education can also promote the further development of autonomy as far as the choice of study objectives, critical appraisal of competing schools of thought, and problem-solving are concerned. It includes individualizing procedures conducive to student autonomy and academic socialization.

Special methods have been developed for use in non-contiguous communication, including counselling, course development, the application of media, and administrative work, which rely on principles of instructional design and dialogue. Conversational approaches and general empathy have been shown to be conducive to students' satisfaction and goal attainment. On the basis of investigations of empathy approaches and other aspects 1, on the one hand, conclude that predictive theories of distance education are possible and that a beginning has been made, yet, on the other hand, concede that empirical studies testing theories/hypotheses may cause both interpretative difficulties and modifications of assumptions without necessarily categorically refuting (falsifying) them. Rumble may well be right when he says that they are unlikely to be 'conclusively falsifiable in the same way that, for example, the discovery of a black swan falsified the theory "All swans are white".' (Rumble 1992: 112).

Theoretical approaches more concerned with the economic, social, political and cultural contexts of distance education are sometimes asked for and are no doubt possible. Attempts in this direction occur in Campion and Guiton (1991), Edwards (1991), Evans and King (1991), Evans and Nation (1992) and elsewhere. See further pp. 212–13 below. Theorizing in this area faces problems regarding the separation of scholarship from value judgements (Holmberg 1992 and Ljoså 1991).

Finally, it is important to recognize that distance education is a separate kind of education, which cannot be regarded as a substitute for conventional schooling because of its openness to adults gainfully employed and/or fully occupied with family life, its independence of face-to-face meetings, classes, and generally of time and place, its combination of mass-communication and individualization, its potential for student autonomy, and its special methodology.

Chapter 10

THE ROLE AND CONCEPT OF EVALUATION

Educators and society at large find it increasingly important to evaluate the various contributions made to education. This is because of the general desire to safeguard the highest possible educational quality and also to ensure that money is invested in a way that yields the highest possible educational output. Evaluation is a general educational concern with some special implications for distance education. The purpose is usually to find out to what extent teaching and learning lead to expected results and acceptable standards. The knowledge acquired by evaluation studies can be used as a basis for improving the teaching–learning system as well as for describing and judging it.

The term 'evaluation' denotes different things in different contexts. Sometimes it refers to the assessment of students for the purpose of awarding marks, sometimes to the judgement of complete educational systems. Evaluating these implies an appraisal of their status in society, the relevance, quality, quantity, and results of their teaching and their impact on education, training, and the labour-market (Tate 1986) including in many cases consideration of their accessibility to various social groups, i.e. equality. This appraisal of the contribution of educational systems is usually related to the costs that they incur. Examples of such evaluation of distance-education systems are given in Keegan (1990 Part IV).

Particularly in the Anglo-Saxon parts of the world, evaluation of educational activities has been to some extent spurred on

by the barbs of an accountability movement which attempts to respond to economic adversity through 'rationalization' of higher education provision and the promises of an

education can also promote the further development of autonomy as far as the choice of study objectives, critical appraisal of competing schools of thought, and problem-solving are concerned. It includes individualizing procedures conducive to student autonomy and academic socialization.

Special methods have been developed for use in non-contiguous communication, including counselling, course development, the application of media, and administrative work, which rely on principles of instructional design and dialogue. Conversational approaches and general empathy have been shown to be conducive to students' satisfaction and goal attainment. On the basis of investigations of empathy approaches and other aspects 1, on the one hand, conclude that predictive theories of distance education are possible and that a beginning has been made, yet, on the other hand, concede that empirical studies testing theories/hypotheses may cause both interpretative difficulties and modifications of assumptions without necessarily categorically refuting (falsifying) them. Rumble may well be right when he says that they are unlikely to be 'conclusively falsifiable in the same way that, for example, the discovery of a black swan falsified the theory "All swans are white"' (Rumble 1992: 112).

Theoretical approaches more concerned with the economic, social, political and cultural contexts of distance education are sometimes asked for and are no doubt possible. Attempts in this direction occur in Campion and Guiton (1991), Edwards (1991), Evans and King (1991), Evans and Nation (1992) and elsewhere. See further pp. 212–13 below. Theorizing in this area faces problems regarding the separation of scholarship from value judgements (Holmberg 1992 and Ljoså 1991).

Finally, it is important to recognize that distance education is a separate kind of education, which cannot be regarded as a substitute for conventional schooling because of its openness to adults gainfully employed and/or fully occupied with family life, its independence of face-to-face meetings, classes, and generally of time and place, its combination of mass-communication and individualization, its potential for student autonomy, and its special methodology.

Chapter 10

THE ROLE AND CONCEPT OF EVALUATION

Educators and society at large find it increasingly important to evaluate the various contributions made to education. This is because of the general desire to safeguard the highest possible educational quality and, also to ensure that money is invested in a way that yields the highest possible educational output. Evaluation is a general educational concern with some special implications for distance education. The purpose is usually to find out to what extent teaching and learning lead to expected results and acceptable standards. The knowledge acquired by evaluation studies can be used as a basis for improving the teaching–learning system as well as for describing and judging it.

The term 'evaluation' denotes different things in different contexts. Sometimes it refers to the assessment of students for the purpose of awarding marks, sometimes to the judgement of complete educational systems. Evaluating these implies an appraisal of their status in society, the relevance, quality, quantity, and results of their teaching and their impact on education, training, and the labour-market (Tate 1986) including in many cases consideration of their accessibility to various social groups, i.e. equality. This appraisal of the contribution of educational systems is usually related to the costs that they incur. Examples of such evaluation of distance-education systems are given in Keegan (1990 Part IV).

