PAPER 10

EVALUATION IN DISTANCE LEARNING

Alan Woodley and Adrian Kirkwood
The Open University, Britain

CONTENTS

INTRODUCTION

Evaluation in the sense of "the systematic investigation of the worth or merit of some object" has several meanings within education. Its most common purpose is to grade and classify students, to sort out the successes from the failures. However, here we are concerned with "Programme Evaluation", where the focus is on how well a particular educational programme, curriculum or teaching method works, how it might be improved and how it compares with alternatives.

This type of evaluation has and continues to be done by formal agencies, such as the Inspectorate in Britain. However, during the late 1950's and early 1960's the pace of innovation increased with the introduction of new curricula (e.g. "New Maths" and "New Science"), new teaching methods, and new programmes aimed at disadvantaged groups. The cost and scale of these innovations led to a demand for more evaluation and this has produced an increase in professional evaluators. Today educational evaluation is a legitimate field of intellectual endeavour and has its own associations, journals, conferences and theorists.

As in any other field, there are controversies and disputes as to the most appropriate aims, methods, theories, etc. To describe and debate the merits of each of the various paradigms in any detail would require far too much space. Instead we attempt to characterise two very different approaches which, while they are "ideal types", serve to indicate the major divisions within the field.

On the one hand there is the "traditional", "positivist-empiricist" or "agricultural-botany" approach to evaluation. This represents the attempt to apply the rules and procedures of the physical sciences to the evaluation area. Parlett (1970) described how this would work in practice:

> "A proper evaluation requires that the experiment or project is structured, from the start, with the design requirements of the evaluator in mind. Only in this way can a thorough-going study be launched. Carefully matched samples must be picked, controls established, and the variables selected for which data will be gathered. The evaluator is likely to require the scheme's organisers to define their objectives, either from scratch or using objectives taken from pre-existing lists like Bloom's.
>
> The next stage is to select or construct appropriate objective tests. Pre-testing, item analysis, checks on reliability, and also, perhaps, the piloting of questionnaires and attitude inventories can then all follow. Next comes the administering of the tests. This completed, the evaluator disappears from the field to analyse his data, only to reappear some months or years later bearing his report."

In its ultimate form:

> "large, finely balanced samples are formed into experimental and control groups; they are tested before the pedagogical "treatment" is applied and tested again afterwards; "before and after" and "between samples" comparisons can then be drawn."

This represents the orthodox position in evaluation, the ideal to which most evaluators strive. However, in recent years there has been a counter movement which has challenged this orthodoxy. Collected writings such as "Beyond the numbers game: a reader in educational evaluation" (Hamilton, *et al*, 1977) "Human inquiry: a sourcebook of new paradigm research" (Reason and Rowan, 1981) and "New directions in educational evaluation" (House, 1986) have documented the weaknesses of the traditional approach and have suggested alternatives.

It is difficult to summarise this new movement because its members are more united by their rejection of the old orthodoxy than by their adherence to a new orthodoxy. Basically the application of physical science methods to complex human phenomena is seen to be invalid, it produces "dead knowledge" which is seldom used and it is immoral in that it frequently involves manipulation of the subjects under study (Reason and Rowan list eighteen objections). The way forward is seen in terms of a more naturalistic methodology and a holistic perspective on human affairs. One has also to acknowledge the complexities of the evaluation process - the fact that there might be several potential audiences with different information needs, the impact of the evaluators own values, the need to adapt one's methodology when, for example, the programme produces unforeseen side-effects.

With the increase in publicly-funded distance education over the last fifteen years, often by the creation of new institutions, has come a corresponding increase in the amount of educational evaluation in this area. As with most innovations, evaluation is required to justify continued funding but also it is necessary because of the invisibility of distance students.

Innovations in distance education frequently involve new teaching methods, new media, new student groups, new curricula and new institutions. In these contexts "evaluation" tends to cover a broad range of activities and employs the full armoury of social science methods of data collection and analysis. Distance education evaluators are aware of the traditional versus new paradigm debate (see Morgan, 1984) and attempts have been made to overcome the problems of conducting essentially qualitative research with distance students. However, distance education evaluation remains eclectic as to its methods and this seems to be inevitable when we consider what such evaluation consists of.

SYSTEM EVALUATION

a) Basic measures of activity

Any system evaluation must begin with certain basic measures. How many courses have been produced? How many students are there? How many applicants had to be turned away? This data is drawn from administrative records and is presented, proudly or otherwise, in annual reports.

b) Measures of efficiency

Allied to measures of activity come those of efficiency. How many students successfully complete the courses? What workload do they attempt? What is the throughput of students? Again this data comes from administrative records (McIntosh, Woodley and Morrison, 1980).

