

## PART 2

### BACKGROUND FACTORS

#### OPERATIONAL REQUIREMENTS FOR VOCATIONAL QUALIFICATIONS AND EXAMINATIONS

##### An operational model

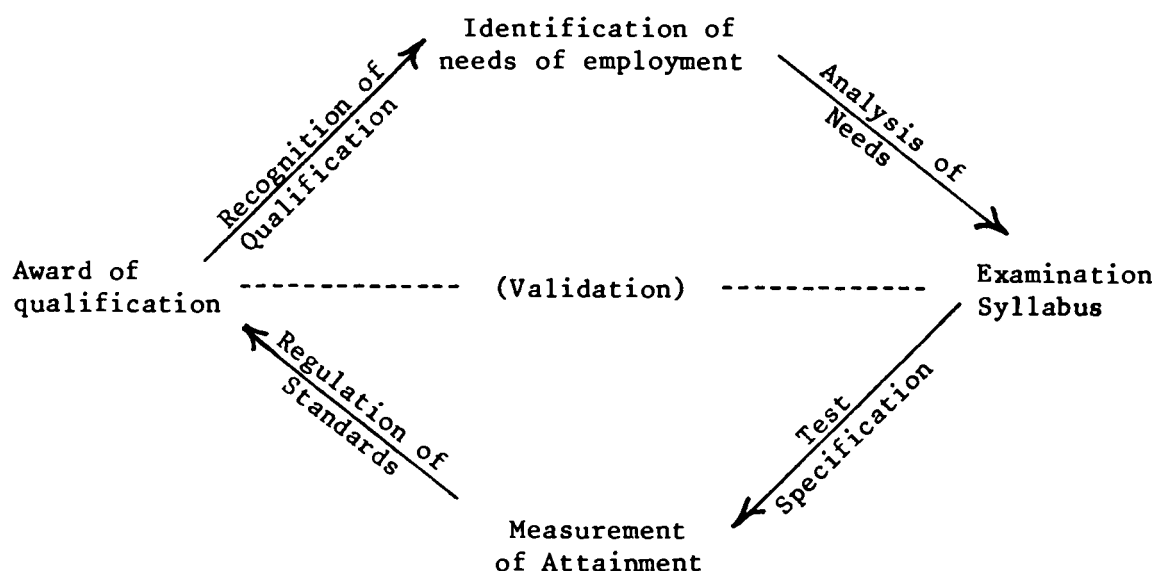
28. In this report specialist terminology (e.g. as used by the Technician Education Council and other qualifying bodies) has been avoided as far as possible. However, in some contexts, and in particular in this section, it is necessary to use some technical terms and so a short glossary is provided in Appendix III.

29. Any system of vocational qualifications and examinations has four essential elements

Identification of the needs of employment  
An examination syllabus  
Measurement of attainment  
Award of a qualification

Some systems have an additional element: validation.

30. These elements can be shown diagrammatically as follows



31. The essential difference between vocational qualifications and examinations and academic ones is that the former have to meet the needs of employment, in terms of both the content of the examination syllabus and the recognition of the qualification as identifying a standard which is acceptable in terms of a job to be done. Vocational qualifications and examinations always have to be related to this reference point which is external to any educational one.

32. Historically in the UK, and hence in the developing Commonwealth, the other three essential elements have not been separated: a qualifying body sets an examination to determine who is worthy of receiving a qualification,

and establishes the syllabus on which the examination is based. Although it can be convenient in practice to combine these three elements, there is no absolute logic in their combination. In this report they are treated as separate activities requiring different resources and techniques, and they could in fact be undertaken by two or even three separate bodies. To some extent this already happens in validation systems.

33. There are often differences in emphasis over the relationship of a college course to examinations and qualifications. Sometimes a view is heard that the course is all important, and at best a qualification is an additional incentive for students to study hard on the course. What is learned is more important than the means of recognition. A logical development of this view is that a qualification is only awarded if there has been attendance at a course and this in turn can lead either to the award of the qualification by the teaching institution (following the university tradition) or by an external qualifying body adopting a validating role and approving the course (content and often resources and staffing) and the method of assessing student attainment.

34. A contrary view is that the qualification is all important, especially where it is a 'licence to practise', and that a course is merely one of several means of preparing a student for a qualification. A local development of this view is that there should be an external examination which can be taken by students whether or not they have attended a course at an educational institution. The student's attainment is measured by the examination and the mode of learning is irrelevant to the award of the qualification.

