DOCUMENT RESUME

ED 044 577 AC 008 695

| | |
|---|---|
| AUTHOR | Wood, Freddie |
| TITLE | The Evaluation of Functional Literacy Projects Unesco Workshop. |
| INSTITUTION | London Univ. (England). Inst. of Education. |
| PUB DATE | Aug 69 |
| NOTE | 153p. |
| EDRS PRICE | EDRS Price MF-$0.75 HC-$7.75 |
| DESCRIPTORS | Communication Skills, *Cost Effectiveness, Curriculum Development, *Functional Illiteracy, Instructional Media, International Organizations, *Literacy Education, *Program Evaluation, Research Methodology, Rural Extension, Workshops |
| IDENTIFIERS | UNESCO, *United Nations Educational Scientific and Cultural |

ABSTRACT

The document contains papers, working group and discussion reports, and statements by participants of a UNESCO Workshop on the evaluation of functional literacy projects, held at the University of London Institute of Education, August, 1969. The objective of the workshop was to enable people engaged in the evaluation of functional literacy and similar programmes to exchange their views and experience and to come to conclusions as to realistic and practical techniques of evaluation which can be applied to their countries. Some of the topics of discussion were: functional literacy--definition and evaluation; experience problems in evaluation; the evaluation of functional literacy; a sociologist's view of experiences and problems in evaluation of functional literacy; techniques of social study for evaluation; evaluation of rural extension programs; and the Unesco Manual on the experimental literacy projects. A list of participants and their addresses are included in the appendix. (PT)

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED0 44577

# THE EVALUATION OF FUNCTIONAL LITERACY PROJECTS

## UNESCO WORKSHOP

London, 3 - 22 August 1969

REPORT

by

Mr Freddie Wood

University of London:Institute of Education
Malet Street, London WC1

CONTENTS

1 Summary Report on the Organisation and Activities of the Workshop 1

2 Ideas and Suggestions Put Forward by the Workshop 5

3 Functional Literacy: Definition and Evaluation John Bowers 15

4 Experiences and Problems in Evaluation - Statements by Participants 28

5 The Evaluation of Functional Literacy: an Economist's View Prof. M. Blaug 35

6 Cost Analysis in the Evaluation of Functional Literacy Dr. R. Jolly 43

7 Experiences and Problems in Evaluation - a Sociologist's View Dr. K. Neys 49

8 Techniques of Social Study for Evaluation Dr. R. Wiener 54

9 The Evaluation of Rural Extension Programmes Dr. J. di Franco 69

10 Visits to the Overseas Visual Aids Centre (OVAC) and the Centre for Educational Television Overseas (CETO) and Address by Mr. J. B. Bowers on Action Research in the Production of Media for Functional Literacy 73

11 Visits to the City and Guilds of London Institute, the National Coal Board and Pressed Steel Fisher Ltd. 83

12 Reports of Working Groups and Discussion 94

13 The Unesco Manual on the Evaluation of Experimental Literacy Projects 132

Appendix A - List of Participants and Staff 137

Appendix B - Workshop Programme 147

# CHAPTER 1

## SUMMARY REPORT ON THE ORGANISATION AND ACTIVITIES OF THE WORKSHOP

Following Resolution 1.335(d), adopted by the General Conference of Unesco at its Fifteenth Session (Paris, 1968), a Workshop on the Evaluation of Functional Literacy Projects was held at the University of London Institute of Education from the 3rd to the 21st of August, 1969.

The objective of the Workshop, as stated in the letter of invitation to participants, was "to enable people engaged in the evaluation of functional literacy and similar programmes to exchange their views and experience and to come to conclusions as to realistic and practical techniques of evaluation which can be applied in their own countries." All participants were expected to have a working knowledge of English.

Invitations were sent by Unesco to the governments of Member States in which functional literacy programmes were known to be operating, including those countries participating in the World Experimental Literacy Programme, asking them to nominate participants for the Workshop. In the event, Ecuador, Ethiopia, India, Indonesia, Iran, Nigeria, the Sudan, Tanzania, Thailand, the United Arab Republic and Venezuela put forward participants. A complete list of participants is provided in Appendix A.

The Workshop was financed by Unesco and its direction was entrusted to Mr. J. B. Bowers, Adviser on Adult Literacy at the Department of Education in Tropical Areas, University of London Institute of Education. Fourteen specialists in evaluation and related fields from Unesco and FAO and from Universities and other Organisations in the United Kingdom, whose names are also listed in Appendix A, served as full-time staff or visiting lecturers.

A small planning group of three staff members prepared the Workshop programme which is attached as Appendix B. This programme was intended to be sufficiently flexible to allow adjustments to be made as the Workshop developed and the needs and wishes of participants become clear. However, apart from small changes in the time-table, it was in the event closely adhered to.

On the first day of the Workshop, the participants elected a Steering and Evaluation Committee which was charged with the task of assisting on the organisation of the programme and appraising its achievements. The Committee met regularly throughout the Workshop and provided a very valuable service to the participants and staff.

ERIC

The programme maintained a balance between meetings within the Institute of Education and visits to organisations and institutions in and near London.

The meetings in the Institute were themselves varied in purpose and format. During the first week, statements by participants on their particular problems and experiences in evaluation were alternated with contributions from visiting specialists or staff members. These covered the following subjects:

Functional Literacy - Definition and Evaluation.
by Mr. J. B. Bowers;

The Evaluation of Functional Literacy: an Economist's View.
by Professor M. Blaug;

The Role of Cost Analysis in the Evaluation of Functional Literacy.
by Dr. R. Jolly;

Techniques of Social Study for Evaluation.
by Dr. R. Wiener;

The Evaluation of Rural Extension Programmes.
by Dr. J. di Franco;

Experiences and Problems in the Evaluation of Functional Literacy: a Sociologist's View.
by Dr. K. Neys.

These lectures were followed by discussions of the problems raised.

Early in the second week the Workshop broke into three working groups dealing respectively with:

Policy, Planning and Organisation in the Evaluation of Functional Literacy;

Methods and Techniques used in the Evaluation of Functional Literacy;

Cost/Benefit Analysis in the Evaluation of Functional Literacy.

On the third day after the start of their deliberations, the working groups made interim reports of their progress to the whole Workshop. Their final reports were presented during the last week.

Mr. J. C. Cairns, Director of the Literacy Division in the Department of Out-of-School Education of Unesco visited the Workshop for two days at the beginning of the second week and held an informal but most useful and informative discussion with the staff and participants on problems of evaluation in Unesco's Experimental World Literacy Programme.

Also during the second week, visits were made to:

the Testing and Development Unit of the City and Guilds of London Institute, an organisation which is actively co-operating with industries and industrial training boards in providing training programmes and testing procedures for all levels of technical and vocational training;

the National Coal Board in London which provided a detailed description of training and retraining programmes in the mining industry and their evaluation;

Pressed Steel Fisher Ltd., at Cowley, near Oxford, to learn about evaluation and training procedures in the motor industry;

the Overseas Visual Aids Centre (OVAC), and the Centre for Educational Television Overseas (CETO) to discuss the relationship of action research and evaluation to the preparation of media for functional literacy.

A number of cultural visits were also made to places of interest in and around London and Oxford.

Towards the end of the Workshop a full day was devoted to the detailed study of the Unesco Manual on the Evaluation of Experimental Literacy Projects. This was introduced by a most helpful talk from the author of the Manual, Dr. H. P. Saksena, Head of the Literacy Evaluation Unit in the Unesco Department of Out-of-School Education.

A narrative format has been chosen for the Report on the Workshop, activities being described mainly in the time sequence in which they occurred. It has not, of course, been possible to give a detailed account of all the Workshop's activities. The intention rather is to present a synthesis of the contributions made by individual participants, describing their problems and experiences, a summary of the lectures given by visiting specialists and of the discussions which arose from these contributions. These are supplemented by the reports of the three working groups, accounts of educational visits and a brief record of discussion on the Unesco Manual.(1)

An analysis of the main ideas emerging from the Workshop follows in Chapter 2.

---

(1) Unesco Manual on Adult and Youth Education: Evaluation of Experimental Literacy Projects. Provisional version document ED/WS/135 Paris, 85 pp.

## CHAPTER 2

## IDEAS AND SUGGESTIONS PUT FORWARD BY THE WORKSHOP

### Definitions

The Workshop began by discussing the meaning of a number of current terms. 'Functional literacy' (as distinct from 'simple literacy') was taken to mean a comprehensive programme of education and training for illiterate (and sometimes semi-literate) adults, with a literacy component built in. Agreement was also reached on the meaning of 'work-oriented literacy', the 'selective-intensive strategy' as applied in the Experimental World Literacy Programme, and 'development' and 'modernisation', which were regarded as essential aims of functional literacy. These terms are further discussed in the next chapter. The Workshop accepted the definition of 'evaluation' used in the Unesco Manual: "evaluation aims at measuring, wherever possible in quantitative terms, and according to well-defined criteria, the major direct and indirect effects of a certain activity, taking into account its objectives as established before its inception." A distinction was drawn between evaluation in this sense and 'action-oriented research' or 'action research', which was defined as "systematic study and investigation, the results of which are fed back directly and immediately to the operational staff to help them to improve their day-to-day work."

### Aims of evaluation in functional literacy projects

An important purpose of evaluation is to provide the financiers, planners and administrators, who allocate funds for functional literacy, with objective information on how these funds have been spent. Related to this is the need to shed light on the inter-relationship between literacy and development and to study its significance, in order to determine the place and priority to be given to literacy in development plans; hence also the importance of cost-benefit analysis. Evaluation should also help to answer the question: "what sort of literacy?" Thus it should study the effects of different methods and media. For example importance was attached to the comparative evaluation, in terms of cost-effectiveness, of work-oriented functional literacy in which literacy teaching is integrated into vocational training, as against simple literacy leading on to separate vocational training. The efficiency of different types of organisation should also be analysed.

## Evaluation and Action Research

In contrast to this long-range purpose of evaluation its more immediate purpose is to help the operational staff within each project - to act as their 'controlling device', and even as their 'defence mechanism' against over-ambitious expectations from outside. The first is 'external' the second 'internal' to the programme. The first is particularly relevant to the Experimental World Literacy Programme, the second to the success of the local project. The staff of projects within the World Literacy Programme must therefore respond to two demands: first, to accept international norms for evaluation established by the Unesco Evaluation Panel, in the interests of world-wide comparability, and second, to adapt these to the local situation. They may therefore need to establish both general objectives and local operational objectives. In any case, effective evaluation requires a much more precise definition of objectives - and in measurable terms - by project planners than has so far been evident, and evaluators must assist them in this task.

In fulfilling its immediate purpose of helping project operators, evaluation takes on the role of action research and the relationship of evaluation and action research was a recurring theme of discussion in the Workshop. It is clear that there are certain anomalies between the long-range purpose of measuring the effects of functional literacy on development and the immediate purpose of helping the operational staff to improve their on-going activities. If it was a question of choice, it was agreed that priority should be given to long-term evaluation, at least in the Experimental World Literacy Projects, and the temptation to concentrate entirely upon the less rigorous methods and more immediate benefits of action research should be resisted. There was, however, general support for the 'intermediate' position taken up by the Unesco Manual, namely that the evaluation programme should serve both long-term and short-term, external and internal, purposes. This would involve three concurrent and overlapping types of activity:

- the measurement of the 'direct and indirect effects' of functional literacy;
- special studies to evaluate operational methods and media, which will provide data both of world-wide interest and of local value;
- action research to help planners and operational staff to control and improve the programme.