The evaluation can move beyond the descriptive to the examination of patterns and causes. For example, postal surveys can be carried out asking students why they dropped out of courses. However, these tend to produce low response rates and answers of dubious validity. A more systematic approach involves a combination of methods including detailed statistical analysis, an understanding of the subjective process of "dropping-out" and an awareness of the different policy options and their likely impact (Woodley and Parlett, 1983).

Other efficiency measures centre on the question of cost-effectiveness. On the face of it distance education is a cheap teaching method, but just how cheap or whether it is cheap at all is still in dispute (Wagner,1977 and Mace, 1978). Disagreements among the economists centre upon difficulties in making comparisons with conventional institutions. How does one allow for the fact that distance students remain economically active while studying? Is it important that mature students have less years of economic activity in which to employ their new knowledge?

c) Outcomes

Measures of adequate learning are usually considered to be covered by formal exams and assessment. However, as we note later in this report, there have been some attempts to measure the development of distance students as learners using study inventories and in-depth interviews over a period of several years. In one case Open University economics students were compared with conventional students by administering a standardised test of economic knowledge (Lumsden and Scott, 1980).

On many distance courses there are no formal exams and follow up surveys have to be carried out to see whether there have

been appropriate changes in behaviour or attitudes. Such a study has been undertaken at the Open University to see whether people carried out the energy-saving measures in their home that were specifically recommended. Elsewhere the focus has been on health campaigns (Hall, 1978) and changes in agricultural methods (Coleman and Opuku, 1968).

Mail surveys of Open University graduates have been carried out to measure the personal, occupational and educational outcomes of their studies (Swift, 1982). Such studies allow one to measure the subjective benefits experienced by individuals and the extent to which the qualification is recognised by other educational institutions and professional bodies. Methodological problems include those of tracing graduates after a number of years, finding appropriate comparison groups and determining which changes were a direct result of their studies.

The recognition of the qualification can also be approached from the other direction. A survey of employers was carried out to establish the standing of the Open University degree and the acceptability of its graduates (McIntosh and Rigg, 1979).

Distance teaching can have other outputs besides the more obvious ones. These include the use of teaching materials in other institutions (Moss, 1979) (Glaister and Carr, 1986); the passing on of materials to other learners (Stainton Rogers, 1984); and effects on the educational motivation of the children of distance learners (Fenster, 1982).

d) Programme aims

Some distance teaching programmes specify their basic aims in terms of what and whom they intend to teach. One of the evaluation needs will be establish the extent to which these aims have been met. The easiest way to measure who the students are is to include detailed questions on the application form and store this data as part of the student record. However, most institutions to not want to make their application forms too daunting or imply that certain information such as ethnic background may influence the chances of gaining a place. It is, therefore, more usual to collect only basic demographic information at this stage and to supplement it by means of special surveys at regular intervals (McIntosh, *et al*, 1977).

When one of the programme's aims is "open-ness" or an increase in social equity then particular attention must be focussed on formerly disadvantaged groups, be they women, ethnic minorities, working class people or whatever. As well as measuring their representation among the student body and looking at the progress they make on the courses, one can also look at the barriers to greater participation. At the Open

University this has taken the form of carrying out surveys of the general public to determine the levels of awareness and knowledge of the institution (Swift, 1980), surveys of people who sent for details but decide not to apply (Woodley and McIntosh 1977), and surveys of applicants who decline the offer of a place (Woodley, 1983).

e) Policy Evaluation

Formative evaluation in the policy area can take the form of market research. Surveys of prospective students and employers can be carried out to measure the likely demand for possible new courses. Surveys of current students can also be used to sound out opinion on various policy options. A recent study tested the reactions of Open University students to the possibility of reduced tutorial provision on higher level courses (Thorpe, *et al*, 1986).

Policy Evaluation can take the form of monitoring. The Open University carries out regular surveys to monitor the financial impact of study on its students, thus gauging the effects of fee increases, changes in local authority assistance, the effects of its own financial assistance schemes, etc (Blacklock, 1982). Other survey data on the ownership of televisions, cassette recorders, home computers, etc, can assist course planning (Grundin, 1983).

Research can evaluate the impact of policy changes. A recent study at the Open University showed that the new policy of "deregistering" non-progressing students actually affected a large number of its own graduates (Heron, *et al*, 1986).