35. The operational model illustrated in paragraph 30, accommodates both points of view, and the many shades between them. For example, in the validation system adopted by the Technician Education Council (TEC) in the UK, the college is responsible for identifying the needs of industry, for drawing up a programme of work to meet such identified needs and a scheme of assessment to measure whether the students have attained the objectives of the programme, and for implementing the programme and the scheme of assessment. TEC approves the programme and the scheme of assessment, monitors the implementation, and awards its Certificates or Diplomas to successful students. The essential elements of the model are included within this system, with the addition of approval and monitoring, i.e. validation.

### Qualifications

36. To be worthwhile, a qualification, especially a technical or commercial one, has to have currency. It has to be respected as providing evidence that at a stated time an individual reached a certain standard. It may, indeed, be given an exaggerated currency and people may expect that the standard is higher and more consistent than can possibly be guaranteed. But like any other currency, a qualification depends on the confidence which the users have in it. If there is a feeling that the standards are variable or the qualification can be obtained by underhand or backdoor means, its credibility will vanish and it will become worthless.

37. There are three main requirements for ensuring the continuing credibility of a vocational qualification. First, the general standard which a qualification implies has to be acceptable to the user (e.g. the employer of the qualified person). Secondly, the decision whether or not a qualification is awarded to an individual has to be consistent and fair. Thirdly, the guarantor of standards and of consistency and fairness has to be beyond reproach, in terms of both accountability and efficiency. These requirements are considered in greater detail in the paragraphs which follow.

## Acceptability of Standards

38. There are three ways of setting a standard:

(a) In norm-referenced standards there is no absolute standard, but the students' results are compared with a norm based on the results of other students, past and present. True norm-referencing implies pretested standardised tests, so that this year's students are compared with previous years' on the same basis. A less rigorous approach is to assume that the standard of students remains the same, and therefore the same percentage should pass, even if the consistency of the examination standard cannot be guaranteed.

(b) In some skills-based examinations (e.g. typing) or in cases where examinations are based on a mastery of clearly stated objectives, it is possible to introduce a criterion referenced standard; e.g. a typist either can or cannot type accurately at 90 words per minute. However, there are many factors which make the introduction of completely criterion referenced standards extremely difficult, for example the comparative quality of the typewriters available, the number of errors considered acceptable. Many apparently criterion referenced standards therefore include within them a normative element. And the criterion itself has to be related to the requirements or expectations of an employer.

(c) Where a new qualification is to be introduced, its standard can receive more rapid acceptability if it can be compared generally to an existing respected qualification, or can achieve recognition as being equivalent to an existing qualification for a specific purpose. The Technician Education Council had to tackle the problem that its award structure was completely different from the National Certificates and City and Guilds. It therefore included the following passage in its 1974 Policy Statement:

'It is expected that the standard of the TEC Certificate will be approximately half way between Parts I and II of the CGLI Technicians' Certificate (T3) or broadly comparable with that of the Ordinary National Certificate. The standard of the TEC Higher Certificate will be broadly comparable with that of the CGLI Technicians' Certificate Parts II and III combined, or with the level of the Higher National Certificate. Any comparability with an existing award in a particular case will depend on the nature of the individual programme'.

Such general statements of intention can be reinforced (as they have in practice with TEC awards) by the specific recognition of the new awards by, e.g. professional bodies (as exempting qualifications) or universities (for entrance to degree courses).

## Consistency and Fairness

39. A qualifying body has to convert the general standards implied by its awards into a series of rules which determine whether the level of performance of any individual deserves a qualification. It will also

usually be necessary to have procedures or rules to determine what happens to candidates who marginally fail to measure up to the standard required for a qualification ('borderline cases') and for particular hard cases (e.g. illness or disablement). The credibility of an award will suffer if the rules are applied inconsistently from individual to individual or from year to year, or if they appear to be unfair to any category of student.

40. The concept of consistency is often also applied to comparability of standards as between different colleges. Clearly it would be unfair if a student attended one college and failed, whereas if he had been educated at another he would have passed - not because of a better education but because of a more lenient college-set examination. Unless an external examination is set on a common syllabus, precise comparability is not possible. If colleges are offering different local specialisms like is not compared with like. It can be argued that precise comparability of standards is not essential for employment and what is required is a guarantee of an acceptable overall standard coupled with as much information as possible on the syllabus covered by the students and the level reached in the various subjects. This is what the Technician Education Council sets out to do with its validation and monitoring procedures and its 'profile' certificate.