Particularly in the Anglo-Saxon parts of the world, evaluation of educational activities has been to some extent spurred on by the barbs of an accountability movement which attempts to respond to economic adversity through 'rationalization' of higher education provision and the promises of an

emerging technology of evaluation that it can provide the means by which 'rationalization' of – if not rationality in – educational provision can be achieved.

(Kemmis and Hughes 1979: 7)

The dangers inherent in this rationalizing approach are evidently that the complexity of the educational tasks may be under-estimated and that academic concerns are replaced by administrative ones. Kemmis has developed a series of principles that are meant to 'create an image of evaluation as the process of marshalling information and arguments which enables interested individuals and groups to participate in the critical debate (the process of self-reflection) about a program' (Kemmis 1980: 4).

The principles and practices of distance-education organizations applied to evaluation at the beginning of the 1990s are described and analysed in Schuemer (1991).

PRINCIPLES OF DISTANCE-EDUCATION EVALUATION

Any appraisal of educational quality must be at least partly based on a micro level, i.e. individual and specific educational endeavours. In distance education, courses can be taken as starting points. Course is here taken to mean a defined part of a curriculum in one particular subject, for example a course of basic algebra for university students of mathematics or a course of English grammar for the upper stage of grammar schools. In this context I use the term 'course,' to denote not only the course materials but also the whole process of interaction between the course materials, the students, the tutors, and the supporting organization as a whole (Thorpe 1979b).

Course evaluation as discussed here is related to, but narrower than, curriculum evaluation. Course evaluation is meant to provide evidence as to whether or not a given course or course draft is reasonable and effective.

The evaluation of distance-study courses includes various attempts to obtain a good grasp of how courses function, how effective they are, and how they are received by students and others concerned. One type of course evaluation concerns the assessment of students' progress; another that of the teaching effects of the system.

Assessing students

Studying the processes of students' learning is a prerequisite for helping them (see pp. 32–6). The assessment of students' progress is needed both to give students feedback so that they know how they succeed, and, in all cases where diplomas or graded certificates are required, to provide the basis for marks. In all examination systems it is important that tests should be both valid and reliable. If continuous assessment (rather than final examining only) is to some extent applied by distance-education institutions, it does not follow that these demands can be neglected. In the light of the literature available on psychometric and other considerations applicable to testing generally Lewis shows the importance of paying due attention to the development of exercises and tests:

Let us suppose we took the trouble to analyse the co-occurrence of mistakes on (say) our computer-marked assignments; students who get question 1 wrong also tend to get questions 3, 8, 17 and 24 wrong. This suggests that the five questions are all actually tapping the same underlying dimension of confusion. This being so, we may be marking the student down five times over for having made just one mistake.

(Lewis 1972: 119–20)

In one respect testing has special relevance for distance education. If students are offered the possibility to start, interrupt, and finish their study when they wish, to pace themselves, and generally to organize their study as they see fit, there must be a great demand for frequent examination opportunities in all subjects in which formal qualifications are required. A mastery-learning system allowing individual students to be examined when they feel they are ready for an examination (as according to the Keller Plan) would seem to be called for. This requires a bank of validated test items. On the Keller Plan, as related to distance education, see Holmberg (1981a); Coldeway and Spencer (1982).

In other respects, examination problems can be disregarded here; distance education causes no problems for assessment other than those occurring in all examination situations. The exceptions are those of an organizational and administrative type, for instance arranging decentralized written and oral examinations

in special study centres, under the auspices of other educational bodies, embassies, consulates, etc. Nevertheless, assessing students' progress is an important matter to distance educators and has been thoroughly examined in this context (Lewis op. cit.).

A special form of student assessment is required when completely individualized study programmes are investigated. The PERC approach (Program Effectiveness and Related Costs) developed at the Empire State College in the US State of New York is a case in point. The individualization is brought about by individually designed degree curricula and learning contracts. The basic evaluation question studied is 'What kinds of students working with what kinds of faculty in what kinds of learning programs change in what ways at what costs?' (Lehmann and Granger 1991: 104). See further Palola et al. (1977) and Lehmann and Holtan (1988).

Evaluating courses

Educators may be interested in evaluating a distance-study course in order to be able to describe it properly, to provide a declaration about its characteristics and known effectiveness. This may serve as a guide either for people considering taking a course on the subject concerned or for study counsellors who are looking for detailed information about study opportunities that might be useful to students whom they are advising. This type of evaluation has been called summative and is a kind of product evaluation. It is a task for unbiased researchers, not for those who are engaged or have been engaged in the development of the course concerned.

Distance educators are not usually primarily interested in summative evaluation, however. They are, above all, concerned with evaluating how well a course can and does help students to attain their objectives of study and how it corresponds to their requirements and expectations. They usually want to investigate courses with a view to improving them. This is what is known as formative evaluation. Formative evaluation does not aim at passing judgement but is to be seen as a component of the development work. The findings of the evaluation are meant to influence, modify, or radically change the course presentation and the study

procedures. Formative evaluation can have different functions, however.

One function of formative evaluation is to make the presentation coincide as closely as possible with the purposes which have led to the creation of the course. This 'ends-dominated' approach seems to prevail. It is primarily concerned with making learning more efficient. However, a learner-centred approach claims to regard the function of formative evaluation differently. The learner-centred approach stresses the desirability of finding out

whether the concepts, procedures and criteria of a subject (have) been made sufficiently accessible, whether the problems involved could have been presented more cogently or clearly, whether the exercises or aids to thinking employed in the course really helped students to appreciate the concerns of the course makers. An evaluator would also want to know to what extent students had been able to pursue their purposes and interests through the course, what habits of thought or ways of seeing the world they might already have that may be getting in the way of aiding their understanding – and so on.