Finally, evaluation studies make take the form of experiments or pilot schemes. A policy option such as the admission of school-leavers to the Open University can be carried out on a trial basis and the outcome evaluated before making a final decision (Woodley and McIntosh, 1980).

f) Organisational evaluation

A distance education programme or institution can be evaluated in terms of its internal organisation and procedures. Just as in any other setting this can involve scrutinising the financial management and general "organisation and methods". More specifically it can involve tasks such as the monitoring of tutors marking patterns and the turn-around time for assignments. Evaluations have also been conducted into the course team approach to distance course writing.

LESSONS FROM SYSTEM EVALUATION IN DISTANCE LEARNING

What sorts of things have been learned from these types of system evaluation? Here are a few examples:

* There has been a very marked drop in the proportion of teachers on Open University undergraduate courses.

* Many of those who have heard of the Open University believe incorrectly that qualifications are required for entry.

* Increased drop-out rates on Open University higher level courses were shown to be due to changes both in the student population and the nature of courses being offered.

* By 1982 a third of Open University students owned or had access to a microcomputer.

The list of useful, and often surprising, findings could be extended almost indefinitely. However, the findings almost always depend very heavily upon the nature of the system under study and of the society in which it occurs. We are very far indeed from being able to construct "iron laws of distance teaching systems". Take drop-out rates for example. By knowing something about the course characteristics and the student population we can estimate the likely drop-out rate for a new Open University course. However, it would clearly be impossible to use such a predictive model as it stands for, say, a distance learning material pack on vegetable growing in a developing country. Rather than the detailed results of particular evaluations, it is the general approaches and techniques that have been developed that should be of interest to evaluators elsewhere.

Good administrative records are a pre-requisite for much system evaluation. By "good" we mean accurate, comprehensive and accessible to the evaluator. In large scale distance learning projects this generally involves the careful design of self-completion application or registration forms for students, creation of computerised records with each student's demographic characteristics and academic record, and statistical packages to enable the evaluator to manipulate and extract the data in appropriate ways. Special surveys are frequently required to address particular problems. Techniques such as mail surveys, telephone interviews, face-to-face interviews and group discussions have all been used in distance learning systems and each method has its strengths and weaknesses. The choice of method (or methods) will depend upon factors such as environment (e.g. extend of telephone ownership), cost and type of information you are trying to obtain. Mail surveys are commonly used because of the low costs and ease of administration but low response rates often cast doubts on their validity.

The impact of system evaluation on decision making is difficult to measure. Often it is naive to even look for a direct relationship between research recommendations and policy decisions because the research findings are only one input into the decision making process. However, it is certain that an evaluation report will have no impact if it arrives after the decision has been made,

weighs several pounds and addresses the wrong questions. Experience has shown that the best evaluation is done in collaboration with those making the decisions, the form and nature of the evaluation being negotiated with all concerned so that it is appropriate to their needs. (This may involve the evaluator persuading them that they are asking the wrong questions!) Similarly the results need to be disseminated in ways that will produce some impact. This frequently means short summaries of results presented in a straight-forward manner rather than dense articles in learned journals.

The great majority of evaluation in distance learning has been carried out internally by members of the providing institution. This runs counter to the concept of "the neutral judge" but (provided that the staff have the appropriate expertise to carry out the studies) can bring with it the advantages of relevance and timeliness. Above all, a programme of institutional self-evaluation is an indicator of a communal wish for self-improvement. The process of carrying out the evaluation can be as thought-provoking as the results themselves.

COURSE EVALUATION

If a primary aim of evaluating instruction is to improve the quality and effectiveness of the teaching and learning involved, there are a variety of procedures that can be adopted to inform the activities of those involved in providing and supporting instruction. In distance education the separation between teachers and students normally involves instruction or learning experiences being mediated through written or recorded materials. These 'published' materials will normally remain in substantially the same form for several (or many) cohorts of learners, although some modifications and amendments can usually be arranged. In turn, this separation entails the absence of direct verbal and non-verbal feedback from students that is so important in all teaching. The evaluation of any distance education teaching materials or course of instruction (whether or not a formal 'course' with clearly defined expectations of outcomes) may seek to provide information that can be used during the process of developing or preparing materials or learning experiences - *formative* evaluation procedures, or information about how well the 'finished' instruction has worked in normal use - *summative* evaluation procedures (Scriven, 1967). In practice, it is often impossible to draw such a clear distinction, but it does provide a useful way of considering methods of evaluation.