The suggestion was put forward that action research should be built into the experimental production of media(1) and instructional systems for functional literacy, with an action research worker assigned to the media production team. Action research in this context would include studies designed to relate the content of media to the needs of the target audience and to the educational aims of the programme, as well as 'pre-testing', or 'operational tests', as they are called in the Unesco Manual. This type of action research, intended to feed back data to the media producers - to improve media during production, was distinguished from evaluation of the effectiveness of media once completed, distributed and used in the programme.

## Organisational Structure for Evaluation

A national evaluation unit for an important functional literacy programme should ideally include a full-time director and full-time or part-time professional staff with both international and national experience in the fields of economics, sociology, psychology and statistics. It should engage short-term consultants in specialised areas and research workers to undertake special studies.

The unit should have working relations with Universities, and research Institutes, both inside and outside the country, as well as with various government departments, agencies and organisations whose support and assistance may be valuable.

Each 'sub-project' in the country should have its own small evaluation unit including a competent social scientist and appropriate full-time and part-time field staff.

It was considered important to resolve the apparent conflict between the need for independence and objectivity in the evaluation unit and the need for close collaboration between the unit and the operational staff. With this in mind two possible structures were envisaged.

- The national evaluation unit could be located outside the project and be independent, in which case a working liaison committee, including members of the operational staff and the evaluation unit, should be set up to ensure collaboration.

---

(1) The term 'media' is used here in a general sense to include all kinds of educational 'instruments' from books through audio-visual aids to radio and television programmes.

- The national evaluation unit might be located within the operational staff, with responsibility to the project director, in which case links should be established with an outside organisation or organisations, such as a National Evaluation Committee or an appropriate University department or Institute, which would lay down norms and examine the results of evaluation.

The question was debated whether or not action research should be a function of the evaluation unit. The working group dealing with Policy, Planning and Organisation suggested that if the evaluation unit was a part of the operational staff it should be responsible for action research, but not if it was established as an independent unit. Eventually, in its discussion of the Unesco Manual, the Workshop took the view that the three activities (long-term evaluation, special studies, action research) should be handled by a single 'Research and Evaluation Unit.'

## Policy and Procedures for Evaluation

There was general acceptance of the evaluation procedures proposed in the Unesco Manual, at least as an ideal to be aimed at. These would involve the Research and Evaluation Unit, during the preparatory phase: in recruiting and training its staff; in the collection of background data, (from informants, documentation and observation); in problem surveys (to detect the principal problems of the project areas); in technical feasibility studies and operational tests (to try out methods and media). It was agreed that these activities were intended to provide information to the unit, as well as data for project planning and development. They were thus in the realm of action research rather than evaluation. The unit should then be involved with the project planning staff in defining the objectives of the project and in selecting experimental and control groups, if the use of controls proved locally acceptable.

Once experimental and control communities and work-units (such as factories, mines, farms) have been chosen and the operational objectives of the programme and its various courses defined, the evaluation unit can establish its methodology, select what the Unesco Manual calls 'indicators, elements and minimum measurements for evaluation', and begin its 'community-level' and 'work-unit-level' base-line surveys. The 'participant-level' base-line surveys cannot however, commence until the participants are enrolled in classes and courses in the operational phase of the project.

Thereafter 'interim' and 'terminal' surveys will be carried out at the three levels, to measure, against the initial base-line data, the changes that have taken place. Comparison with control groups, as well as continuing observation through so-called 'process studies', should then make it possible to isolate the effects of the independent variable of functional literacy.

Special studies on educational methods and media, and on organisation and administration could be conducted during the operational period and every endeavour should be made to analyse costs and benefits of the functional literacy programme and of specific methods used in it.

Whilst accepting this general scheme of action research and evaluation the Workshop was under no illusions regarding the difficulties of applying it in the conditions prevailing in most of the projects of the Experimental World Literacy Programme, let alone in other functional literacy projects. There was general gratification that Unesco had abandoned the impossible aim of trying to measure the effects of functional literacy at the national level - on the economy of a whole country* The Workshop recognised the importance of evaluation at the levels of the local community and the work unit, but was more inclined to the view expressed by Professor Blaug that the evaluation team would have a substantial achievement to its credit if it could arrive at conclusive measurements of the impact of such a programme on individual participants, plus some effects on the work unit.

## Methods of Social Research

The Workshop assigned one of the working groups and a considerable proportion of its time to discussion of the methods and techniques of social research which were applicable to evaluation and, more specifically, of how they should be applied.

This study encompassed methods of sampling, selection and training of interviewers, interviewing design and interviewing, observation, sampling, selection and use of control groups, assessment techniques, achievement tests, sociometric techniques, interest scales and motivation analysis. It also dealt with the special problems of comparative evaluation of alternative methods in functional literacy. It emphasised the point that evaluation studies should concentrate upon behavioural change rather than attitudinal change. An evaluator is not so much interested as to whether farmers have changed their attitudes, for example, to education or agricultural innovation, as in whether they have changed their behaviour in these fields as a result of the functional literacy campaign. Quite apart from the difficulty of measuring attitudes, the assumption that a person's behaviour changes to match changes in attitude does not always hold up in practice and does not justify measuring the intermediary rather than the end stage of behavioural change.

---

* It should be stated that Unesco did not set before itself the aim of measuring the effects of functional literacy at the national level or on the economy as a whole.The misunderstanding has perhaps arisen because the Guide recommended collection of some data at national level.The purpose of this was not to establish benchmarks for comparison with similar data at a later date but to get an understanding of the context in which the programme was introduced. Please see in this connection page 134,last paragraph.

## Problems met in the Evaluation of Functional Literacy

Early in the Workshop participants gave accounts of evaluation and action research in projects of functional literacy and related fields. Despite substantial economic and social differences between the countries studied, a wide range of similar problems recurred.

Most crucial perhaps was the shortage of competent research and evaluation staff at all levels - international, national and local, and the consequent problems of recruiting the evaluation unit. Related to this was the shortage of time in the one year preparatory period. In several countries, where vernacular languages were preponderant, language policy was a cause of conflict and language differences handicapped the recruitment and operations of the staff. Governmental attitudes to evaluation were not always helpful, and a particular difficulty arose in achieving co-ordination between departments concerned with development and those concerned with literacy.

The process of data gathering was handicapped by a variety of factors - the absence of documentary material and reliable census reports, the physical problems of communication in rural areas, the resistance of rural people to providing information on their economic status, even the resistance of factory managers to releasing data on production or industrial processes which might benefit competitors.

Finally there were the problems of defining measurable objectives and suitable indicators and, more difficult still, of establishing controls, where this meant withholding the benefits of the project from certain groups and communities, whose members must nevertheless submit to interviewing and observation.

Other problems were brought up by visiting lecturers. The economists in particular stressed the very real difficulties of measuring the impact of functional literacy on production. At the level of the work-unit - factory, mine or farm - this was an extremely complex exercise and serious doubts were expressed on the possibility of isolating the effects of functional literacy from those of other factors, such as improved seeds, fertilisers, land reform, communications, changes in the market for produce and variations of climate in the rural sector and machinery, management and materials in an industrial setting. It was significant that in the comparatively sophisticated situation of British industry it was generally considered impossible, and perhaps even undesirable, to measure the effects of training as an independent variable - to evaluate its impact in quantitative terms - on production in the work-unit.

Even at the participant level the isolation of the effects of training on production was not an easy exercise. One possibility was to concentrate on the individual's earnings, but this would only be valid if the labour market worked perfectly competitively. Another approach was through job-analysis, but this would require that the skill-content of every job be analysed, together with the means of acquiring the necessary skills and whether literacy was required for this purpose. The aim envisaged by the Unesco Manual of measuring the contribution of the individual worker to 'the production or other purposes of an enterprise' was regarded as unattainable.

The problems of cost-benefit analysis were also discussed, amongst them the difficulties of quantifying 'opportunity costs' and 'implicit benefits'. The possibility of arriving at a cost-benefit ratio for functional literacy, by encompassing all benefits in a single number and contrasting this with all costs, again in a single number, was considered illusory and a warning was given against drawing simple conclusions from the study of cost-effectiveness in isolation. These difficulties and dangers were not, however, regarded as diminishing the importance, or precluding the possibility, of cost-benefit or cost-effectiveness studies, to which the Workshop gave careful attention.

## Summary of Suggestions emanating from the Workshop

Although the problems of evaluating functional literacy were recognised to be many and serious there was a real determination in the Workshop that they must be resolutely faced. The success of the Experimental World Literacy Programme depended upon it, since there could be no value in experimentation without evaluation.

Although the Workshop was not called upon to give advice or make recommendations, it may be useful to summarise at this point some of the more interesting ideas and suggestions which will be recorded in greater detail in the following chapters of this report.

As a palliative to the shortage of evaluation staff and a contribution to their efficiency it was suggested that Unesco might:

- approach Universities to provide integrated courses suitable for the training of project evaluators;
- bring evaluators together in high-level conferences;
- provide systematic training for evaluation teams in each country;
- give more clear-cut instructions on methodology to be applied;
- list the special studies and experiments it wishes to see undertaken in projects;
- issue a bulletin on functional literacy and its evaluation.

Universities should be encouraged not only to provide training but also to take a more active part in research and evaluation in functional literacy projects.

More attention should be paid in evaluation programmes, and in training for them, to data processing and analysis.

National evaluation units should provide orientation to operational staff in each project in order to give them an understanding of the value, aims and requirements of evaluation, as well as training in evaluation procedures which they will have to apply. In particular, teachers and instructors should be trained in building testing techniques into training programmes.

The 'preparatory period', for a functional literacy project, originally one year, should be flexible enough to permit the evaluation unit to establish its design and methodology, select its experimental and control groups, and to start its base-line surveys.

The various government departments and other agencies whose collaboration would be necessary to the evaluation, as well as to the operation of the projects, should commit themselves to full co-operation, if possible by signing the initial agreement with Unesco.

The definition of general and operational objectives, in behavioural and measurable terms, for each project, sub-project - each training course or extension programme - was regarded as a prerequisite for effective evaluation and as needing more serious attention than heretofore.

Evaluation should focus primarily on the effects of functional literacy on participants and their families, secondarily on its effects in the work units and communities and not at all on the impossible task of measuring its effect at the levels of the region (e.g. whole project area) and country.

A variety of tests must be devised to aid the selection of participants and matching control groups, and to measure achievement of participants at succeeding stages of courses and programmes. The Workshop examined tests used in the training of operatives in the British coal mining and motor industries. A suggestion was made that Unesco should commission a vocational training expert to prepare a manual on job analysis and test design.

An important and manageable area of study might be the extent to which literacy effects the efficiency of extension and training programmes, and another, related to this, the comparative evaluation of composite work-oriented functional literacy as against simple literacy leading on to separate vocational training. These might be fruitful areas for cost-effectiveness studies. Evaluation of instructor and teacher training and performance will also be a crucial means of improving the quality of the programme.