(Mace 1976: 27)

The question is whether there is really a conflict between the two aims; effectiveness and adaptation to the student's situation. The latter appears to be a prerequisite for the former.

It is important to evaluate not only the pre-produced course as a product but also the tutoring and counselling belonging to it. This process evaluation is looked into in Thorpe (1988), Ganor (1991) and Naylor, Cowie and Stevenson (1990).

BASES OF EVALUATION

Comparing achievement with objectives and performance standards

How well students achieve in relation to the purposes of the course is a common basis for the judgement of the merits of the course. The effectiveness is then often measured by the use of pre- and post-tests. This type of evaluation may concern not only whole courses or course units but also individual parts of

the presentation which concern specified items of knowledge and proficiency and for which desired performance standards have been laid down. The degree of attainment of each individual study objective is measured. If a sufficient majority of the students attain the performance standard, all is well. If not, either the course (the printed and/or recorded or broadcast presentation), the non-contiguous two-way communication, or some other course component must be revised on the point or points concerned. Alternatively, the objectives must be revised on the basis of the assumption or conclusion that they are not realistic.

This approach relies to a great extent on the definition of so-called behavioural objectives, i.e. study objectives which express what the student is expected to be able to do after he has completed a course or a course unit. These objectives can be derived from the known requirements of a job, from an analysis of what will be necessary for a new type of task, or from existing curricula and/or examination requirements. Since the 1970s distance educators have tended to go in for (or pay lip-service to) consistent endeavours to determine in advance what the students should be able to achieve and have expressed in quantitative terms how well they should be able to do it. The degree of objective attainment is then assessed with a view to improving the course.

Performance standards are not always specified in advance. Some evaluation based on defined learning objectives measures students' achievements after the course to find out what the performance standard is for each objective and does so without any pre-conceived requirement levels.

Consulting experts

An entirely different basis for the judgement of distance-study courses is the opinions of experts. This evaluation procedure means that the course is submitted to criticism by subject specialists recognized as educational authorities. No less than 29 out of 33 respondents to a questionnaire sent to 79 European distance educators in 1980 declared that they use this type of evaluation (Holmberg 1981b). There can be no doubt that expert opinions can influence courses in a very useful way by testing the intelligibility of texts, identifying difficulties and suggesting improvements. There are certain dangers connected with this procedure,

however. First, it is very probable that only the course materials rather than the whole learning experience, inclusive of the two-way communication, is studied and commented on and that the course text is criticized on the basis of the criteria of a traditional educational text, without due attention being paid to the particular requirements of distance-learning students. Second, the experts who are consulted may well be regarded by course authors as a new target group; authors may write with a view to satisfying the expected opinions of the experts rather than the needs of students. This may lead to overloading and to unnecessarily scholarly presentations.

The use of internal experts is less problematic. Tutors' comments on course presentation, after they have graded and commented on a number of assignments, have proved very helpful. Some distance-education institutions expect their course authors to function as tutors of their courses for at least some months. This leads to useful feedback for course-improvement purposes and facilitates using student feedback (on which see Nathenson and Henderson 1980).

Consulting future employers and/or teaching bodies

Those employers or teachers in universities or institutes of further training who are to receive students may also be used as evaluators. It is then their task to relate the course and what it teaches to the requirements of the job or the further study at which the student is aiming. It is probably true to say that this group of evaluators have influenced course development in connection with the introduction of large-scale curriculum reform rather than by suggestions relevant to individual courses. In my experience it is difficult to induce the evaluators to specify the actual requirements of the job or further study concerned. They are more likely to express conventional views and say what is expected of them on the basis of academic traditions and other conventions. Whether this is due to a deficient awareness of the actual needs or merely to difficulties in articulating them seems uncertain. However, if this snag, which largely concerns finding the right representatives of the recipient organizations for the evaluation tasks, can be overcome, the procedure should be very valuable.

Investigating students' views

The students' own opinions of the course that they are taking are, of course, of paramount importance. Studies of students' attitudes to courses have been undertaken by a great number of institutions. Students may be asked what they think about a course generally, how motivating and interesting it is; they may be requested to comment on specified course items, on a particular method, etc. Very often questions are asked with a view to improving both the course presentation and the procedures that are applied to the interaction between students and the supporting organization. In some cases, as well as indicating where there is a problem, students may contribute to diagnosing the difficulty. An early example from the Open University illustrates such questions (see Fig. 8).

Twenty-six of the 33 European respondents referred to studies of students' opinions as a type of evaluation that they applied.

EVALUATION METHODS

Distance educators all over the world base their evaluation on one or more of the types of criteria discussed above (and others). While most seem to consider a comparison of students' achievements with the course objectives to be an essential type of evaluation, they adopt this procedure less often than that of asking the parties concerned about their opinions. Students' attitudes and expert opinions seem to play a very important part in the evaluation activities of distance-education institutions.

For obvious reasons the publicly funded, large distance-education institutions are able to apply more sophisticated evaluation procedures than most of the institutions that have to finance their work by means of students' fees only. As expected, the Open University has contributed much to the art of distance-course evaluation, not least to a method of early formative evaluation, referred to as developmental testing.

Developmental testing

This is a term that denotes try-out procedures characterized by small experimental groups taking courses in preliminary versions before these are offered for more general use. Developmental

a. Was section X clear to you?