#### Efficiency and Accountability

41. It is not enough for a qualifying body to make the right decision. It has to communicate the decision, directly or indirectly to the candidate, and print his certificate or diploma in such a way that it will be difficult to counterfeit. It needs a record keeping system so that if, in subsequent years, there are queries about an individual's right to a qualification, these can be dealt with by the qualifying body.

42. There also has to be a degree of accountability. In the UK, for historical reasons, technical and commercial examinations were largely administered by educational charities with no apparent accountability, but which in practice had a very high degree of accountability within their committee structure. In course of time the international reputation of their awards has reinforced this 'informal' accountability. In other countries, and in the recent establishment of new award-making bodies in the UK, accountability has been achieved by making them part of the public education service (even if financed by examination fees), either with government imposed constitutions or by actually becoming government departments. Yet another form of accountability is achieved by professional institutions, which are accountable to their membership for the qualifications which convey membership rights.

#### Resource requirements for qualifications

43. A qualifying body therefore requires staff who are

- (a) sensitive to the requirements of the users
- (b) able to develop (directly or through consultants) norm-referenced or criterion-referenced standards
- (c) able to negotiate recognition for its awards
- (d) competent administrators (with access to adequate security printing and record keeping and, in many cases, data processing)

- (e) accountable for the decisions which affect the careers of individuals.

### Measurement of Attainment

44. When a student's future career depends on gaining a qualification, and gaining a qualification depends on the measurement of his attainment, then great trouble must be taken to ensure that such measurement is as accurate, fair and consistent as possible. It is not sufficient to take the line that any competent teacher of a subject can set an examination paper and mark it, and that all students with 40% or better can pass - because such an action gives no guarantee of accuracy, consistency or fairness. The measurement of human attainment is a complex task calling for a high degree of professionalism. There is a large literature covering the necessary techniques.

45. In particular, measures of attainment should be valid, reliable, efficient and have beneficial rather than detrimental side effects. A more detailed explanation of these qualities is to be found in the literature on examinations; a short bibliography is included in Appendix VI. In this section we are concerned with the operational requirements to achieve these qualities.

46. The ways in which these qualities will be achieved will depend in part on the examination methods used, as defined in Part 1:

- (a) whether the examination questions are banked or traditional
- (b) whether the examination is set and marked:
  - (i) internally without validation
  - (ii) internally with validation
  - (iii) externally.

### Validity

47. An examination or a question is valid if it tests what it purports to test and if what it purports to test is itself appropriate (i.e. whether the syllabus itself is valid - see paragraph 66).

48. Assuming that the syllabus is valid, then the validity of the examination paper as a whole depends on it having a balanced relationship to the syllabus. The questions which have to be asked are:

- (a) Is there a wide coverage of the syllabus?
- (b) Is there an appropriate balance between syllabus topics and abilities?
- (c) Is the coverage consistent from year to year?

49. In turn the validity of a question depends in part on whether it is set on the syllabus and in part on whether it has been designed correctly. The sort of questions which have to be asked are:

- (a) Is the content realistic and important or merely trivial?
- (b) Does it test appropriate knowledge and abilities?
- (c) Does it avoid dependence on inappropriate abilities

(e.g. is it in practice testing writing ability although the real need is to test theoretical knowledge)?

(d) Is the response expected from the student consistent with the abilities to be tested? (e.g. if a sketch is required is this meant to be illustrative or to prove that the student can draw?)

(e) Is the response expected from the student appropriate to his knowledge at that stage of his educational progression?

(f) Is the question intelligible to the student? (e.g. is it reasonably 'culture-free'?)

(g) Is the question ambiguous? Does it give away the answer?

(h) If the question involves a problem or equation, can it be solved from the given data?

50. The answers to the questions in paragraphs 48 and 49 require the exercise of judgement. The disciplines of establishing a test specification and editing or moderating, enable the judgement of several experts to be brought to bear. If the system involves banking, the discipline of pre-testing or field testing then enables the students' responses to the questions to be gauged. There is no absolutely certain formula for ensuring validity, but the above disciplines, properly applied, are the normal methods adopted by professionals in the field to ensure that examination questions are as valid as possible.