Evaluation at the participant level may aim to measure three kinds of change: attitudinal, (e.g. acceptance of the importance of fertilisers); behavioural, (e.g. correct use of fertilisers); economic, (e.g. increased crop production). The second is likely to be the most useful and manageable.

The difficulties of measuring the economic consequences of functional literacy in terms of productivity have already been mentioned; for this purpose evaluation should be at a low-enough level (e.g. one production line rather than a whole factory) for the changes measured to be attributable to the training given.

Systems analysis and cost-effectiveness studies may usefully be applied to various forms of organisation. The cost-effectiveness of various instructional systems should also be evaluated.

In view of the difficulty - and importance - of cost-benefit analysis, pilot experiments to pre-test techniques might be carried out in one or two selected projects, where objectives were defined in measurable terms, where explicit and quantifiable benefits could be anticipated and where the efficiency of budgeting and book-keeping could be relied upon.

The usefulness of evaluation must ultimately depend on the way its findings are communicated to decision makers. For this purpose a report is needed which can be understood and interpreted by the layman - the person with little or no training in the social sciences. At the same time social scientists will wish to have fuller details of the design and methodology used and the data obtained. Rather than two separate reports, the best solution was thought to be a short and readable report with full supporting appendices

# CHAPTER 3

## FUNCTIONAL LITERACY : DEFINITION AND EVALUATION

by

JOHN BOWERS

The Workshop is concerned with the evaluation of functional literacy, so it is suggested that as little time as possible should be taken in discussing functional literacy itself. Nevertheless one cannot realistically discuss evaluation without considering what is to be evaluated. This working paper, therefore, deals with definitions of 'literacy', terms used in describing it, and some of the problems which are likely to arise in evaluating it.

### Adult Literacy and Schooling

Although the primary school is the most important institution for eradicating illiteracy, the Workshop is concerned with the evaluation of adult literacy and related programmes of adult education and training. Unless otherwise stated, 'adult' will be taken to mean 15 years and over - the age normally used for census purposes.

### Definitions of Literacy and Related Problems of Evaluation

Literacy is defined as 'the degree to which an individual possesses the mastery over symbols in their written form, or is able to encode and decode written messages - to write and to read.' (1) One of the first aims of literacy evaluation may be to measure how far participants in a programme have achieved this: a process that may conveniently be called 'the evaluation of reading and/or writing ability.'

For census purposes, Unesco has defined a literate as 'a person who can read with understanding and write a simple message related to his everyday life.' This level of attainment might suitably be described as 'elementary literacy.'

This definition contains two relative terms - 'understanding' and 'simple', the one relevant to the reader, the other to the message. The reading ability of a reader naturally relates to the readability (the level of difficulty) of the message (reading matter) to be read. If the evaluation is to test the reader's level of reading ability, it will be necessary to establish standards of difficulty in the test reading matter (texts) to be read. If it is concerned to evaluate the message, the reading matter (e.g. in pretesting reading materials), it will be necessary first to establish the level of reading ability of the 'target audience' - the readers for whom it is intended.

---

(1) Rogers, Everett M. 'Modernisation among Peasants.' Holt, Rinehart & Wilson, p.429. See Ch.4, p.72. (Available for study.)

This is not as simple as it sounds, especially if it is desired to standardise levels of literacy (and tests) for different cultures and languages.

Levels of literacy among adults are sometimes equated with levels supposed to be attained by children after a stated number of years of primary schooling, but this, for obvious reasons, is also relative and of doubtful value.

Numeracy

Literacy is often regarded as including some skill in arithmetic and even algebra and geometry (the third of the '3 Rs'). Evaluation may, therefore, also be concerned with testing 'numeracy.'

Interpretation of Visual Images

Another skill, which is complementary to literacy and numeracy is mastery over visual symbols other than words and numbers - especially representations of 3 dimensional 'reality' on 2 dimensional (flat) surfaces - pictures, photographs, diagrams, maps, charts etc. Illiterates generally have great difficulty in the perception and interpretation of visual images, and this may have to be taught and evaluated.

Here again we must distinguish procedures for testing the perception and understanding of the viewers from those designed to evaluate the communication achieved by the visual images (e.g. audio-visual aids.)

Languages and Literacy

Literacy is essentially an extension of the use of language from speaking and hearing to writing and reading. What language? In multi-lingual areas this may be a difficult question and a sound language policy may be an essential basis for a literacy programme. There may, for example, be a conflict between the psychological advantages of making a person literate in his mother tongue and the social, political or other advantages of doing so in a language of wider communication: e.g. the official national language. If functional literacy is to be 'work oriented' (see below) the obvious choice may be the language of work or of training for work. It is, however, more difficult, and takes much longer, to teach literacy in a second/foreign language than in the learner's mother tongue. If, however, the mother tongue is an unwritten language anything up to three years of work by linguists will be needed to analyse its dialects, give it an appropriate alphabet, establish its grammar, prepare word lists and perhaps a dictionary, before it can be used for literacy.

A literacy programme may, therefore, have to comprise:

a. linguistic research on languages spoken to determine language policy;

b. linguistic research (e.g. word, syllable, letter frequency counts) on written languages as a basis for the production of teaching materials; (1)

c. linguistic research on unwritten languages;

d. comparative evaluation studies of literacy teaching methods (e.g. in the mother tongue as against a a second language; of teaching a second language through the mother tongue as against direct teaching of the second language;

e. tests of linguistic ability (speaking/understanding; reading/writing) in a second language.

'Simple' Literacy versus 'Functional' Literacy

Much of what has been written above refers to literacy *per se* - as an end in itself - as an extension of the ability to communicate - irrespective of the purposes for which this ability is to be used. The term 'simple' literacy seems appropriate to distinguish this ability from 'functional' literacy. A more common practice is to distinguish 'elementary' literacy from 'functional' literacy, but this is a false antithesis, for there can be elementary levels of functional literacy; though, of course, the more advanced it is the more likely it is to be 'functional.' It is therefore suggested that the contrast should be drawn between 'simple' and

---

(1) In the Vale do Rio Doce literacy project in Brazil a computerised study of syllable frequency in Portuguese is said to have halved the time taken to teach literacy to miners. A report in French is available.

'functional' : 'elementary' and 'advanced' literacy.

Before turning to definitions of functional literacy there is a further term to be defined.

## 'Permanent' Literacy

It is obviously important that the skills of reading and writing, once acquired, should not be lost. Any acquired skill is liable to fade if it is not used; the ability to speak a foreign language, for example. Regression into illiteracy is a well known phenomenon. So 'permanent' literacy is not likely to be permanent unless it is functional and is continually used. One of the reasons for making literacy functional may then be to ensure that it is permanent.

An interesting research project is now being undertaken in Tunisia by a team from Harvard University (USA) to study the retention of literacy in various environments and after various periods of initial study. (1)

## 'Functional' Literacy

A building is said to be functional if it is designed and constructed to serve as efficiently as possible the specific purposes for which it was conceived. Thus, in architecture, the term 'functional' is contrasted with 'generally useful' or 'decorative.' The word carries connotations of both efficiency and of specificity in serving defined aims. In relation to functional literacy the question therefore arises - what aims? literacy for what? Literacy has already been described as an extension of the ability to communicate. The first answer therefore is literacy for better communication.

Literacy gives a person the power both to send messages, by writing, and to receive them, by reading. Rogers (2) points out that it also allows a person - 'to control the rate of message input' (i.e. to read and absorb information at his own speed in his own time;) to 'store and retrieve print information for delayed use,' (thus substituting for or reinforcing memory;) and that it 'unlocks more complex mental abilities,' such as the ability to 'generalise through symbolisation' (to think and express oneself in general and abstract terms.)

---

(1) Report expected by September, 1969. See also Rogers, op cit., Ch.4.

(2) Rogers, op. cit., pp.70-72.

Evaluation of these powers of communication and related mental processes may be more important than simply testing a person's reading and writing ability.

But the answer: 'literacy for more efficient communication' only postpones the question, which now becomes: 'communication for what? why? of what?'

## Definitions of Functional Literacy

'Functional literacy' was defined by an international Committee of Experts on Literacy, convened by Unesco in Paris in June, 1962 as follows:

"A person is literate when he has acquired the essential knowledge and skills which enable him to engage in all those activities in which literacy is required for effective functioning in his group and community, and whose attainments in reading, writing and arithmetic make it possible for him to continue to use these skills towards his own and the community's development."

This is a rather elliptical definition. It essentially says: 'Literacy is required (is functional) for what literacy is required for.' But it does suggest that functional literacy must be efficient, specific to the needs of particular groups, and permanent.

A more recent description of functional literacy is given in the Final Report of the Teheran Conference: (1)

"Adult literacy, an essential element in over-all development, must be closely linked to economic and social priorities and to present and future manpower needs. All efforts should therefore tend towards functional literacy. Rather than an end in itself, literacy should be regarded as a way of preparing man for a social, civic and economic role that goes far beyond the limits of rudimentary literacy training consisting merely in the teaching of reading and writing. The very process of learning to read and write should be made an opportunity for acquiring information that can immediately be used to improve living standards; reading and writing should lead not only to elementary general knowledge but to training for work, increased productivity, a greater participation in civic life and a better understanding of the surrounding world, and should ultimately open the way to basic human culture."

---

(1) Unesco: Final Report ED/217 World Conference of Ministers of Education on the Eradication of Illiteracy, Paris 1965.

The term 'functional literacy,' when it was first used perhaps twenty years ago, generally meant literacy at a sufficiently advanced level to be used for practical purposes, such as reading newspapers, extension literature, instructions etc., writing letters or simple reports, and where numeracy was included doing elementary arithmetical operations. It now tends to imply not simply a level of literacy that is useful but literacy that is actually used to acquire useful knowledge and skills. Indeed the term seems to be applied not so much to the learner's state or level of literacy but more to the aims, strategies and methods of literacy teaching. Thus a number of new terms, which will now be considered, are currently used to describe 'functional literacy.'

'Work-oriented' or 'job-oriented' functional literacy

The major pilot projects of the Unesco 'Experimental World Literacy Programme (see below) are described in official documents as 'Work-oriented.' This appears to carry some or all of the following implications:

a. that the primary aim of literacy teaching is training for work; (thus the question: 'better communication for what?' is answered 'for technical and vocational training');

b. that the content and vocabulary used for teaching reading and writing, and the operations learnt in arithmetic, are largely or wholly focussed on the learner's job or livelihood, the problems he meets and must learn to solve, the technical terminology he needs, the tools he uses and the skills he must learn;

c. that literacy is used (for job training) while it is being acquired and acquired by being used;

d. that 'literacy skills' (3 Rs) are taught more or less (i) at the same time (sometimes in paid working-time,) (ii) in the same place (often the work-place,) (iii) by the same people (often the job-training instructors,) (iv) using the same media (manuals, work sheets, visual aids) as the technical knowledge and manual skills they are intended to communicate.