Yes ( )
No ( )

b. If not, please say briefly exactly what was not clear. (Space for answer)

c. Can you suggest any way in which section X could be improved? (Space for answer)

d. Without referring back to section X, what do you think it is important for you to remember? (Space for answer)

e. Were the examples on pp 00–00 clear to you?
Example 1 Yes ( )   Example 2 Yes ( )   Example 3 Yes ( )
No ( )   No ( )   No ( )

f. If not, please say briefly exactly what was not clear. (Space for answer)

g. Were there: too many examples? ( )
about the right number of examples? ( )
too few examples? ( )

h. Did you do exercise Y? Yes ( )
No ( )

i. If you did, which parts did you get wrong?
(a) ( )   (b) ( )   (c) ( )   (d) ( )   (e) ( )

j. Did exercise Y have: too many parts? ( )
about the right number of parts? ( )
too few parts? ( )

k. If you did get any parts of exercise Y wrong, what did you do about it? (Space for answer)

Your comments: (Space for answer)

Figure 8 Open University example of specified questioning of students' attitudes
Source: Henderson and Nathenson 1976: 37

testing is thus a type of preliminary formative evaluation. In literature on programmed learning, formative evaluation very often means the same thing as developmental testing, as described here. For distance-education institutions, however, it may be practical to limit the latter term to evaluation work done

before the course is offered to students as part of a syllabus and to use Scriven's term 'formative evaluation' for the investigation, with a view to modifying later editions, of how a course already on the market functions.

Inducing experimental groups to take a preliminary version of a course naturally offers many possibilities to try out alternative presentations, to ask the experimental students for their views, and to check on their attainments in relation to the objectives specified. There are difficulties, however, in that these experimental students do not work under the same conditions as normal students do, that their use of the course may differ from that of the normal students, and that they may pay less attention to the course than regular students do. Experiences of developmental testing in the Open University seem, at least partly, to bear out these objections whereas, to my knowledge, no traditional correspondence schools which apply developmental testing have reported on difficulties of this kind. An OU study showed that experimental students spent 30 to 50 per cent less time studying than actual students who later studied the same units (Henderson and Nathenson 1976: 33; see also Bartels and Wurster 1979: 2).

It seems to be fairly common practice in general educational evaluation to pay a fee to experimental students for their contribution to developmental testing. This may motivate them to do a reasonable job; on the other hand, it may lead to differences in attitudes and requirements between them and the regular students. Any such differences, of course, detract from the relevance of the findings to which they contribute. It is interesting to note that only one of the 33 above-mentioned respondents to my questionnaire pays their experimental students, whereas both the Open University and the FernUniversität compensate theirs for their work. Nine of my 33 respondents declared that they do developmental testing. The number of experimental students engaged varies between two to five and 300 to 500.

Some interesting studies of developmental testing were published at an early stage, among them Henderson and Nathenson (1976) and Henderson et al. (1977). The former find it important in developmental testing

that a complete educational package . . . should be tested; that the evaluation should combine both objective and subjective data and be capable of identifying faults in the

material and generating potential solutions; that testers, and the conditions under which they study, should closely approximate the target student population; and that the developmental testing programme should operate strictly within the constraints of the . . . course production system.

(Henderson and Nathenson 1976: 33)

[They call testers those whom I call experimental students.]

Evaluation as illumination

The information acquired, through questionnaires and through analyses of students' achievements evidently does not give a complete picture of the students' situation, their interaction with the course, and their particular difficulties. For this reason, attempts have been made to find out not only what students' learning conditions are but also what their life in general is like in relation to their studies. This approach, which thus pays particular attention to the study milieu, is sometimes referred to as 'evaluation as illumination'. Typically 'in illuminative evaluation there are three characteristic stages: investigators observe, inquire further, and then seek to explain' (Parlett and Hamilton 1972).

In-depth interviews with individual students are made, with a view to understanding what happens to them, to their life while they study, rather than sampling the general reactions of the students to the course under study, which is what happens when questionnaires are used.

Work of this kind is sometimes combined with efforts to make students participate in the evaluation. It is considered important that the students should not be objects of study only but should be subjects themselves. Participative work of this kind can lead to a method of action research, as stressed in a FernUniversität study of the lives of distance students (Abels et al. 1977). It should be stressed that Parlett and Hamilton themselves do not combine their approach with action research. The evaluator, as they see him,

makes no attempt to manipulate, control or eliminate situational variables, but takes as given the complex scene he encounters. His chief task is to unravel it; isolate its significant features; delineate cycles of cause and effect; and comprehend relationships between belief and practices, and

between organizational patterns and the responses of individuals.

(Parlett and Hamilton 1972: 16)

Quantitative analyses for formative and summative evaluation

Course development and tuition (the latter word being used in its British sense, meaning tutoring, teaching, not in its American sense) are based on assumptions of what will work. These assumptions are tested more or less systematically. Statistical methods to test their validity are commonly applied by educationists in terms of experiments and studies of what has actually happened in relevant situations. This applies to distance education as well as to traditional forms of education.

If the purpose of course evaluation is to improve the courses evaluated, and that is evidently the main purpose behind distance course evaluation, then what is being sought is a theory of what the optimal procedure is in each case. Such a theory will enable the consequences of measures taken to be predicted, that is methods and/or media to be used in a particular way, etc. Further, it may be possible to indicate what can be done to attain certain purposes, i.e. to prescribe a technology. The evaluation may or may not falsify the theory.

Course evaluation includes both the description of what works and the search for causal relations. For both functions there are useful quantitative, statistical methods. Descriptive statistics may organize the findings in a handy form, including graphical forms of presentation (e.g. the histogram) and frequency distribution.