Formative Evaluation

a) Critical Commenting

Whether or not the teaching materials are being prepared by a course team, it should be possible to arrange for peer review

of draft materials. At an informal level this can simply involve one or more colleagues reading, listening to or looking at draft materials and providing comments in terms of the suitability of content (e.g. is it factually correct and supported by examples or evidence, up-to-date, at an appropriate level for the intended students, etc.) and the style of presentation (e.g. does it adopt an engaging style, employ appropriate sequencing and strategies, clearly present information and/or arguments, etc.). The reactions of colleagues can be augmented by adopting the more formal procedure of inviting one or more experts in the field to act as Assessors to comment on the draft materials.

Where arrangements have been made for systematic critical commenting, with teachers or writers reviewing the materials prepared by all the others working on the same course or programme, there is the potential to improve not only individual teaching materials, but also the overall course of instruction. However, it is possible that a 'course team' approach can introduce new problems when writers come to revise their draft materials, as discussed in further detail below.

b) Developmental Testing

As the name implies, Developmental Testing takes place during the preparation phase and involves trying out draft teaching materials with students. The feedback obtained is used to guide and inform writers' revisions to the materials before they are committed to print or tape (Nathenson & Henderson, 1980). Such testing may range from a fairly informal student try-out of a single piece of teaching, to an elaborate procedure for testing draft materials for a whole course of instruction.

Students study the draft materials in the usual manner and may be asked to undertake any other requirements, for example submitting assignments, attending tutorial sessions, etc. and possibly sitting an examination upon completion of the course. Their comments on and reactions to the teaching can be collected by means of questionnaires and/or interviews and observations and are fed into the process of revising the course materials for 'final' presentation.

Experience of developmental testing at the Open University (Henderson, *et al*, 1983) indicates the strength of the procedure for the revision of materials within the overall structure of the course and that these can be of benefit to both course writers and students. It is, however, not particularly suitable for enabling major structural changes to be made to the course. In an attempt to allow for greater flexibility, a number of other procedures have been tried that are part formative and part summative (Henderson, *et al*, op.cit.). These involve collecting feedback from students and tutors on a short-term 'published' version of course materials,

to inform revisions to be made for subsequent presentations.

c) Revising Draft Materials in the light of Formative Evaluation

A number of procedures have been discussed for the collection of evaluative comments on the quality and effectiveness of teaching materials during their preparation or development. It should be recognised, however, that evaluation is rarely uncontroversial and that course writers may experience problems when trying to take account of evaluative information and data in the revision of their teaching. For example, it may be difficult to decide which source of information should be given greatest weight (especially if conflicting views have been expressed), comments on later parts of a course or programme may not be as extensive as those on earlier parts, etc. The actual strategies adopted by writers for the revision of draft materials will, in part, be determined by emotional factors, as the process is not simply a technical task (Riley, 1984).

Summative Evaluation

No matter how well distance teaching materials have been prepared, it is not until students study and try to learn from them that it is possible to determine how good an educational experience results. It is all too easy to consider the 'product' of course development in distance education to be the materials that are delivered to students by one means or another, whereas it is really the interaction of learners with those materials and other resources, possibly including tutors and fellow students (Thorpe, 1979). Summative evaluation procedures are intended to provide information about a course or materials in use.

a) Feedback from Tutors

In many situations, those who have been responsible for preparing distance education materials also provide tutorial support to their students. In these circumstances there exists a direct channel for students' reactions and comments to be made known to the course writer(s) and for any problems with the instructional arrangements to be discovered. However, many distance education tutors may not have been directly involved in the development of teaching materials for the course(s) they tutor. Where such tutors support a course by marking assignments, running tutorial sessions (face-to-face or by means of telecommunications), etc. mechanisms can be implemented to collect, on a systematic basis, evaluative comments from them on a range of issues, e.g. their reactions to the teaching materials, accounts of problems their students have encountered in their studies and assignments, etc. (Ryan, 1982). The experience of tutors in making the course work can provide particularly useful information for subsequent modifications to or adaptations of the teaching materials and instructional arrangements.

b) Feedback from Students

The collection of feedback from students undertaking a course can provide information that may be of value to (i) students of the course in that or subsequent presentations, (ii) students of a replacement course and (iii) students taking other courses prepared by the same writer(s). It may be possible to implement some revisions during the presentation of a course as a result of students' comments, for example by providing a supplement to update information or to clarify a problem area. More frequently, the student feedback from one presentation of a course helps to determine revisions for subsequent presentations. After a course or programme of instruction has been presented in substantially the same form to many cohorts of students, feedback may be collected to inform decisions about remaking or replacing the course. Information gained by course writers about the success (or otherwise) of approaches and strategies employed in their distance teaching may prove to be of great value when they prepare further courses.