This integration of literacy teaching with technical/vocational training in industry and crafts, and with extension and training in agriculture, takes a variety of forms. Primers and readers used for literacy may also teach technical knowledge and skill (e.g. the Ibadan (Nigeria) Functional Literacy project for tobacco growers;) there may be no primers and readers but only 'work-sheets' and 'charts,' based on job training requirements or work problems, from which literacy is learnt (e.g. Algerian, Tunisian, Malian, Brazilian projects.) Learners may participate in constructing their own training materials, choosing what they wish to learn; or action-research may be used to feed back their needs and wishes to teams who prepare the training materials. Literacy teaching may be simultaneous or consecutive with technical/vocational training.

This trend has led to a rather illogical terminological inexactitude, largely propagated by Unesco documents; for the term 'functional literacy' has come to mean, not literacy that is functionally related to, or aims to promote, technical/vocational training, but the whole amalgam or combined programme of technical/vocational training-cum-literacy. In other words 'work-oriented functional literacy' means technical/vocational training with a literacy component. The word 'literacy' of course, has sales value so perhaps the inexactitude can be forgiven, but it gives rise to considerable confusion.

Nor is the picture yet complete, for the Teheran definition (para. 22) speaks also of 'social and civic' education, 'general knowledge,' 'understanding of the surrounding world' and 'basic human culture.' And what about health education, home economics, nutrition, family planning? Thus 'functional literacy' comes to mean comprehensive education and training for illiterate (and even semi-literate) adults, with a literacy component built in. It is in this sense that the term is used in the Unesco manual and it is suggested that the Workshop accept this rather inexact but practical definition of 'functional literacy' and that it use the term 'simple literacy' to mean the teaching of the 3 Rs alone.

### Literacy, Adult Education and Training

So the answer to the question 'literacy for what?' becomes more complex. Literacy for improved communication, communication for education and training (which is more than communication, since it involves the complex processes of learning and behaviour change.)

### Evaluation of Adult Education and Training

The evaluation of 'functional literacy,' in this sense of an amalgam of adult education and training is likely to involve the testing of participants' achievements, far beyond the evaluation of reading and writing ability and numeracy. The Workshop will need to consider what tests and other instruments of evaluation will be needed to measure the effectiveness of job training, agricultural extension etc. in communicating new knowledge and skill and bringing about desired changes in the participants' behaviour and practices. Discussions on testing methods used in British industrial training should have some relevance to this. A distinction may also have to be made between evaluation of the effectiveness of the whole 'functional literacy' amalgam (the total education-training-cum-literacy programme) and evaluation of the literacy component. This raises the question: just what is to be the independent variable in the evaluation of functional literacy projects? It may, for example, be interesting to compare oral training and extension for illiterates without a literacy component with the functional literacy approach.

## Functional Literacy, Development and Modernisation

A further question remains: education and training for what? The usual answer is 'development' and it is often said that the ultimate aim of functional literacy is development. But what is 'development?' Rogers (1) defines it as 'a type of social change in which new ideas are introduced into a social system in order to produce higher per capita incomes and levels of living through more modern production methods and improved social organisation.' This is a definition of social and economic development. What about individual development? Rogers calls this 'modernisation,' which, he suggests 'parallels at the individual level what development represents at the national level.' 'Modernisation,' he writes, 'is the process by which individuals change from a traditional way of life to a more complex, technologically advanced and rapidly changing style of life. Modernisation is not identical with Europeanisation or Westernisation; it is not necessarily 'good' for all members of less developed countries, since it provides 'new opportunities and prospects at a high price in human dislocation and suffering.' (2) Perhaps one of the aims of functional literacy for the individual should be to promote modernisation - or individual development - with understanding and intelligent participation, in order to minimise dislocation and suffering.

---

(1) Rogers, op. cit., Ch.1 pp.17-18. (An extract will be available.)

(2) Rogers, op. cit., Ch.1 p. 15.

## The Selective-Intensive Strategy for Functional Literacy

Unesco, starting with the World Congress of Teheran, has encouraged its Member States to experiment with what it has called the 'selective-intensive' strategy. This is based on the view that if literacy is to be functional it must be efficient and specific. The strategy is selective in the sense that literacy projects are planned for and located in specific sectors of the national development plan (generally priority sectors) specific areas of the country (generally development areas.) These sectors and areas are generally selected on the criteria that within them:

a. mass illiteracy appears to be a handicap to modernisation and development;

b. the infra-structure and organisation exist to enable functional literacy to be effectively used to promote modernisation and development:

this means where there is the need, the motivation and the outlets for functional literacy.

The strategy is selective in another sense in that programmes, methods and media are adapted to the needs of specific occupation groups (i.e. 'work-oriented.') This implies that specific programmes and media have to be prepared for comparatively small groups and even for specific jobs - a requirement which raises questions of cost-effectiveness. This may need evaluation (e.g. the comparison of 'simple' literacy followed by specific training with functional literacy which integrates and synchronises literacy teaching with technical/vocational training.)

The strategy is 'intensive' in contrast to the traditional 'extensive' 'mass' literacy campaign. It favours quality rather than quantity, stresses efficiency and generally requires longer and more elaborate training. This raises serious questions for evaluation of cost per capita against returns.

Although the selective-intensive strategy has gained many adherents, the 'extensive' campaign approach is still used, and has recently proved successful in eradicating mass illiteracy in Cuba. An impassioned plea for it is made by Sir Charles Jeffries in his book 'Illiteracy - a World Problem.' (1)

---

(1) See Unesco Courier, April, 1968, p.26. Extract from 'Illiteracy - a World Problem' by Sir Charles Jeffries, Pall Mall Press, London.

## The Experimental World Literacy Programme

Unesco launched in 1966 an Experimental World Literacy Programme. Fifty-two Member States have announced their intention of participating in it. Major pilot projects have been launched in twelve countries (Algeria, Ecuador, Ethiopia, Guinea, India, Iran, Madagascar, Mali, Sudan, Syria, Tanzania, Venezuela.) These Pilot Projects and a number of other 'Micro Experiments' (e.g. in Algeria, Brazil, Chile, India, Jamaica, Nigeria, Tunisia and Upper Volta) are designed to apply and evaluate the 'selective-intensive strategy of work-oriented functional literacy.' "The basic aim of the Experimental World Literacy Programme is to shed light on the nature of the inter-relationship between literacy and development and to study its significance." (1) Each project has an evaluation unit.

The Unesco Manual on Evaluation of Experimental Literacy Projects (ED/WS/135) which has been circulated to all participants of the Workshop, is intended to serve as a guide for evaluators in these projects. It should be considered as a working document and as background information for the Workshop's studies. The Workshop is not, however, bound by the policy and methods set out in the Manual, which should be considered critically, along with other sources of information.

## Aims of the Workshop: Study of the Evaluation of Functional Literacy

The aims of the Workshop are set out in document WEFL/5, and the Steering and Evaluation Committee has been asked to refine them and evaluate the degree to which they are achieved. In the broadest terms the Workshop is concerned, not simply to consider the evaluation of the pilot projects of the World Literacy Programme, but to study, discuss and discover policies, organisation and methods for evaluating functional literacy, and for building action research into functional literacy projects, appropriate to participants' countries.

---

(1) Unesco document 15C/52 of 12th September, 1968: "The Position as Regards Functional Literacy Pilot Projects."

## DISCUSSION ARISING FROM MR. BOWERS' ADDRESS

The Workshop agreed that the discussion following the address by Mr. Bowers might usefully range across the entire field of functional literacy and its related problems and not be exclusively concerned with evaluation. This would enable a common understanding to be arrived at from the start by all participants on the scope and terminology of functional literacy and would obviate the need to divert the discussion subsequently from the essential topic of evaluation.

### The equation of levels of attainment between adult literacy programmes and the formal school system

Doubts were expressed as to whether adult literacy courses could or should be correlated as far as levels of attainment were concerned, with levels of literacy attained by children at recognised stages within the formal school system. It was admitted that in many countries adults were often motivated towards literacy classes by the desire for conventional education and for certificates comparable with those gained in schools. This might apply, for example, where husbands had generally proceeded further with formal education than their wives. With literacy classes designed to satisfy needs of this sort, the parallel with the formal school structure might not be inappropriate.

It was noted that in parts of Latin America many adult literacy courses were in fact delayed primary school courses. Several participants argued that the concept of work-oriented functional literacy tended to over emphasise purely economic aspects at the expense of human and social aspects and that it might be desirable to extend the scope of functional literacy beyond an exclusive concern with training in particular skills. For example, it was felt to be especially important that communities undergoing rapid social and economic change should have an understanding of the process of development. Such understanding might serve to obviate some of the often painful social side effects of rapid development. It was, however, agreed that educational programmes associated with the formal school often had the effect of developing a predominantly abstract and verbal frame of thought which often disregarded the social and economic realities of particular countries. The value of the work-oriented approach to functional literacy teaching might lie in its curtailment of verbalising and its emphasis on practical skills and economic realities.

In general it was agreed that the equation of levels of adult literacy with specific levels of primary school education might be useful in general literacy programmes but was a rather doubtful exercise where literacy teaching was work-oriented. Here it was agreed that a clear definition of the attainments sought by functional literacy teaching should be arrived at and that tests should be devised to find out if those objectives were achieved. Such tests could not be applicable to children of ten years of age working within an academic system of education.

## Selection of participants for functional literacy programmes

There was some discussion of the problem of selecting suitable individuals for functional literacy training aimed at specific vocational areas. Testing techniques should be worked out which would indicate the latent abilities even of illiterates, although this presented particular problems. Several participants stressed the need for such tests. A further need was for tests which would enable the selection of uniformly matched control groups for evaluation purposes.

## Age groups

Two issues were discussed: first whether children out of school from the age of ten years upwards should be taken into programmes designed for adults, and secondly what form of distinctive treatment should be given to children within this age group. Young people who had not been able to progress within the formal school system often tended to appear in adult classes and this, on the whole, was not to be encouraged. There were, however, serious practical difficulties in removing such young people from adult classes.

## Work-oriented functional literacy and practical subjects within the general educational curriculum

The Workshop considered the Ghandian approach to 'Basic Education,' with its special emphasis on practical activities, and its possible parallels with work-oriented functional literacy. It was agreed, however, that the purpose of the Ghandian approach was to counteract the verbal emphasis in academic schooling by the inclusion of practical activities in the curriculum. This approach was essentially educational in its purpose, not functional or work-oriented. It did not aim at providing a training for employment.

## The evolution of work-oriented functional literacy

The process whereby Unesco had evolved the present concept of work-oriented functional literacy was traced from the earlier concept of 'fundamental education,' through a phase when the eradication of mass illiteracy by means of a World Campaign was insisted upon. The new approach formulated at the Teheran Conference in 1965, emphasised the integration of literacy teaching and vocational training and aimed at special groups which might be expected to make an important contribution to the process of development. It was stressed that this approach was new and complicated, and demanded a high degree of co-operation between Government departments and many other agencies, and required specialised materials to be prepared for particular groups and was therefore costly. Because of this high cost factor, attention had been drawn particularly to the need to evaluate work-oriented functional literacy in cost/benefit terms. It was expected, therefore, that the Workshop would give particular attention to this problem.

### Instructors for functional literacy programmes

A major problem was discerned concerning the personnel who would carry out functional literacy teaching. Many programmes were dependent upon conventionally trained school teachers who often knew less about the practical aspects of the vocational subject taught than did the illiterates who were being taught. Three possibilities were seen by which this problem might be tackled:

teachers might be subjected to intensive training courses in particular vocational subjects;

the extension workers and vocational instructors might be engaged in the functional literacy classes;

technical expertise might be concentrated into the production of effective educational media; the programme would then depend heavily upon those media to get the technical message through to the participants.