When testing predictions, i.e. statements that a change in one respect (the so-called independent variable) will lead to change in another respect (the dependent variable), it is of vital importance that no seemingly irrelevant variable is allowed to change systematically with the independent variable. This could happen if, for instance, in the division of students into an experimental group and a control group gainfully employed students were allocated to one condition and full-time students to another, or, if in a study of instructional design or communication frequency one group were made to consist of students taking part in supplementary face-to-face sessions, whereas those belonging to another group, with which the former is to be compared, benefit
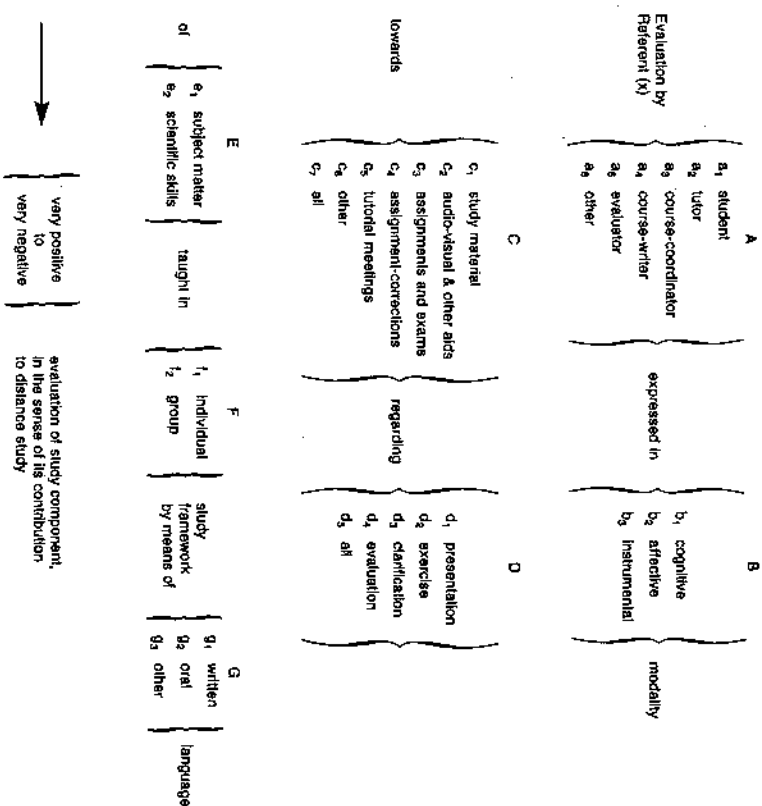
exclusively from mediated teaching. Here we are concerned with statistical inference trying to establish the cause of differences between findings. Conclusions of this type can evidently influence course development radically. Procedures for various types of quantitative investigations are in use. They include scales for measuring the attitudes of students (and evaluators), such as the Likert scale (Likert 1932) and Osgood's semantic differential leading to 'polarity profiles' (Osgood 1952).

Quantitative methods are chiefly used for analyses of students' achievements. They may concern the assessment of students for examination purposes, but seem especially important in formative evaluation for the purposes of testing the effectiveness of procedures, methods, and media. As this is related to the objectives of the course under examination, it has been found useful sometimes to formulate the objectives in the form of end-of-course examination questions and problems: 'By defining the objectives of a course in terms of the tests that the student should ultimately be able to pass, we do much to ensure that our tests are both relevant and equitable' (Lewis 1972: 119).

Measurement techniques have been developed for the assessment of students' achievements and the evaluation of courses and programmes. A study of the practices in this respect of sixteen distance-teaching organizations occurs in Chia (1990).

A very interesting – and evidently highly rewarding – approach is applied at the Open University of Israel. It is based on Guttman's facet theory, 'Specification of course content and its instructional objectives in "course maps" serve as a basis for preparing a teaching syllabus, establishing a computerized bank of questions and assessing all course components' (Ganor 1991: 80). Not only pre-produced courses but also student support and the assessment of students' achievements are subject to this evaluation. The information collected through the evaluation work is used as a foundation for staff development in the sense of didactic training. The technique used implies specifying what is to be investigated in a so-called mapping sentence. An example is shown in Figure 9. Descriptions of this approach occur in Ganor (1988, 1990 and 1991).

Evaluation by Referent (x)

A
a1 student
a2 tutor
a3 course-coordinator
a4 course-writer
a5 evaluator
a6 other

expressed in

B
b1 cognitive
b2 affective
b3 instrumental
} modality

C
c1 study material
c2 audio-visual & other aids
c3 assignments and exams
c4 assignment-corrections
c5 tutorial meetings
c6 other
c7 all
} towards

D
d1 presentation
d2 exercise
d3 clarification
d4 evaluation
d5 all
} regarding

E
e1 subject matter
e2 scientific skills
} of

taught in

F
f1 individual
f2 group
} study framework by means of

G
g1 written
g2 oral
g3 other
} language

→

very positive to very negative

evaluation of study component, in the sense of its contribution to distance study

Figure 9 An example of the 'mapping sentence' applied to course evaluation at the Open University of Israel
Source: Ganor 1991: 83

Reference (A) expresses attitudes in modality (B) towards study component (C) regarding didactic aspect (D) of course material (E) taught in study framework (F) by means of language (G). Range of evaluation (R) is from very positive to very negative, in the sense of study component contribution to distance study.

COMPLETION AND DROP-OUT RATES

Course completion is usually taken to imply success, whereas drop out is interpreted as failure. In distance education this understanding is valid only to a limited extent, at least if by course completion is meant the submission of all of the assignments of the course concerned. If by success we mean goal attainment,

then only knowledge of individual students' goals can help us to decide if a student and a course of study have been successful or not. Céline Lebel illuminates the relations between dropping out and failure after an initial well-considered declaration that when a student – for whatever reason – does not complete a course 'il ne s'agit pas d'un échec académique mais d'un abandon transformé en échec' (Lebel 1989a: 51).

