English schools, but students who have not achieved the expected level (level 4) at the end of year 6 (end of key stage 2) are given additional classes during year 7 (and if necessary, year 8) in secondary school and are given annual 'progress tests' until they do so.

Perhaps the most radical feature of the assessment system in use in England is that before the curricula for the national curriculum subjects were determined, it was decided to define the levels of achievement for each subject independently of age. A total of ten levels of achievement are currently used for reporting the achievement of students in the first three key stages. Students assessed at the end of the first key stage (i.e. at the age of 7) are generally assessed at level 1, 2 or 3 (with the median student awarded level 2), although some students achieve level 4 and students who fail to reach the standard associated with level 1 are reported as 'working towards level 1'. At age 11 most students are awarded levels in the range from 2 to 6, with level 4 as the 'expected level' (currently achieved by approximately 75% of the cohort). At age 14, most students are awarded levels 3 to 8, though again, the weakest students will be reported at lower levels than this, and in some subjects there is a higher level awarded to students whose performance is well beyond that associated with level 8 which is called 'exceptional performance'. Although the levels awarded at different key stages are based on different Programmes of Study, the intention is that the levels are comparable, so that level 3 might describe a high-achieving 7 year old or a low-achieving 14-year old.

Of course such a model may fit some subjects better than others. In Mathematics, for example, it is plausible to think that a high-achieving 7-year-old might give the same sorts of responses to test items as a low-achieving 14-year-old. In other subjects, however, depending on the sort of test, this may be less likely, and this is manifested in lower dispersion of levels for such subjects at each age. However, even where such a model does not fit particularly well, the effect of requiring that achievement be defined in a way that is, as far as possible, independent of age, had substantial effects on the subject definition. In particular, since it was intended that each level subsumed the previous levels, the subject working groups were forced to articulate a notion of progression in their subject; in other words, they were forced to answer the question, "When someone gets better at this subject, what is it that gets better?"