Those three possibilities were not, of course, mutually exclusive and might often be combined in various ways in a single programme.

## CHAPTER 4

## EXPERIENCES AND PROBLEMS IN EVALUATION - STATEMENTS BY PARTICIPANTS

In the original letter of invitation, Unesco asked each participant to be prepared to give a short talk to the Workshop on his experience in evaluation and on the evaluation activities which were in progress in his country. These statements revealed a wide range of common problems, as well as a variety of expedients which had been, or were being, adopted in order to tackle them. This chapter of the report brings together the main points discussed and the problems outlined by the participants, together with a lengthier account of one evaluation exercise described.

### Data-gathering

Several participants expressed doubts as to whether in practice it was possible for data to be gathered at the earliest stages of the life of a project which would serve as a baseline survey for scientific evaluation. Basic data could, however, be gathered, with recognised limitations, on particular problems within the project area which might be very useful in the planning of project operations.

Particular problems facing such preliminary work were the scarcity or absence of documentary materials on development problems, the lack of reliable census reports and the physical problems of gathering information, due to the inaccessibility of many parts of the project areas, the lack of roads and the scattered nature of many settlements. There was a particular problem with small farmers who were resistant to the provision of information on their holdings, crops, animals, returns, etc. They were suspicious of interviewers and of government generally. A related problem was found in certain industrial situations where factory managers were hesitant to provide information on productivity or on particular factory processes, which they regarded as potentially valuable to competitors. On one instance cited, time and motion studies, which could be of great use in the evaluation of functional literacy in a particular industry, were known to exist but could not be made available to the evaluation team because of the insistence of the factory management that this information should remain secret.

### Motivation studies

Some participants emphasised the need for research into the problems of motivating people, at whom literacy programmes were aimed, towards active support for, and participation in, these programmes. It was pointed out that in many rural areas there exists a powerful interest among adults in the formal education of their children and that this motivation is often stronger than that which they themselves might have towards literacy. In certain rural areas where development had not proceeded very far, the older generation, particularly in the instance cited the older men, were very difficult to convince that literacy, whether traditional or functional, could give them anything.

Examples were cited where interest in literacy was particularly strong among women. Their motivation was, however, not based on the desire to perform a particular task more efficiently. Often the formal education of women had lagged behind their menfolk and in such simple literacy was regarded as a means of compensating for the earlier neglect of their education relative to that of men. In another instance quoted, the opportunity was being taken to introduce functional literacy in an industrial setting aimed at female workers with a substantial element of home economics as a supplementary motivating factor.

## Government policy and evaluation

In certain cases, the aims established by governments were too idealistic and included elements which by their very nature could not be objectively evaluated. Gains were often claimed as a result of literacy in terms such as 'increased self-confidence,' 'receptivity to new ideas,' 'ability to solve practical problems.' It was very doubtful whether realistic evaluation procedures could be established to measure such factors.

In cases where it was clearly intended that a functional literacy programme should contribute directly to development, it was regarded as of the greatest importance that the signatories to the agreement between Unesco and the country in which the programme was to be established should include representatives of those ministries and agencies already involved in development work and upon whom the programme might subsequently depend for technical support. Many of them might already possess information and insights of crucial importance to both the operational and evaluation activities of the programme and ought therefore to be enlisted in support of the programme from the beginning.

Particular problems for evaluation arose in cases where, although it was agreed that the literacy programme should be aimed at areas of the economy scheduled within national development plans for accelerated development, the literacy input went ahead in isolation and other development inputs did not appear. Again the need was stressed for co-ordination between the departments of government involved in development work and those directly concerned with the literacy programme.

## Personnel problems

These were located in three areas:

International professional evaluation personnel;

National professional personnel;

National supporting staff to the evaluation teams.

With regard to the international evaluation experts, it was pointed out that these had in several cases arrived late, often after the initiation of operational activities. This had resulted in failure to carry out initial data-gathering before the launching of operational activities and had led to the adoption of operational plans which did not provide for evaluation procedures.

Qualified national counterparts were very scarce, there being little provision within established degree structures in most universities in the countries concerned for the kind of composite training required of an evaluator. A plea was made that Unesco should approach selected universities in developing countries with a request that courses should be established which would provide the kind of broad and practical training in the social sciences which was required. It was also recommended that Unesco should involve national counterparts as far as possible in international and other conferences at which evaluation was discussed.

With regard to supporting staff, particular problems arose where the general level of formal education was comparatively low. For several evaluation exercises, interviewers had had to be recruited who had little formal education and therefore, despite preparatory training programmes, little real understanding of the task that they were carrying out. Various administrative problems arose from this situation. The professional staff of the evaluation unit, already bearing a heavy work-load, had to devote a substantial period of time to providing a limited training for interviewers. In cases where it was necessary to make use of secondary school students as interviewers, this necessarily implied that interviews had to be carried out at weekends, when the students had some leisure time available, and not at times which might be more appropriate in terms of the evaluation exercise itself.

## Language

A major problem was presented by language policy in literacy programmes. It was already clear that there was a potential conflict between national languages, which were supported by the whole weight of government, and local vernaculars, which, despite all pressures, remained strong in many areas. The identification of literacy programmes with languages which were little spoken and locally unpopular could gravely prejudice the success of the programme; and, related to this, the identification of evaluation staff, particularly interviewers, with an unpopular language policy could also gravely prejudice the work of evaluation.

## Control groups and indicators in rural areas

In a number of countries there were now very few rural areas which were not continuously subject to the widest range of educational influences from outside. This was particularly so with the rapid development of the media of mass communication. Especially difficult problems were posed when it was sought to isolate one particular educational influence and to estimate the contribution of this influence alone to the development process.

Crucial importance was seen in the selection of suitable indicators directly relevant to the rural situation. In the determination of such indicators, evaluation units should be able to draw upon the active co-operation of the social science research specialists attached to local universities. The selection of these indicators should ideally be done well before the launching of the first literacy classes.

## Action-oriented research in support of the operational programme

Particular areas were defined where action-oriented research could have a constructive influence in programme planning and be of direct and immediate service to the overall activities of the programme:

the educational effectiveness of media, ranging from flashcards to educational television;

the costs and benefits of particular media;

the selection and training of instructors;

the location of areas of society likely to be responsive to functional literacy programmes;

the location of problems upon which the literacy programme might focus.

## The timing of functional literacy programmes and evaluation activities

Doubts were expressed by several participants as to whether Unesco was trying to go too fast in the launching of functional literacy programmes and in their evaluation. It was felt that the stipulated period of one year between the signing of the Agreement between Unesco and the host country and the launching of the first classes was unrealistic. This 'preparatory year' was supposed to include the recruitment of evaluation personnel (both international experts and national counterparts), the carrying out of preparatory surveys, the preparation and testing of educational materials. Two years was likely to be a more reasonable time for the completion of these activities but doubts were expressed whether any precise time limit could be determined, first, because each country was unique in its nature, and secondly, because of the known difficulties and delays in recruiting staff. It was noted that in one project where there had been a full complement of evaluation staff, it had still not been possible to complete the recommended programme within a single year.

## An experiment in the evaluation of the use of television in farmer training in the Delhi area of India

Many of the problems and questions discussed by participants were revealed in an account of the evaluation by the Indian National Council for Educational Research and Training of the use of television in the training of farmers in the Delhi area in the techniques required for the introduction of high-yielding varieties of crops. Despite the fact that this experiment was not concerned with functional literacy as such, the techniques employed were undoubtedly relevant and this was the only example provided of a completed project evaluation.

Agricultural television was set up in the Delhi area in 1967 in order to disseminate information on improved farming techniques. Eighty villages within twenty-five miles radius of the city were selected for the project. Various interested governmental agencies combined to set up the programme. Television sets were provided to the selected villages by the Department of Atomic Energy; the technical content of programmes was provided by the Agricultural Research Institute; All India Radio was responsible for the production of programmes; the Delhi local administration organised television clubs in twenty per cent of the villages, selected at random, the purpose being to ensure that farmers viewed agricultural programmes on television regularly; the evaluation of the impact of this programme was entrusted to a body not committed to the programme - the National Council for

Educational Research and Training.

Telecasts were made to farmers twice weekly, the content of the programmes being tailored to the agricultural seasons, cropping patterns and practices. For evaluation purposes, the interest lay in assessing impact and obtaining feedback for the modification of the content of the broadcasts. In the first case, it was important to discover the knowledge gained by farmers; their attitudes to farm practices and how this had changed, their adoption or otherwise of new practices and the reasons for this. In the second case, it was important to discover their reactions to the broadcasts as they were in progress, isolating for example difficulties over language and feeding back the suggestions of the recipients to the agencies responsible for production.

Ten villages were selected for each phase of the project, carefully matched on levels of development. The social character of the village was analysed and also the levels of literacy in the village. A control group of villages was selected, carefully matched to the television group. Care was also taken that the control villages should be as detached from the television villages as possible.

Before the broadcasts began a baseline schedule was prepared. Interviews were conducted with farmers listing their personal characteristics, agricultural knowledge and practices and attitudes to agricultural change. The results indicated the agricultural knowledge of each individual and provided a measurement of his likely attitude to the television broadcasts.

Evaluation of the rate and degree of adoption of changes by participants was a longer process. Interviewers were selected from holders of masters degrees in the behavioural and social sciences. The investigators received special training emphasising the social factors which might inhibit effective interviewing. A programme assessment schedule was then prepared and applied in the television villages. Information which derived from the programme assessment schedule was immediately sent to the agencies responsible for producing the programmes and used to modify content.

The measurement of the impact of the television programmes was carried out by means of baseline and terminal surveys of both the television group and the control group. Results showed that those villages which received agricultural television programmes had positive gains in terms of knowledge, attitude and adoption of new practices.

The Delhi experiment seemed to show clearly that the experimental factor, in this case farmers' training through television produced significant differences in the attitudes of farmers to agricultural change and in the application of improved techniques. If the groups studied were well matched then these results would be valid. It was necessary, however, to check whether there were other influences which might have contributed towards this difference in attainment. The investigators, therefore, sought to isolate any extraneous influences which might have produced exaggerated differences by probe questions seeking to discover from respondents the source of any new information which they claimed to have received. In addition, investigators carried out participant observations by means of overnight stays in the villages.

## CHAPTER 5

## THE EVALUATION OF FUNCTIONAL LITERACY : AN ECONOMIST'S VIEW

by

PROFESSOR M. BLAUG

Professor Blaug indicated that he would concern himself mainly with the problems of evaluating functional literacy programmes from the economist's point of view and particularly with reference to the suggestions put forward in the Unesco Manual. He stressed that there had to date, been no serious attempt to measure the returns of functional literacy programmes in terms of resultant increased productivity and that this was a particularly difficult exercise involving complex techniques which should not be over-simplified. In particular, he expressed grave doubts whether it was possible for four levels of effects to be measured, as suggested in the Unesco Manual. It was realistic to look for economic consequences of functional literacy at the level of the participant and possibly at the level of the work unit. At the level of the community the exercise was dubious; at the level of the whole economy, it had little practicality.