Distance-education programmes are often used by individual students who do not declare either their ultimate goals (self-actualization rather than the acquisition of competence may be their aim) or the period over which they intend to spread their study. Thus it is often impossible to say for certain, unless the students are conscious of their ultimate goals and have made their study intentions known, whether non-completion means interruption, or drop out in the sense of failure, or if it accords with their intentions or plans. The well-established accountant who takes a course in automatic data processing in order to get to know its basic principles and terminology so that he may be able to communicate with computer staff may well reach his goal by thoroughly reading the first few course units, submitting solutions of assignment problems based on these and, as far as the rest of the course is concerned, limiting himself to browsing and looking up items of special interest to him. From a statistical point of view, however, he will appear to be one of the drop outs. This situation is highlighted by a reply given by a successful inventor to a question as to why he had not completed his course:'I am a busy man. I took this course to learn how to solve a certain problem in advanced physics. When I learned that, I stopped sending in lessons' (James and Wedemeyer 1959: 93).

As all practitioners are aware, many successful graduates and students who have successfully passed professional and school or university examinations have refrained from submitting all of the assignments that belong to the courses preparing them for examination. They cannot possibly be referred to as drop outs since they have reached their goals, not only from their subjective point of view but also in the sense of examining bodies. They simply have not made use of all of the assistance offered by their distance-teaching organizations. This may, as explained to the author by a great number of successful former distance students, be due to pressure of time or to personal confidence of adequate examination standards before course completion. Data from a

comprehensive research project confirm 'that submission of assignments is not perfectly indicative of how far a student has reached in his studies' (Bååth 1984b: 31; see also Bååth 1980: 87ff). Bååth here confirms a conclusion of Wedemeyer's (Wedemeyer 1971: 556).

The above-quoted paper by Bååth gives a short but balanced, and in my view accurate, summarizing picture of completion and drop out. After recognizing that it is difficult 'to determine exactly when a slow student should be considered a drop-out' ... 'as in a distinct majority of correspondence schools students are enrolled continuously and are often allowed to keep a very flexible time schedule' (p. 31), and after discussing the completion concept, Bååth sums up:

Although there are some uncertainties connected with the notions of completion and discontinuation in the distance education context, it is fairly safe to say that the completion rate in this type of study is often fairly low ... When non-starters (i.e. students who have registered for the course but who have not sent in one single assignment for correction and comment) are included among non-completers, drop-out rates round 50 per cent are not unusual. The variation between schools and even between courses at the same school is, however, considerable. As an extreme, from 0 per cent to 100 per cent completers have been found in different courses of the same Institution (Ward 1954).

Non-starters are sometimes as frequent as – or even more frequent than – 'real' drop outs. On the whole, most of the dropping out normally happens during the very first part of a correspondence course.

(Bååth 1984b: 32)

The completion issue has been investigated by a great many scholars. Up-to-date reviews of the completion/drop-out problem are provided in Cookson (1990) and Schuemer and Ströhlein (1991).

A general characteristic is that dropping out, when it occurs, usually happens at the beginning of the study. Sometimes the first few months of study should evidently be regarded as a trial period. In agreement with this thinking, the British Open University requires an introductory period of study and, after this period, a reconfirmation of study intentions before a student is

regularly registered. Those who drop out during the introductory period are not included in the university statistics.

Considering the typical students of distance-teaching organizations (adults with families, jobs, and various commitments), it is not surprising to find that the reasons given for discontinuation are in the majority of all cases the pressure of duties, work commitments, travel, illness, lack of time, and similar circumstances.

Some attempts to explain the occurrence of drop out in higher education are well known. Tinto (1975) refers on the one hand to such self-evident causes as lack of interest, academic ability and goal commitment, on the other hand to poor social and academic integration, i.e. to feelings of not belonging. Bajtelsmith (1988), who pays special attention to distance education, modifies this approach and stresses the influence of circumstances outside the study situation, such as family and work. Schuemer and Ströhlein (1991), partly with reference to research done in Germany, also query the applicability of Tinto's approach to dropping out in distance education. In a careful study of drop-out occurrence in university distance education Peters (1992) implies a need for active student support to counteract the drop out problem. However, he does not regard its introduction into the German distance teaching university, the FernUniversität, as feasible, the reason being that the 'active support of students to reduce the drop-out level has no tradition in German universities' (Peters 1992: 258).

Nevertheless in all of the cases where dropping out occurs without goal attainment, counter-measures are evidently desirable (which is also stressed by Peters). In my view the best possible assistance that can be given to students, and thus an antidote against unwished-for discontinuation, is the empathy approach that produces conversation-like real and simulated communication (see p. 55) and personal relations between students and tutors (see pp. 125–7). The findings that Stein, Posz, and Brady (quoted on p. 126) report on support this strongly. Encouraging reminding letters have also proved to be helpful, as shown by Rekkedal (1972b) and others. Further, as discussed on pp. 122–5, both a short turn-around time of assignments that are submitted for correction and comment and a suitable frequency of non-contiguous communication can be of great importance.

However helpful counter-measures may be, the really decisive factor for course completion is the student's personality. Data culled from three German studies carried out in the 1980s show that:

1 The agreement between personal interest and course offer (degree structure) is the most decisive factor for success (continuation of study) and failure (drop out) (Bartels 1982: 11; Bartels 1983: 16).

2 Students inclined to work on their own rather than collectively, i.e. those who do not feel any handicap of isolation but rely on their own initiatives to establish contacts when desired, tend to be successful (Bartels 1982: 18), whereas most drop outs suffer from learning in isolation (Bartels 1983: 24–5).

3 A certain amount of resignation concerning the chances of professional promotion is common among the drop outs (Bartels 1983: 7).

4 The drop outs have 'greater problems co-ordinating the requirements of their jobs, families and study than those continuing their study and are less capable of sustaining heavy workloads and changes in job situation; the latter are more prepared to accept that their personal lives suffer during their time of study' (Bartels et al. 1984: 94).