Students may be invited to comment on their experiences of studying a particular course by means of questionnaires and/or interviews. Feedback may be sought on each unit of instruction or part of a course. Questionnaires are probably most widely used for collecting feedback from students and can include several different types of question. At the simplest level a questionnaire may just ask students to indicate what they thought were the 'worst' and the 'best ' aspects of the teaching and to specify any particular problems they encountered. At the other end of the scale, questions may be posed on all aspects of the course, possibly using multiple-choice response categories that can be coded for computer analysis. Questions may seek to obtain information about the following main areas of interest:

* Extent of Utilisation. In distance education it is important to know exactly what is being studied and what use is being made of the teaching materials and other resources. Students may be asked to indicate which parts of the course or programme they have studied, which components they have used, how much time they have spent on their studies, etc. They may also be asked to report on any problems they have encountered in obtaining the course materials or in gaining access to resources.

When interpreting the responses to these questions, it must be remembered that students' answers are subjective due to their individual 'orientation' to their studies (Taylor, *et al*, 1981a). They may all have different reasons for studying a particular course or set of materials, have varying amounts of relevant prior knowledge and experience in the area of study and be undertaking their studies in a variety of financial, geographical, domestic and, possibly, occupational circumstances. Questionnaire responses should not be

considered without reference to possible differences in such student characteristics. For example, a student may report having spent considerably less than the notional amount of study time on a particular unit of instruction, due to a high level of existing knowledge about that aspect of the course. Another student of the same unit may report having spent considerably longer than the notional study time, because of great personal interest in the topics and a willingness to explore them at great length.

* Overall View of the Teaching. Students may be asked to rate the teaching of a particular unit of instruction in terms of its interest, perceived relevance or usefulness, level of difficulty, etc. They may also be asked to rate individual components of a course (e.g. teaching text, audio-tape, etc) in terms of their relative usefulness.

Again, when it comes to interpreting the responses, account must be taken of the subjective nature of students' answers. The 'relevance', 'interest', 'difficulty', etc. of any unit of instruction will mean something different to each person studying it and an individual student's rating will largely be determined by his or her orientation. For example, a topic that is considered by one student to be highly relevant (perhaps for vocational reasons), may be rated at the opposite extreme by another student with different orientation.

The analysis of responses on rating scales requires some caution. Very often 4-, 5- or 6-point rating scales are used (e.g. 1 = Very Interesting, 2 = Fairly Interesting, 3 = Of Little Interest, 4 = Of No Interest). It is not valid to average the responses from a group of students to the same question (i.e. reporting an 'average rating' of 2.1 for a particular unit of instruction), because the difference between points 1 and 2 on the scale is not the same as that between points 2 and 3, etc. (They are usually 'Ordinal Scales'). Responses should be reported in a form such as "x% of students found the unit 'Very Interesting' or 'Interesting'".

* General Style of Presentation. Course writers may be keen to receive students' comments on the style of presentation, both in terms of layout, design, etc. and the coherence and clarity of the teaching. Perhaps more importantly, students could be asked to comment on the extent to which the teaching style or strategy had enabled them to become actively engaged in learning from the materials.

Here again, it is important to take account of students' orientations to their studies when interpreting responses. For example, students with an 'instrumental' or 'extrinsic' orientation may prefer a didactic style of presentation, while other students may not.

* Specific Content Issues. It is important to know how well the teaching has achieved its aims and objectives. To this end, information about students' problems with key concepts, ideas and relationships, etc. can be of great value to course writers when it is time for revisions to be made.

Students may be asked to indicate on a multiple-point scale how well they felt particular key points had been taught, or may simply be asked to report any specific problems they had encountered. Simply to know that difficulties exist may not be enough. It may be possible for students to suggest whether problems arise from the inherent difficulty of certain ideas, etc., or from the course writer(s) explication in the learning materials. Any indication of the nature of specific difficulties can be of great value when revisions are made to the teaching.

At the Open University extensive use has been made of questionnaires for the collection of feedback from students during course presentation (Lockwood, 1983; Ballard, 1983). The questionnaires used for courses in their initial presentation contain some items that are course specific and others that allow comparisons to be made between courses for certain aspects of the information, e.g. workload, use made of course components, etc. Depending upon the number of students taking a particular course, questionnaires may be sent to all or to just a sample of students. If a series of questionnaires is to be used for a course or programme of study (for example, one questionnaire for each part of a course), problems of diminishing response rates may be encountered, with fewer students providing feedback on later parts of the course. In such circumstances, it has been found that the decline in responses can be reduced by using 'panels' of respondents, i.e. sending the questionnaires for each part of the course to a different sub-set of students, so that each individual receives only one form to be completed. An increase in responses to the questionnaires on later parts of the course will normally compensate for any reduction in the total number of respondents for each part.