A particular feature of evaluation attempts which had been made so far was that these had been conducted inwards. The initial task had been to try and determine the effects of programmes on the community and thereafter to work towards the work unit and the participant. The revised Manual on Evaluation sought to reverse this, starting with the individual Participant and working outwards. The notion, however, persisted that measurable effects of a functional literacy programme in isolation could be traced on the community. Since the community would be subjected to very many other influences, the isolation of the impact of literacy in itself was not likely to be a profitable exercise. It was proposed that discussion should be concentrated primarily on measuring the returns from functional literacy programmes to the new-literate himself. Thereafter some consideration should be given to whether there were spill-over effects into other sectors. It could be expected that the work unit, consisting of an aggregate of new literates, might be affected. If conclusive measurements of the impact of such a programme on individual participants were achieved, plus some effects on the work unit, then the evaluation team would have a substantial achievement to its credit.

It was still an open question whether, when a worker was made literate, his productivity was immediately increased, or whether literacy alone made not much difference, since, essentially, productivity was dependent upon capital equipment, factory organisation, management, relationships with other workers etc. If this latter postulate is true, scientific measurement of the impact of functional literacy was not possible - the process is too slow, diffuse and lengthy. The real impact of functional literacy would then lie in the building up of a 'critical mass' of literates which, when supplemented by other inputs - seeds, fortilizers, land reform, roads, bridges - would accelerate development and affect productivity. Many functional literacy programmes do not actually involve the other elements in this package, at least not as they were usually conceived. Therefore it seemed important that those seeking to evaluate the economic effects of functional literacy projects should bear in mind the possibility that they did not achieve significant economic benefits. The process of how development is initiated is one regarding which there is little firm knowledge. Most economists would, however, argue that one factor, for example functional literacy, cannot in itself start the process; what is necessary is a combined attack on many fronts. It was essential, therefore, that evaluators of functional literacy should approach their task sceptically, bearing in mind that there might be no measurable economic effects.

How, therefore, should greater individual productivity be measured? In the case of the individual industrial worker, his employment setting involved machinery, management, material, as well as the human factor. The isolation of the individual's contribution to output from other factors of production was very difficult. The only possibility was through an examination of the individual's earnings, each employer defining the individual's financial worth to the enterprise in terms of what he produces. Ideally, the individual should be removed from the enterprise and productivity scrutinised in his absence. This was, however, usually impossible, although a similar exercise might be possible when a factory is in the process of increasing its recruitment. An acceptable calculation is also only possible if a labour market is perfectly competitive, meaning that an individual's earnings are equal to his individual contribution to output. In the case of developing countries, labour markets were far from perfectly competitive and were usually either tradition bound or hedged around with legislation governing wages. In this situation it was very unlikely that individual earnings would closely correspond to the individual's contribution to output. The first approach to measurement of individual productivity was therefore to examine earnings before and after the application of functional literacy.

Thereafter, there were two possibilities:

- to examine the working of the labour market in a particular community to discover whether the market worked competitively. This exercise was only possible for an expert labour economist. The techniques involved were difficult, but it might be possible over a lengthy period of time to discover whether competitive forces were driving earnings down to the level of productivity. Having clarified these issues, it might be possible, through comparing the individual's earnings before and after participation in a programme of functional literacy, to reach a valid assessment of the economic impact of the programme on the individual;

- another possible approach was through job analysis. This was almost as difficult. It required that the skill content of each job performed by the individuals studied should be analysed. Consideration should be given to the means of acquiring the necessary skills for the job and to whether this process called for literacy or whether it could be learned on-the-job. In such situations it was likely that the effect of literacy would be to curtail the training process. The period of time saved could be costed in terms of wages to trainer and trainee. An overall calculation could be arrived at in the saving in the cost of training brought about by the literacy programme. Literacy therefore in itself would not, in this way, raise productivity. It might save training time. Literacy was not always necessary for the improvement of industrial output. Industrial development was possible with illiterate workers. Literacy could, however, bring about savings both in the cost of training and the degree of supervision necessary. It might also improve the quality of workmanship and reduce the rate of industrial accidents and breakdowns.

With reference to the Unesco Manual (p.69) on the Evaluation of Experimental Literacy Projects, particularly in those recommendations concerning the extent of the contribution of functional literacy programmes to production, several criticisms were offered. It was not thought possible that the contribution of an individual worker to the 'production or other purposes' of an enterprise could be assessed in terms of three grades. To carry out the exercise suggested would involve a complex statistical computation estimating the relationship of all factors of production in the individual enterprise and its output. It would require an observation lasting possibly ten years. There would still be no valid way of clarifying the individual's contribution on a graded scale. All factors of production within an enterprise combine to produce output. It was therefore meaningless.

On page 70, item (b), it was agreed that the care and maintenance of machinery was measureable. A method of obtaining information frequently used was to ask the supervisor of the job. The result was often a biased answer. For a complete answer the evaluator would need to understand both the details of the techniques of production plus the technology of the machine. In a long-term project confined to a skilled industrial process involving relatively few machines, an accurate measurement using expert evaluators might be possible. In a diverse project involving a variety of processes and machines, it was very unlikely that a machinery observation over a period or periods would give an accurate view.

On page 70, item (c), doubts were expressed on the meaning of 'identification with the enterprise and degree of discipline in work.' The experience of many industries suggested that complete discipline of the workers in industry would be arrived at if the management were free to dismiss. In industries where the working force was organised, such a situation was very unlikely. Without the power to dismiss, discipline of the type suggested was unlikely to exist. The facility for obeying orders was not therefore likely to be influenced by the presence or absence of literacy. 'Identification with the enterprise,' was not regarded as measurable. A 'productivity contribution test,' including all the factors mentioned above, could not be regarded as reasonable or useful.

On page 72, item (b), with regard to self-employed participants, it was pointed out that in all developing countries there is a substantial sector of the labour force consisting of craftsmen, artisans, service workers, between the small body of wage earners in industry and the large body of workers in agriculture. They are a most diffuse category with erratic income levels and very traditional patterns of production. It could not be expected that a functional literacy programme would have an effect on their equipment or materials, which again were largely traditional. The impact of a programme would perhaps be seen in the selling practices and on the keeping of accounts. Gross income might therefore go up as a consequence of a literacy programme and this should be regarded for this category, as a sufficient measure of the effect of literacy. It was suggested, however, that this sector was not likely to produce large enrolments in functional literacy programmes. It was a sector of the labour market which was, to a large degree, tradition bound, conservative, with skills and crafts often passed down from father to son.

With reference to the impact of functional literacy programmes (page 60) on work competence of those engaged in agriculture, the particular difficulties of bringing about modernisation and change in agricultural practices was stressed. Yet this was regarded as of the greatest importance because of the large size of the agricultural sector of the labour force in developing countries and because this was the area where the need for progress was most acute. The root of the problem lay in the fact that the individual worker performed so many different tasks. In general, management and labour came from one and the same source. Agricultural productivity was affected by tools, seeds, fertilizers and the weather, of which the weather was probably the single most important explanation of year-to-year variations in output. The entire process of agricultural production was many times more complicated than the industrial process and the flow of technical information to the producer, the individual farmer, was more critical in agriculture than in industry. In the case of industry information was available and could be used. In the case of agriculture, although information on soils, seeds, fertilizers, cropping patterns, markets, was becoming increasingly available, it was usually beyond the resources of the individual farmer either to understand or to use. The purpose of the agricultural extension service was to bring information to the farmer. Agricultural extension was, next to the weather, the most critical input. Literacy should be seen in relation to agricultural extension - as a means of facilitating its work. This was regarded as more important than the direct effect of literacy on agricultural production. Literacy among farmers increases the servicing scope of extension workers and improves the quality of communication between extension workers and the farmer.

In this way literacy could have an immediate impact on the individual farmer and on the work unit. The measurement of this process was, however, complicated. It might be more readily measured in terms of crops grown if the farmer produced more of what he produced before. This was, however, not the usual pattern. The effect of making the farmer literate might often be to involve him in the production of new crops, in which case his income might rise and the money earned might be used as a measure of his increased productivity. It should be remembered, however, that in developing countries many farmers were not entirely engaged in production for the market. There remained a substantial subsistence element in their work. Their primary task was often to produce the means of survival - food. Cash was earned through selling the surplus. Where a substantial subsistence element remained, the effect of literacy might be to increase human consumption quantitively and qualitively. This was not measurable in monetary terms. There was, however, a need to assess changes in home production. The evaluation of the effect of functional literacy on agricultural production was therefore concerned with the measurement of a total mix, including the non-monetary element. The situation was made more complicated by the fact that many small farmers in developing countries grew a very wide range of crops, often in very small amounts. They were also subjected to a continuing series of influences which might change their agricultural practices - new seeds, better extension, alteration in world commodity prices. In the latter case it might be possible for a farmer, through a functional literacy programme, to be better informed, to grow more but to earn less. There were great dangers in any effort at evaluating the impact of functional literacy on agricultural productivity in settling for superficial indicators of increased productivity.

Professor Blaug concluded by assuring the Workshop that there was no precise formula on how functional literacy should be evaluated. General guidelines could be offered. What was important was that the evaluator should know exactly what to look for and that in the evaluation effort there should be a concentration on unambitious, realistic targets.

## DISCUSSION ARISING FROM PROFESSOR BLAUG'S ADDRESS

An example was quoted of Western Nigerian tobacco farmers who were subject to the activities of an existing extension service related to tobacco growing. The service was provided by the Nigerian Tobacco Company. A decision had been taken by the Company to maintain the quantity of output at its present level but to work for the improvement of the quality. This was sought by first controlling the acreage which each individual farmer could cultivate, and second by putting in functional literacy directly related to tobacco cultivation and processing. The project evaluation was being carried out by the Institute of African Adult Education at the University of Ibadan; final results were not yet available. The main problem revealed so far was to determine why farmers were not adopting agricultural innovations to a greater degree and to isolate those factors which induce the farmer to adopt such practices.

This experiment seemed to show that if functional literacy was to be fully effective, it was necessary to ensure that it interacted with other factors, for example extension, provision of fertilizers, equipment. The literacy element in itself was not regarded as sufficient to bring about major increases in productivity.

A particular characteristic of this project, as far as evaluation was concerned, was that the Nigerian Tobacco Company provided a carefully graded system for the purchase of tobacco leaf. Prices paid to the farmers varied with the quality of the leaf sold. By scrutinising the amounts paid to tobacco farmers within the entire tobacco area and within the control area a price differential in the earnings of farmers receiving functional literacy and those outside the programme might be revealed. Despite this favourable situation, it was not expected that the Western Nigerian experiment would produce dramatic results when the data were finally analysed. It was believed that other factors were inhibiting the full adoption of the innovations promoted by the functional literacy teaching. In particular, since each farmer's area of tobacco was limited and his earnings therefore from this crop never likely to be very high, the farmers would not fully identify with the task of tobacco cultivation.

A particular feature of agricultural development was that in the extension services there existed a large measure of practical expertise on those factors which made farmers receptive to new ideas. Evaluators of functional literacy projects in agricultural areas should, therefore, talk with the technical officers of the agricultural extension service from whom much useful practical information on the attitudes to innovation of local farmers would be available.