Apart from the last characteristic, which is concerned with physical and mental strength, and, partly, the general level of satisfaction with the study facilities, this summary indicates the dependence of success on strong study motivation generally and motivation for distance study in particular. As to the handicap of isolation, it is interesting that a Swedish study showed that 24 per cent of the drop outs considered their failure to be due to 'incapacity to study on one's own' (Wångdahl 1980: 54).

That strong study motivation is decisive for success is confirmed by experiences in many different contexts (thus explicitly in Sewart (1983: 168); see McIntosh et al. (1980: 51) on the first intake at the Open University). It should be noted, however, that there are different kinds of motivation and that the study motivation of each individual is likely to be rather complex.

A study by Rekkedal already referred to indicates, on the basis of statistical evidence, that:

1 Practically no relationship could be established between

students' domestic background and discontinuance (Rekkedal 1972a: 17); this is remarkable, as distance students generally stress the importance of encouraging support from husband/wife and other family members (see Bartels 1982: 14 and 1983: 20, confirming this).

2 Older students 'survived' to a greater extent and achieved better results than younger students (ibid. p. 26), which, as far as the first statement is concerned, agrees with Donehower's study of 1968; as to the second statement, Donehower 'found that the oldest group (only 9 students more than 60 years old) received the lowest marks; except for these oldest students, the achievement rose with increasing age of group at least up to about 45 years of age' (Rekkedal 1972a: 26).

3 Not unexpectedly, there were positive correlations between the levels of previous education and both survival and achievement.

The greater success of older and better qualified students and of students already familiar with intellectual work of some kind, as well as the lack of influence of the domestic background, can all be related to motivation. Older, mature, and well-informed students may be assumed to be less likely than younger students to enrol unless they are strongly motivated. Good basic education, relevant prior knowledge, reading habits, and similar background conditions naturally confer advantages and make for initial good results; the maxim 'Nothing succeeds like success' interprets the motivational influence of this. Strength of will, self-discipline, and similar qualities are evidently connected with motivation.

The decisive influence of motivation for goal attainment in distance education is stressed by Sewart in a statement that puts other aspects referred to in perspective:

In the final analysis, we are left with the conclusions that neither age nor distance nor domestic environment nor any other quantifiable term stands out as a salient feature. It is motivation above all else which, despite physical and general social and environmental problems, brings success.

(Sewart 1983: 168)

THE ECONOMICS OF DISTANCE EDUCATION

Sweeping statements about the economics of distance education are hardly possible. Such statements must be differentiated. They

may represent the factual situation that applies in most cases (the private correspondence schools), in well-defined extensive areas (such as the distance-teaching universities), or in applications of distance education to specific functions only (as is often the case in personnel training).

There are different kinds of distance education, and it is important to realize that even seemingly parallel systems include different components and media. The British Open University includes the use of television and radio for its teaching, whereas the German FernUniversität does not. Both these universities run a number of study centres where students are continuously offered tutorials and various media facilities. This is a type of service that is not provided by probably the vast majority of distance-study institutions in various parts of the world, the publicly funded and private correspondence schools. These in their turn may or may not make use of the telephone for oral tutorials. Some distance-teaching institutions provide video-recordings and films to be used during face-to-face sessions or individually in study centres; others limit their media provision to printed material and audio-tapes (cassettes) and/or discs. Whereas some insist on bringing students together for concentrated residential courses, others do not organize any activities of this kind. Some institutions work with student bodies limited to a few hundred or even fewer students. Others are mass-education institutions with many thousands of students. The differences in relation to economics between the two last-mentioned types of providers of distance education are illuminating.

In one application of the former at the university level, the Australian New England system, parallelism with on-campus study is considered valuable and even of vital importance. Periods of residential teaching play an important part. Further, distance 'students are taught and examined by the same staff that teach internal students. ... They study the same courses as those offered internally.... They take the same examinations at the same time as internal students in examination centres' (Smith 1975: 163).

This parallelism has been extended to a requirement that distance study and on-campus study should have the same student/ staff ratio.

It has been found at New England that one lecturer can

satisfactorily teach 50 students in one course, where three courses provide a full-time year of study. In other words, three lecturers are required to teach 50 students full-time, producing a student/staff ratio of about 16:1. Allowing for a drop out of about 20 per cent for external students, this produces a student/staff ratio of about 13:1, which is very close to the ratio considered reasonable for Australian universities generally.

(Sheath 1972: 288–9)

This is a very different approach from the one applied by the large correspondence schools and sometimes by the distance-teaching universities, when courses are developed for thousands of students and various kinds of technology, labour-saving devices, and division of labour are used to attain economies of scale.

The differences mentioned, and others, show that any reference to the costs of distance education must be qualified by a description of what kind is meant. Evidently, pure correspondence study, relying exclusively on the written word as its medium, can be very inexpensive if it is offered on a large scale, so that the cost of each course can be spread out over several thousand students. More sophisticated systems require higher costs per student, even if large numbers of students are provided for. The problem area concerned with the economics of distance education is surveyed in a profitable way in Perraton (1982); Kaye and Rumble (1981); Rumble (1986). Much discussion on economic issues has taken place. A summary with comments relevant to publicly funded tertiary-level institutions is given in Keegan (1990: Part IV).

In order to come to grips with the economics of distance education, comparisons with other forms of study may be helpful. If we compare the costs of reaching a particular educational goal, for instance a degree, by distance education with the cost of attaining the same qualification by conventional study, we should be able to draw important conclusions. Then it is essential to compare both input and output. The input would be the total cost (students' fees, government or other financing and subsidies, the loss of income incurred by students who give up work for study, etc.), whereas the output would be the degree or other study goal reached and possibly even its economic value.