### Reliability

51. A reliable examination is one which for similar groups of students produces results which will be consistent from one occasion to another. If it fails to do this, then any qualification given as a result of the examination will differ in standard from year to year.

52. Reliability is measured on a mathematical scale and, unlike validity, can be stated to a precise degree of accuracy. There are, however, major practical problems of estimating the reliability of attainment tests: different question papers have to be used on each occasion for reasons of security; there is a time limit to the amount of remarking and analysis which can be undertaken before a result is issued; and students cannot be made to sit a second set of papers to compare their results.

53. One approach to the solution of these practical problems is to bank questions - usually objective test items. By keeping a record of the history of a question in terms of how well students have performed, or by pretesting a new question, an estimate of the reliability of questions and therefore of an examination paper as a whole can be built up before the examination is taken. Pretesting, however, requires a reasonable number of students if it is itself to be reliable. City and Guilds aims at 300, and it seems likely that once the pretest population falls below 120 there will be problems in establishing the reliability of the items.

54. In traditional examinations the approach is different. If the student population has the same characteristics as that of previous years then, if the examination is completely reliable, the same proportion

should pass at a given pass mark. Taken a stage further, the pass mark can be adjusted so that the same proportion pass each year, i.e. a correction is made to compensate for any loss of reliability. Even if the numbers of the candidates are about the same, there may be other factors affecting their success rate: e.g. a change in employment patterns, standard of recruitment, course structure (including the introduction of a lower or higher course). A college with good records of previous classes and knowledge of present students is in a better position to gauge how this year's class compares with the past than a more remote external examining body. Indeed the latter should only follow this sort of procedure if the student entry is large enough for the examining body to be reasonably confident that individual or local variations will be ironed out in the 'sample'. Theory, based on 'the standard error of the mean', does not suggest any absolute size for the student entry, but shows that once it drops below 500, there is a severe loss of reliability and that 100 is certainly too low for the use of this procedure.

55. The third approach to reliability is that adopted by the Technician Education Council. There are three coarse grades (Pass with Merit; Pass; Fail) and most students can be allocated quickly and reliably into one of them. All cases of doubt are treated as borderline and considerable time and trouble is spent considering them. This approach is more readily undertaken by a college setting and marking internally than by an external examining body.

56. It will be seen therefore that standards can only be maintained if the measurements of attainment are reliable, and that achievement of reliability requires a knowledge of basic theory, maintenance and analysis of records, and an insight into the practical problems of examining.

### Efficiency

57. There are two aspects of efficiency. Firstly, an efficient scheme of assessment is one which makes the best use of the resources spent on it and makes no unnecessary demands on students, teachers or examiners. The second aspect relates to the efficiency of examination administration.

58. In any system of measurement of attainment (whether internal or external) the following operational stages are necessary:

- (a) Drawing up a test (examination) specification
- (b) Compiling a paper (if a bank is used) or having an examiner set a paper (in a traditional system)
- (c) Editing (bank) or moderating (traditional)
- (d) Printing
- (e) Arranging for the invigilation of candidates
- (f) Marking
- (g) Analysing the students' results
- (h) Reporting the students' results
- (i) If a system is based on banking, analysing and recording the way in which the questions performed

(j) Feedback and evaluation.

59. If the examination is external there are the following additional stages

(a) Determining and publicising an examination timetable

(b) Arranging for candidate entry and collection of any fees

(c) Despatch of question papers

(d) Arrangements for return of candidates' scripts and marksheets (and if a banking system is used, question papers)

(e) Dealing with queries.

60. The development of an efficient administrative system for examinations which lead to qualifications is thus a major undertaking, even if the examinations are internal, but particularly if they are external.

#### Side Effects

61. All methods of measuring attainment have some influence, either beneficial or harmful, upon the related courses. Inappropriate or inefficient methods, or failure to follow recognised examination practice, may lead to undesirable side effects. End of course examinations, for example, are sometimes said to encourage rote learning at the expense of critical thinking, but this is likely to be because the examination questions are poor and do not test the right abilities.

62. External examinations can have powerful side effects outside the confines of the classroom. Publishers are more likely to invest capital in producing text books if there is an external examination syllabus. The dissemination of new techniques is often much quicker if they are taken up by an external examining body; the prime examples are City and Guilds and multiple-choice testing and the Technician Education Council with learning objectives (although the latter was through a validation system rather than examinations).