In industrial projects, the value of time and motion studies in relation to individual jobs was discussed. Such studies were, however, rare. In the case of Britain, for example, it was estimated that only approximately 5% of factories had detailed time and motion studies available. In developing countries the estimate was perhaps 1%. To provide such studies in order to facilitate the evaluation of the impact of functional literacy in industry, was a skilled, difficult and expensive job requiring professional attention for several months. It would be necessary to scrutinise the operation of individual machines in detail. Each machine might take two to three days. In a textile factory there would be perhaps thirty to forty different types of machines. This process having been completed, a measure of output per job would be available which would be of value in assessing the effect of literacy on the operatives. The process of providing this data would, however, raise the cost of evaluation to a very high level.

Where the industrial plant was imported and set up entirely from an external source, it might be possible to take into consideration data which applied to the equivalent machinery in the country of origin. There were, however, many dangers in this. For example, although in a steel mill, whether in India or in Germany, the basic machine processes took the same time and used the same skills, there were major differences in the entire organisation of the factory, particularly in bringing the materials to the workbench and taking the finished products from the workbench. In India this would be done by hand, in Germany by machine. In Germany this aspect would perhaps involve 5% of the labour force, in India 60% Comparability, therefore, from the viewpoint of job analysis was not really practical.

The Unesco Manual did not pay much attention to the problem of job analysis. It was suggested that Unesco should commission an expert in vocational education to prepare a handbook of job analysis, setting out what is involved in analysing the skill component of jobs in industry. Such a handbook would be of great importance to the evaluation of functional literacy in industry.

## CHAPTER 6

## COST ANALYSIS IN THE EVALUATION OF FUNCTIONAL LITERACY

by

DR. R. JOLLY

The problem considered was how the basic concepts of cost analysis might be applied to functional literacy programmes. There were some doubts whether comprehensive cost/benefit analysis would be possible if, by this, it was sought to encompass all benefits in a single number and contrast this with all costs, again in a single number, giving a cost/benefit ratio which would indicate clearly whether functional literacy was a profitable pursuit in economic terms. Functional literacy programmes were extremely complex, having political as well as economic dimensions. The issues posed could not be summarised in a single number.

Even when the exercise was confined to purely economic benefits the margins of uncertainty remained very great and the validity of a single number computation was still doubtful. Rigorous cost analysis was however valuable about parts of such programmes. The results of this could be fed into a broader framework of decision-making which would act as a guide in the more effective use of economic resources, or produce whatever benefits accrued from functional literacy programmes at a cheaper price.

### The general principles of cost calculation

To the layman, apparently this was not a very complex exercise. It could be seen as a matter of accounting, i.e. the costs would be equal to the sum of the expenditures on a particular programme. For the economist this was a very unsatisfactory approach. A clear definition should be seen in terms of 'opportunity cost', defined in the Unesco Manual as 'the value that would have been gained if the resources had been expended in the next best use.' This concept could be illustrated with reference to the teachers, the students and the classrooms required for the operation of a functional literacy programme.

With reference to the teachers, consideration should be given to the value of these teachers if they were not used in the functional literacy programme, but were used elsewhere. In a situation where skilled manpower was fully employed the value of the teaching force in terms of production might be equal to their earnings elsewhere. This was, however, not always a useful exercise. Where there was large-scale unemployment there might be no productive jobs available outside the functional literacy programme. Moreover, the wage structure, as in many developing countries, might be unrelated to productivity, for example where the wage structure was inherited from an earlier colonial era.

With reference to the students, these far outnumbered the teaching body. They were not usually paid for participating in fuctional literacy programmes. The simple accounting process therefore would tend to underestimate the cost of their use. In the case of primary school students, there was perhaps not much potential value, although in, for example, agricultural situations in Tropical Africa, even small children could make a valuable economic contribution. As far as adults are concerned, to remove them from their everyday work and put them in a classroom situation would mean a loss of production and therefore an opportunity cost. In this case, there might be a difference between functional literacy programmes operating in the day time, involving the extraction of participants from their work, and those which are operated in the evening; even this, however, did not always operate neatly. Tobacco farmers, for example, carried out flue curing at night. It was necessary to examine carefully what the participant gave up in order to participate and cost it if possible; even leisure should be examined. There was, for example, a distinction between leisure time used for home production, as against home consumption and recreation.

The difficulty of the exercise underlined the basic meaninglessness of trying to put a single figure on all costs or all benefits. It also underlined the value of units of measurement other than monetary units. Several elements in a cost analysis could be identified which might best be expressed in non-monetary terms. In essence, it was very important not to ignore manpower costs. To ignore opportunity cost of unpaid students (or other manpower) was to assume they were zero - usually a worse estimate than to make a crude calculation of the loss of production from the withdrawal of labour for participation in functional literacy programmes.

With reference to classrooms, consideration should be given to the estimation of the value of the use of buildings in the best alternate way. If classrooms remained empty and there was no competing use, then the opportunity cost was zero.

## An illustration : the Cuban Literacy Campaign

These features were illustrated with reference to the Cuban literacy campaign. In 1961 the Cuban Government decided upon a mass literacy campaign. Out of a population of approximately seven million, 22% of the adults (those over 15 years of age) were illiterate, a fairly low rate for Latin America. Illiteracy was spread uniformly throughout all age groups. The objective of the Cuban campaign was to make everyone literate within one year, a decision based essentially on political factors. One million illiterates took part in the campaign and it was claimed that 700,000 were made literate. To carry out this bold plan, all formal schools were closed from April to December, 1961. Five-sixths of the teachers were used as organisers in the literacy campaign. Teams of 'Brigadistas' were organised to operate in the main in the rural areas, led by organisers, representative teachers, local government and the political party. Direct teaching was mainly carried out by secondary school and university students, although some upper primary school students were also involved. The teaching process was based upon the idea of 'each one, teach one.' Each Brigadista would live and work with a family, devote two hours per day to literacy instruction and the remainder of the time to assisting the family with its everyday work. In the main, positive results were achieved within three to five months.

The opportunity cost of this programme is in part calculable in terms of the loss of eight months' formal schooling although this is complicated by the fact that the following school year was extended. In addition, it is necessary to calculate how much production was lost while the illiterates were studying. Because classes were organised out of working hours, and students helped with the normal daily work, this possibly was not too much. However, in a country which was prepared to take such a bold step, as the Government of Cuba in its literacy campaign, even time after normal working hours might have been organised for productive work. If not for the literacy campaign, then for something else. Evening hours therefore can have an opportunity cost.

## Cost calculations - some detailed points

The application of the concept of opportunity cost does not provide any simple answers. It is simply a way of looking at costs comprehensively. In attempting to calculate the value of alternative activities to functional literacy classes, many uncertainties are posed involving political and social factors as well as technical and economic factors. The value of alternative activities cannot be measured within a static framework.

The question of who met the costs was also important.

The students themselves might pay by sacrificing earning opportunities. Or the factory owner might decide to bear the costs. This still meant, however, that the student body was not producing what they might otherwise have produced.

Reference was made to a literacy experiment in a Brazilian factory in which the workers, highly motivated towards their employment because of a functional literacy programme, had made up their production losses by harder work after the literacy classes. Did this mean that the literacy class had no opportunity cost? The answer depends on what would have happened in the absence of the literacy classes. Perhaps any exhibition of concern or interest in the working force might have had such a positive effect and productivity would have risen in any case.

Some consideration was given to the relative costs of work-oriented functional literacy, in the sense of a complete amalgam of literacy teaching and vocational training, as against simple literacy followed by specific vocational training programme. From the viewpoint of cost analysis, there appeared to be definite benefits in simple literacy programmes. Many more literates would thereby be available for particular technical training programmes. Such an analysis did not take into account, however, problems of motivation. It was apparent that motivation towards literacy was greater in highly selective situations and less, if literacy teaching was not related to meaningful situations. Simple literacy therefore could be expected to be highly wasteful. It was therefore clear that there were no simple conclusions to be drawn, for example, from consideration of cost effectiveness in isolation. It was one approach which should be taken into account with others.

Particular problems were seen in how cost effectiveness analyses were presented to governments. It was clear that such an analysis should include aspects which could not necessarily be presented in numerical terms. Even so there was a danger that decision-makers might be too readily attracted to the precisely-quantified aspects of such an analysis at the cost of ignoring the other aspects. A great deal was therefore dependent on the way in which cost analyses were presented. Judicious weight should be given to unquantified elements.

It was felt that, in the presentation of cost analyses as an aspect of evaluation, a need for a sense of order of magnitude was very valuable. Too much time should not be spent on small items of detail. The large costs were the manpower costs - the costs of the students and teachers involved in such programmes. Care should be taken to obtain as accurate a calculation of manpower costs as possible. For the other costs a rough approximation was probably enough. The manpower items could be expected to amount to 70-90% of the total costs.

In considering manpower costs, the employment situation was basic to any analysis. In the first place, it was important to consider whether the teachers would be alternatively employed. For the students, although in many rural situations in developing countries there was throughout the year a considerable degree of under-employment, this varied widely with the seasons. At a particular time, there was acute labour scarcity. During these seasons the individual participant might be faced with considerable personal conflicts regarding his adherence to the functional literacy programme.

Conclusion

The opportunity cost concept was concerned with the detailed calculation of the value of resources used. The main conclusions that it offered for the evaluation of functional literacy programmes were:

on the costing process itself, emphasising the inadequacy of simple accounting, although simple financial costs were also of importance to governments, particularly during times of severe budgetary constraint;

that costs could be related to effectiveness. When thorough measures for estimating gross costs had been carried out it might then be possible, for example, to consider how the same end could be produced with a more economic use of resources. Applied to education, this might mean variation in size of classes or alternatives to the timing of programmes, at particular times of the day or at particular seasons. These adjustments could lead to significant reductions in the opportunity cost of the programme.

In planning cost analysis in relation to functional literacy programmes a number of considerations must be borne in mind:

Cost calculations must take into account the entire economic structure, in particular the nature of the labour market and the wage situation.

Particular attention must be paid to the quantities of manpower involved. The detailed timings of this involvement should not be measured in monetary terms until a late stage in the entire calculation. A final estimate in monetary terms should be arrived at with the assistance of a labour economist using such devices as the shadow wage estimate.

Alternative estimates are often preferable to cost analysis giving a single simple answer. They could illustrate the range of costs to a variety of agencies at various levels.

It was essential in any analysis to have accurate time budgets as well as cost budgets, particularly with reference to students and teachers.

Raw data of this nature should be provided at an early stage in the operation of the project, not at the end of the project.

Discussion on the address by Dr. Jolly was deferred since the subject of cost-benefit analysis was to be considered by a working group.

## CHAPTER 7

## EXPERIENCES AND PROBLEMS IN EVALUATION - A SOCIOLOGIST'S VIEW

by

DR. K. NEYS

Dr. Neys reviewed the structure of the functional literacy programmes in Ecuador and in Venezuela, with particular emphasis on the evaluation procedures in these two countries. A particular problem facing all projects within the World Experimental Programme was that of comparability. For evaluation purposes it was necessary that projects should be similar. They could not, however, be identical because of the differing natures of the societies concerned. Similarity could be arrived at in the indicators by which impact of the literacy programme was measured, in the way in which results were presented and in the techniques of planning. The distinction between evaluation and assessment was stressed, evaluation requiring the application of scientific measurement procedures based upon objective data posed particular problems regarding personnel, time and finance. The structure of the literacy programme in Ecuador was examined as a useful illustration of this process. Although this programme was both functional and experimental more emphasis has been placed on the functional aspect so far.