Student Interviews. Questionnaires can be criticised as a method for collecting student feedback on the grounds that they elicit responses to questions that are of interest to the course writer(s) or evaluator(s) and not necessarily on issues that are considered important by students. The constraints of questionnaire design can in part be overcome by including space for students to comment in an open-ended way about their reactions to and experience of studying a course. Interviews with students, conducted either face-to-face or by telephone, can provide a richness of response difficult to achieve by use of questionnaires alone. Depending on the type of information sought, a highly structured interview schedule may be utilised or just a short list of key areas to be covered in a relatively free-flowing discussion.

Although interviews may be more time-consuming to conduct and analyse than questionnaires and may elicit responses from a smaller number of students, in-depth interviews are particularly useful for gaining a view of courses from the students' perspective (Taylor, et al , 1983).

c) Revising Courses in the light of Summative Evaluation

By a variety of means, a wide range of evaluative comments and information can be collected from students, tutors and, possibly, other interested parties. Each of the various methods described above has its strengths and weaknesses and for any particular course one or more of the procedures may be considered to be most appropriate. For example, a questionnaire survey of students may provide an indication of significant problem areas and lead to some follow-up interviews being undertaken with tutors and students in order to determine the nature of the problems and to elicit suggestions about how they can be overcome. However, before any procedures are initiated, it is important to establish the extent to which it is possible to make changes to the teaching in the light of any evaluation. To a large extent, this means considering what information can be acted upon (and is therefore worth collecting), and when it should be collected.

It may be very difficult to amend teaching materials in their 'published' form due to the physical format (e.g. as printed booklets, audio- or video-tape, etc.) or a lack of resources (human or otherwise) to implement changes. Even if evaluation indicated a major problem with the teaching materials, it would be difficult to do more than change the sequence of presentation, omit a part or make it optional, provide additional study guidance, etc. However, in such cases there is little point in conducting a thorough evaluation of all aspects of the course or programme of study - it would be better to concentrate on aspects that can be acted upon for future presentations, for example tutorial and assessment arrangements.

Timing the collection of evaluative information is similarly important. It is important to collect information and data at a time when it can be of most use. Very often, evaluation studies are undertaken during the initial presentation of a course in order to inform changes for future presentations. However, there may be reasons to believe that the first cohort of students is not typical of the students who are likely to study the course, for example they may be more highly motivated to take that particular course. In those circumstances, a second or subsequent cohort may be considered more appropriate. If changes are to be implemented during the current or in a subsequent presentation of a course, the evaluative information must be available in time to be acted upon. If amendments are to be made to printed or recorded materials, production time is necessary - if tutorial or assessment arrangements are to be

revised, adequate notice is required. When deciding how and when to evaluate distance education courses, it is necessary to consider what changes are possible given the time and resources available.

CROSS-SECTIONAL STUDIES

Within an overall programme of distance education it may be considered desirable to undertake an evaluation study of a particular innovation or component used in a number of courses. Such studies are likely to be concerned with drawing out generalisations from the use of a particular aspect of the teaching, or with establishing the effectiveness of a particular strategy or teaching medium. The research may be used in the formulation or revision of policies for the institution and, as such, could contribute even more to the development of future courses than to those evaluated in the study. Any such study may utilise one or more of the methods already described for the collection of information, i.e. questionnaires and interviews. Studies of how students use certain resources may also involve direct observation of individuals or groups in the process of studying (at residential schools, tutorial sessions or in their own homes), or shortly afterwards by means of stimulated recall (McConnell, 1985).

The role of particular course components has been of concern to the Open University and other institutions using a multi-media approach to distance education. Research into the use of audio-visual media (Grundin, 1985), tutorials (Kelly,1981; Kelly & Swift, 1983), and computer assisted learning (Scanlon, *et al*, 1982) has involved collecting information from students on their access to and use of particular components and resources, as well as eliciting their views on the contribution made to the teaching and the overall effectiveness of courses.

The responses given by students to questions about access to resources or facilities for their studies can be highly subjective. For example, a student living in an urban area may consider that the journey across the city to a study centre makes access 'poor', while a rural student may regard access as 'good', even though a much longer journey is necessary (Evans, 1986). In cross-sectional studies, however, it is often important to consider situations from the students' perspective.

The introduction of an innovatory teaching strategy may also give rise to evaluative research across a range of courses, for example the use of project work to encourage greater independence in distance education (Henry, 1979).