Let us from these points of view look at the most sophisticated

distance-education system known, that of the Open University in the UK. Its size in the UK and importance as a pattern for other distance-education institutions makes such a study particularly interesting. The economics of The Open University was thoroughly investigated at an early stage.

A study by Wagner in 1972 of the costs of the Open University in relation to conventional British universities showed the following results:

1 The average recurrent cost per equivalent undergraduate in the Open University was found to be a little more than a quarter of what it was in conventional universities.
2 The capital cost per student place in the Open University was found to be only six per cent of what it was in conventional universities.
3 The average recurrent cost per graduate in the Open University was found to be equal to that in conventional universities, provided that the Open University had a drop-out rate of 85 per cent. (It is actually less than 50 per cent).
4 The resource cost per equivalent undergraduate in the Open University was found to be about one-sixth of that in conventional universities.

A later study by Wagner (in 1977) confirmed the findings of his first study:

If the drop-out rate in the future does not differ significantly from the past then the average cost per graduate is likely to be below half that at conventional universities.

Finally the resource costs measure the cost to the economy and include therefore the output lost by full-time students not being in employment. This of course increases the Open University's advantage because all its students are part-time. The figures indicate a ratio of 5:1 in the Open University's favour.

(Wagner 1977: 365)

The fact that the highly sophisticated multi-media system of the Open University compares very favourably with conventional universities would seem to indicate that distance education generally can be very economical. To what extent this applies to all procedures and media applied is less certain. What we do not know, for instance, is whether the costs of study-centre activities

or television programmes or various kinds of face-to-face support, all very expensive in relation to the use of printed and written communication, contribute to the effects of the system in relation to their costs. This is a subject well worth investigating. It has been the subject of some research but there is no conclusive evidence available.

In large-scale systems, the costs per student are normally low. However, some small-scale projects have also proved highly cost effective. A remarkable example is the University of Surrey staff-development programme for university lecturers in Southern Asia (see pp. 151 and 167–8).

Perraton, who has carefully studied the economics of distance education on the basis largely, but by no means exclusively, of experiences of teaching in developing countries, cautiously summarizes his findings by saying that

it is possible only to claim that there are circumstances in which distance teaching looks attractive from an economic point of view. Economies of scale are possible. But distance education characteristically has high fixed costs and, with relatively low student numbers, its costs can be higher than those of conventional education.

(Perraton 1982: 61)

For someone who, like the present author, is mainly concerned with the applications that make full use of the potential of distance education as referred to on pp. 161–5, it is possible to go a little further than Perraton. There can be no doubt that distance education, as applied to large student bodies, is characterized by very favourable cost-benefit relations provided that the distance-teaching element consistently predominates. A number of case studies illustrating this are reported on in an earlier book of mine (Holmberg 1985a) and also in the first edition of the present work. It is primarily the arrangements for face-to-face sessions, such as study centres, residential schools, and classes of various kinds, that modify or negate the validity of this statement, i.e. non-distance supplements. It is true that use of sophisticated and costly media and technology also in some cases detracts from the favourable cost-benefit relations, but this does not change the overall picture of distance education as economical.

A CONCLUDING GENERAL APPRAISAL OF DISTANCE EDUCATION

This chapter on distance-education evaluation has been concerned both with how this type of education is evaluated and with the outcomes of evaluating studies. The latter concern would seem to warrant a concluding remark.

There are different types of distance education which make more or less consistent use of the potential of its special characteristics. Within these types there are also more or less successful practices. Like conventional types of education, distance education cannot be described as intrinsically either effective or ineffective, good or bad. It opens up a number of possibilities, however, and it does so in ways that are different from those of conventional education.

The preceding chapters will have shown that distance education is applicable, and has been successfully and economically applied, to many educational tasks and many different target groups. It mainly serves adult students, the secondary and tertiary stages of formal education, vocational and professional basic and further training, and self-actualizing study with or without purposes connected with academic credit or labour-market interest. Methods have been developed that strengthen its personal relevance to individual students and make it effective from the aspects of goal attainment, intellectual and emotional development, and involvement in serious study, as well as from those of energy, time, and financial resources invested. Distance education can safely be described as a useful and flexible kind of education with special potential for student autonomy.

Distance education has been shown to have generated theoretical considerations, which are concerned with its particular character, and hypothetico-deductive approaches to its educational effectiveness. After early attempts at distance education in the eighteenth and nineteenth centuries, followed by about a hundred years of experiences with organized distance education and a great number of scholarly studies into its theory and practice, we are entitled to describe it as an established mode of education in its own right.

11

THE ACADEMIC DISCIPLINE OF DISTANCE EDUCATION

Considering the integration of educational, psychological, organizational-administrative, sociological, philosophical, technological and economical aspects of the study of distance education it is possible to regard it as an interdisciplinary field of study. It can also be assigned a place under the comprehensive discipline of education. It has been shown to be a separate entity, however, which can only to a limited extent be described, understood, and explained in terms of conventional school or university education, classroom teaching, or group activities (see pp. 164–5), and so it makes sense to describe it as a special discipline with its roots in education, which, in its turn, is rooted in philosophy. The criteria for a university discipline are usually that there is a body of research encompassing and defining it and that it is taught as a university subject. 'Obviously, an academic discipline is an area of academic interest, and one that poses sufficient problems to stimulate research, and one that leads to the publication of journals in the subject area' (Sparkes 1983: 179).

RESEARCH ON DISTANCE EDUCATION

That there is a body of published research on distance education will have been made manifestly clear by this book with both its constant references to research and its bibliography. However, the study of distance education is evidently benefiting from knowledge and theory developed in disciplines that were established earlier. Most of the research done on distance education could be ascribed to these, for example to general education, pedagogics and andragogics, philosophy, psychology, sociology, history, and economics.