63. The side effects of examinations cannot always be predicted in advance and the elimination of undesirable side effects and the encouragement of beneficial ones calls for particular qualities of sensitivity and knowledge on the part of staff responsible for examinations - whether internal or external.

#### Resources for the Measurement of Attainment

64. The measurement of attainment requires personnel who are

(a) expert in the theory and techniques of educational measurement

(b) expert in their own subject matter for the setting and marking of papers

(c) efficient administrators, capable of handling a complex, multi-stage system with high security requirements



- (d) sensitive to the effect of examinations on the students and teachers, and indeed on industrial performance.

Measurement of attainment also requires physical resources such as duplicating or printing facilities, secure storage, and often optical mark reading and computing facilities.

65. Probably no external examining body can hope to have all such expertise or physical resources 'in house' and so there is a great dependence on part-time assistance, and use of other outside services (e.g. printing and computer time). An educational institution may have all the necessary expertise 'in house', but the academic tradition often prevents it from being applied - hence the need for external moderation.

### Examination Syllabus

66. An examination syllabus is a bridge between the needs of employment and what it is practicable to test in an examination. At one end it depends on an analysis, or a series of options, of what the needs of employment (including career progression) are and are likely to be, given the time-lag between syllabus production and the emergence of the first qualified personnel. At the other end the syllabus is a statement of what the examination will cover. It is now normal, but not universal, to set it out in 'learning objectives' - what the student can do at the end of a course - and to accompany it by a test specification to show the weight to be given to the various parts of the syllabus and to the learning categories involved (e.g. knowledge, comprehension, application, synthesis, psycho-motor skills). In between, account has to be taken of the structure of the educational system, the resources available to the educational establishments and the probable characteristics of the students. An examination syllabus is almost always a compromise, but it is essential that its content has validity for the employment and educational system it services; otherwise the examinations based on it will themselves not be valid.

67. Because of the need to bring together the various factors described above, it is normal for examination syllabuses to be put together by committees with a membership able to express the conflicting points of view. It is not uncommon for most of the work to be done by small groups, exchanging comments by correspondence where long distances are involved.

### Resources for Syllabus Development

68. Syllabus development requires substantial resources, to provide the initial analysis of industrial need, to write the content, to service committees or postal circulation and, in particular, to turn general curriculum statements into specific learning objectives, and to draw up a test qualification.

### Validation

69. Validation is the recognition by an external qualifying body of college courses and examinations. It can take many forms. At the minimum the college submits its examination syllabus for approval, then its question paper and finally a sample of marked scripts. At the other end of the scale the validating body concerns itself with course content, college staffing and physical resources and may send a team to visit the college before the course is approved. It may then carry out regular monitoring of the running of the course and of the assessment or examination of students, using a

combination of visiting moderators and statistical analysis of performance. The main benefit of a validation system is that it enables local initiatives and special requirements to be rewarded by external qualifications, but only approximate comparability of standards can be achieved.

### Resources for Validation

70. The resources required by the qualifying body for validation will therefore vary according to the extent of involvement of the validating body in the work of the college. At the least it can be a simple extension of external examinations; City and Guilds, for example, has simple standard procedures for 'The Administration of College-Set, Institute-Assessed Syllabuses and Examinations'. Towards the other end of the scale the Technician Education Council has a strong team of specialist advisory officers working with validating committees for each main subject area, a large cadre of part-time moderators (and a few full-time) and a complex data processing system; and even the Technician Education Council keeps to the minimum the visits to colleges by course approval teams.

71. Validation also puts a greater responsibility on the colleges, and the requirements of a validating body can impose a strain on college resources, particularly if it changes from a system based on external examinations to one based on validation. Again, the college resource requirements will depend on how far the validating body involves itself with the detail of the work of the college.

### Conclusion

72. The decision-making and administrative aspects of operating a qualification system, although vital to the future career of individuals, require far fewer resources and less highly trained or specialist staff than either the measurement of attainment or validation and their supporting administrative systems. The measurement of attainment, if undertaken externally to the students' college, requires a candidate entry of several hundred if the requirements of reliability and norm-referenced standards are to be properly met. The implications of both these points are considered further in Part 4 of this report.