DEVELOPMENTAL STUDIES

Evaluating the effectiveness of teaching and learning within a

distance education programme or institution may also involve undertaking research into the understandings developed by students as a result of their studies. Qualitative changes in the understanding of key concepts and relationships formed the focus of a study of Open University social science students (Taylor, *et al*, 1981b & 1981c), based upon a method for evaluating the content of students' learning developed in Sweden by Dahlgren (1978).

The study was concerned with finding out not *how much* students know, but *what* they understand about particular concepts and principles. A group of university students taking an introductory social science course were individually asked a set of questions about key concepts and principles taught in the course. They were interviewed before commencing their studies and again after completion of the course. The questions were asked in everyday language, in order to establish what the students understood of these ideas, in their own terms. Analysis of their answers revealed that for each of the key ideas a number of qualitatively different conceptions existed and that it was possible to categorise them. Students would be expected to hold conceptions that were qualitatively more sophisticated after completing the course.

In the course assessment, all the students interviewed both before and after taking the course gained enough marks to pass. However, the conceptions held by students (as revealed in the interviews) show a different picture. For each individual concept or principle, some students demonstated an improvement in the sophistication of their conception after taking the course, some showed no marked difference and a few held less sophisticated conceptions at the end! Hardly very encouraging outcomes. The findings of this study informed the writers of the replacement course, not only by identifying problem areas in the teaching, but also by illustrating the different levels of pre-course understandings that students were likely to have.

This approach to establishing the success (or failure) of individual teaching units in enabling students to develop more sophisticated understandings can provide evaluative information that is of great value when revising or replacing teaching materials. However, such an approach can only succeed with collaboration and co-operation between course writers and evaluators in determining the key ideas that form the focus of the study, in agreeing the categorisation of conceptions and in acting upon the evaluative information that is available.

REFERENCES

Ballard, A. (1983) Obtaining Feedback on Continuing Education Short Courses *Teaching at a Distance Institutional Research Review* 2, pps 113-132.

Blacklock, S. (1982) What the OU Cost the Undergraduate Student in 1981 (mimeo). Survey Research Department Paper No. 229, The Open University.

Coleman W.F and Opuku A.A. (1968) An African experiment in forums for rural development, Ghana 1964-5. Reports and papers on mass communications no 51 Paris. Unesco.

Dahlgren, L. O. (1978) Qualitative Differences in Conceptions of Basic Principles in Economics: A Contribution to the Discussion of the Validity of Examination Results. *4th International Conference on Higher Education,* University of Lancaster.

Evans, T. (1986) A Report on Interviews with Deakin University Off-Campus Students about the Use of Audio-Cassettes and Video-Cassettes in their Studies (mimeo). Distance Education Unit, Deakin University.

Fenster, E. (1982) *College Attendance by Working Adults and its Effects on the Educational Motivations of their Children* . Detroit: To Educate the People Consortium.

Glaister, B. and Carr, R. (1986) Education Open to All. *Open Learning* Vol 1, No 3, pps 50-52.

Grundin, H. (1983) Audio-visual Media in the Open University (mimeo). IET Papers on Broadcasting No 224, The Open University.

Grundin, H. (1985) Report on the 1984 Audio Visual Media Survey (mimeo). Institute of Educational Technology, The Open University.

Hall, B.L. (1978) Mta ni afya: Tanzania's health campaign Washington DC: Clearing house on Development Communication.

Hamilton, D., Jenkins, D., King, C., MacDonald, B. and Parlett, M. (eds) (1977) *Beyond the Numbers Game: A Reader in Educational Evaluation* . London: Macmillan.

Henderson, E. et al (1983) Developmental Testing for Credit: A Symposium *Teaching at a Distance Institutional Research Review* 2, pps 39-59.

Henry, J. (1978) The Project Report (mimeo). Institute of Educational Technology, The Open University.

Heron, M., Kelly, P. and Marshall, J. (1986) Student Progress and Non-progress (mimeo). Regional Research and Development Papers, No 15, The Open University.

House, E. (ed) (1986) *New Directions in Educational Evaluation* . Lewes: The Falmer Press.

Kelly, P. (1981) An Overview of Student Use and Appreciation of Tuition (mimeo). RTS Research Group, The Open University.

Kelly, P. and Swift, B. (1983) Tuition at Post Foundation Level in the Open University - Student Attitudes Towards Tuition (mimeo). RAS / IET, The Open University.

Lockwood, F. (1983) Evaluation of Courses in Presentation *Teaching at a Distance Institutional Research Review* 2, pps 5-13.

Lumsden, K. and Scott, A. (1980) An Output Comparison of Open University and Conventional University Students (mimeo). Edinburgh: Esmee Fairbairn Research Centre, Heriot-Watt University.

McConnell, D. (1985) Learning from Audiovisual Media: Assessing Students' Thoughts by Stimulated Recall *Journal of Educational Television* Vol. 11, pps 177-187.

McIntosh, N., Calder, J. and Swift, B. (1976) *A Degree of Difference: A Study of the First Year's Intake to the Open University of the United Kingdom.* Guildford: Society for Research into Higher Education.

McIntosh, N. and Rigg, M. (1979) Employers and the Open University (mimeo). Survey Research Department Paper No 162, The Open University.

McIntosh, N., Woodley, A. and Morrison, V. (1980) Student Demand and Progress at the Open University - the first eight years. *Distance Education* Vol 1, pps 37-60.

Mace, J. (1978) Mythology in the Making: Is the Open University Really Cost-effective? *Higher Education* Vol.7.

Morgan, A. (1984) A Report on Qualitative Methodologies in Research in Distance Education. *Distance Education* Vol 5, pps 252-267.

Moss, C.D. (1979) The Influence of Open University Distance Teaching in Higher Education. *Teaching at a Distance* 14, pps 14-18.

Nathenson, M. and Henderson, E. (1980) *Using Student Feedback to Improve Learning Materials* London: Croom Helm.

Parlett, M. (1970) Evaluating Innovations in Teaching, in Butcher, H.J. and Rudd, E. (eds) *Contemporary Problems in Higher Education*. London: McGraw-Hill, pps 145-150.

Reason, P. and Rowan, J. (eds) (1981) *Human Inquiry: A Sourcebook of New Paradigm Research* . Chichester: John Wiley and Sons.

Riley, J. (1984) The Problems of Revising Drafts of Distance Education Materials *British Journal of Educational Technology* Vol. 15, pps 205-226.

Ryan, S. (1982) Developing Tutor Feedback. *Teaching at a Distance* 22, pps 9-14.

Scanlon, E. et al (1982) Computer Assisted Learning. *Teaching at a Distance Institutional Research Review* 1, pps 59-79.

Scriven, M. (1967) The Methodology of Evaluation, in Tyler, R. *et al, Perspectives of Curriculum Evaluation* . Chicago: Rand McNally

Stainton-Rogers, W. (1984) Alternative Uses of Materials: Research Findings. *Teaching at a Distance* 25, pps 58-68.

Swift, B. (1980) Trends in Awareness and Beliefs about the Open University among the General Public (mimeo). Survey Research Department Paper No.185, The Open University.

Swift, B. (1982) What Open University Graduates Have Done (mimeo). Survey Research Department Paper No.230, The Open University.

Taylor, E., Morgan, A. and Gibbs, G. (1981a) The 'Orientation' of Open University Foundation Students to their Studies. *Teaching at a Distance* 20, pps 3-12.

Taylor, E., Gibbs, G. and Morgan, A. (1981b) The Outcomes of Learning from the Social Science Foundation Course: Students' Understandings of Price Control, Power and Oligopoly (mimeo). Study Methods Group Report No. 9, The Open University.

Taylor, E., Gibbs, G. and Morgan, A. (1981c) Students' Understandings of the Concept of Social Class (mimeo). Study Methods Group Report No. 10, The Open University.

Taylor, E., Morgan, A. and Gibbs, G. (1983) Students' Perceptions of Gains from Studying D101. *Teaching at a Distance Institutional Research Review* 2, pps 133-147.

Thorpe, M. (1979) When is a Course not a Course? *Teaching at a Distance* 16, pps 13-18.

Thorpe, M. *et al* (1986) Effective Study in the Open University: The Human Dimension (mimeo). Student Research Centre Report No. 1, IET, The Open University.

Wagner, L. (1977) The Economics of the Open University Revisited. *Higher Education* Vol. 6.

Woodley, A. (1983) Why they Declined the Offer. *Teaching at a Distance* 23, pps 2-7.

Woodley, A. and McIntosh, N. (1977) People Who Decide Not to Apply to the Open University. *Teaching at a Distance* 9, pps 18-26.

Woodley, A. and McIntosh, N. (1980) *The Door Stood Open: An Evaluation of the Open University Younger Students Pilot Scheme* . Lewes: The Falmer Press.

Woodley, A. and Parlett, M. (1983) Student Drop-out. *Teaching at a Distance* 24, pps 2-23.