Problems arising from this programme concerned the need to integrate support for the programme from various ministries. There were also human and technical problems, particularly in the evaluation process and also material problems largely concerned with equipment and teaching aids. Most progress had been made in the provision of appropriate teaching aids.

In Ecuador the general emphasis for the literacy programme was in vocational/technical activities, the immediate need being to make a contribution towards the solution of the prevalent problems of under-employment and unemployment. In contrast, the Venezuelan programme was more culturally oriented, although the vocational training of women and experiments in agricultural extension and co-operative organisation will receive due stress.

In this situation, what was the role of evaluation? The evaluators were faced with two demands. The first of these emanated from Unesco, which, through its Evaluation Panel sought to apply to projects within the World Experimental Programme, international norms for evaluation. The

Evaluation Panel was an independent unit, not itself part of the Unesco structure and therefore fulfilling the general requirement of evaluation, i.e. to be directed by an outside organisation. The role of the evaluation unit within the local project was to adapt these norms to the local situation. The second demand arose from the operational activities of the programme which, had severeal precise short term requirements geared to the needs of the programme, for example, action-research on teaching aids. Recognising these pressures it was of the greatest importance that the evaluation unit, at the earliest possible stage, establish first its general objectives in accordance with Unesco policy and secondly its operational objectives in accordance with the particular orientation of the local programme. At the last meeting of the Evaluation Panel in Paris, great stress had been laid on the need to define clearly operational objectives, as this had not been done in many projects. Without this thorough evaluation was not possible.

Internally the evaluation unit had direct linkages with the directorate of the literacy programme and with the operational activities of the programme. Externally it was necessary for the unit to have working links with those organisations which might be of assistance in its purpose, for example, universities, government statistical offices, various organisations involved in the development process. In particular organisations such as those could satisfy short-falls in the professional staffing of the research and evaluation unit. Moreover, another type of difficulty might arise if personnel were used, from the operational side of the programme, to make good those short-falls. Such personnel could not be expected to be completely unbiased in their approach to the project.

Particular problems faced by the research and evaluation unit concerned (a) evaluation and planning; and (b) methods and techniques. In the first case problems of comparability between projects were related to the way in which results were presented, because of the need for some uniformity of structure. In the second case problems centred on the selection of appropriate indicators for measurement in two dimensions, i.e. of the impact of the literacy programme on the environment and the analysis of the organisation of the programme, including such factors as cost/benefit, the phasing of the various processes within the project. So far the Ecuador evaluation had been mainly concerned with impact-analysis and had not ventured in any great degree into systems analysis.

Particular problems were also seen regarding the provision of basic data. It was felt that projects should provide basic data for evaluation simultaneously with the starting of operational activities. In many projects the work of the evaluators was being outstripped by that of the operators. It was essential, therefore, that evaluation be built into programmes from the very beginning and correlated with the operational side.

Particular difficulties were foreseen over the isolation within a given development process of the factor in this process which was contributed only by functional literacy: the independent variable. In many developing countries the agricultural population, for example, were already subjected to various influences which acted as an incentive towards development. The need was stressed for well-matched control groups and the continual survey of these groups, as well as for the use of certain techniques such as depth-interviews and life histories of learners.

Particular attention was given to the problems of selecting indicators. There was a need for a practical structured list of indicators; for example in the Venezuela project, 60 indicators had been determined upon, divided into areas of economic, social and cultural activity. Having determined upon these indicators, two sets of instruments were envisaged - one which would measure the immediate basic effects of the functional literacy programme, the other which would attempt to measure long-term changes which could not entirely be identified with the programme. Some doubts were expressed as to whether structural changes could be measured within the five year span of the World Experimental Programme project. The particular value of a wide range of indicators was seen in its potentiality for measuring the effects of the total range of activities of a Functional Literacy Project. The indicators should reach beyond the specific purposes of the evaluation of the immediate impact upon the living standard of the participants.

The stages of measurement were reviewed

The initial task was the establishment of a baseline survey. This could be carried out in various ways; the Ecuador experience contrasting to some degree to that of Venezuela. In Ecuador a research programme lasting 1½ years had been caried out to produce baseline data on a fairly scientific basis. In Venezuela it had been decided to short-circuit this process as far as possible. Three phases were envisaged in the Venezuela baseline survey:

initial stock-taking; (March - July, 1969)
planning; (Aug. - Sept. 1969)
the benchmark survey. (Oct. - Dec., 1969)

The stocktaking phase, lasting for four months is aimed at collecting data of local resources and problems with the objective of isolating 'situations,' that is, complexes of problems regarding which functional literacy might contribute towards a solution. 'Situations' revealed fell initially into three categories, social, economic and cultural, and gave rise to three problem charts of the area selected which should be combined into one master chart of problem-situations to which functional literacy activities should be attached. This process was essentially a quick survey, and depended heavily on local informants and was not fully scientific. Since, however, a full survey would take more than a year, it was regarded as an adequate means whereby problems upon which the programme might focus could be located. The techniques used in carrying out this phase were:

a series of interviews with selected informants;

documentary analysis of economic and social data, to be found in census data, specialised studies, etc.;

a questionnaire on social problems mailed to local informants selected institutionally;

observation visits by evaluation unit personnel.

In Ecuador apart from all this, lengthy field research on levels of living, including sample studies of local families was carried out by 100 part-time field personnel using an interview schedule.

Some doubts were expressed whether questionnaires were of much value in largely illiterate areas, particularly if these were mailed to the respondent. A difference was admitted if the questionnaire was placed in the hands of a trained interviewer. The relatively poor response to the questionnaire was regarded as an indication of this argument. Further doubts were expressed on the selection of local 'leaders' based on their institutions. This might lead to a biased view of local social problems being fed into the planning of the literacy programme. It was felt that any such bias could be corrected by observational visits.

Planning. In Venezuela the evaluation unit had participated in the planning process, whereas in Ecuador this had been left to those in operational activities. It was emphasised that the evaluation unit ought to take a strong hand in the planning of the operational side of the project.

Benchmark survey. This phase, lasting approximately three months, provided a more specific element to the base-line survery, producing data which could be specifically quantified and upon which scientific periodic evaluation could subsequently be accurately based:

samples of learners covered by an interview schedule providing data on the levels of living of individual families. It was recommended that a large initial sample should be selected in view of the high fall-out expectation of learners for the duration of the project;

samples of the communities. This would involve rapid survey-studies;

samples of organisations within the region, such as trade unions, co-operatives, factories.

## Units of analysis

The units of analysis recommended in the Unesco manual for evaluation were national, regional community/organisation and the individual family. In the Ecuador and Venezuela projects, particular emphasis had been placed upon the individual family and to a lesser extent on the regional unit. Activities impinging upon these units fell into two categories, (a) pilot activities; and (b) experimental activities. Pilot activities mainly involved assesment without scientific controls. Experimental units represented a relatively small proportion of the entire activities of the programme, but were subjected to rather rigorous controls. It was suggested that a third category existed which might be classified as 'incentive activities.' These were non-experimental, were under some slight observation and had the purpose of obtaining the backing of the local population for the programme generally.

## CHAPTER 8

## TECHNIQUES OF SOCIAL STUDY FOR EVALUATION

by

DR. R. WIENER

Since all the techniques that social scientists use are relevent to evaluation, Dr. Wiener faced the problem of condensing a whole year's social science course on methodology into the space of one hour. He proposed to follow the outline laid down by Dr. Neys for a base-line survey, with the three stages - stocktaking, planning and benchmark studies and to try to show the research methods which are relevent to each stage.

### Sampling

Turning first to the stocktaking survey, the first problem is one of sampling. Which regions, communities and households are we going to sample in order to stock pile our lists of problems and basic data? It is important to first define our terms. The term 'population' would be used in the technical sense not to mean the whole population of the country but all of the people to whom the study applied. For example, in the evaluation of this Workshop the population consists of all the people who have taken part in it. Any one member or group of members from the Workshop makes up a sample of the population. One draws samples because it is obviously cheaper to be able to obtain the information one requires from part, rather than all, of the population. In any sample that one draws one must try to make sure that it is as representative as possible of the population. The aim is to make the sample as accurate a predictor as possible of the population. The accuracy of a sample depends on three factors: homogeneity, size and cost. By homogeneity, we mean similarity. Suppose we consider this group as a group of people concerned with the evaluation of functional literacy, then any single person will be typical of the group as a whole, i.e. a sample of one would provide a good estimate of the population. Suppose, however, we consider this group in terms of the continents that people come from. Now one can see that in this case no single person would be representative of all of those continents. Now in the latter case when we have a heterogenous group there are two things that we can do to make a sample more representative: the first of these is to increase the size of the sample. For example, if we draw five people from the group then there is more chance that they will be representative of all the continents concerned. But obviously, the larger the sample becomes the more expensive it becomes to interview all of the people in it. To get round this problem we can stratify the sample. That is to divide it into a number of different homogenoous layers. For example, in this case we can divide the group into Asia, Africa and Latin America; thus we would have produced three homogeneous strata and we can then draw from random one member from each and this would obviously increase the accuracy of the sample as an estimate of the population. In all cases one should try to make sure that the individual people or units are selected on a random basis. By random one means that each person has an equal chance of being chosen. This is important not only because the majority of statistical tests depend upon this randomness but because otherwise one introduces a bias into the sample.

One should note that there is not necessarily a right or a wrong means of drawing a sample, designing a questionnaire and so on, there are various techniques that a person can use and often only experimentation will show which techniques are most appropriate.

Let us assume that our objectives and the time available do not enable us to draw a complete and detailed random sample at both this stage of the project and the next - the benchmark stage. However, even if our sample at this stage will not be as accurate a predictor of the population as the final sample we still want to make sure that it is representative of all the different sub-groups within the population which might have different problems. For example, we might expect towns to have different problems from villages, inland regions from coastal ones and so on. The best way of making sure that the sample represents all of those sub-groups is to stratify it. For example, we could divide the regions into inland and coastal ones and select at random one from each. Within each region we could divide the communities into industrial sectors, villeges and small towns and select at random one from each. If at the household level we believe that literate and illiterate people would have different problems and that middle and lower class families would also, we can then stratify the sample on this dimension and select one or two families at random from each.

## Interviewers: Selection and Training

The next step that one might have to do is to select one's interviewers. It is advisable to select more interviewers than one needs at this stage and for each interviewer one should obtain information on: physical make-up; educational attainment; general intelligence; special aptitudes; interests; disposition; circumstances. During the training and the pilot survey some interviewers will not prove as good as others and can then be dropped from the interviewing panel. If one now refers back to these seven points and then one compares how the interviewers who dropped out compared with those who did best, scored on these points, one will then be able to use the scores achieved by those who did best as selection criteria for any subsequent batch of interviewers required. In training interviewers one needs to stress the following points:

Rapport: this refers to the relationship between the interviewer and the respondent. It is important for the interviewer to introduce himself in such a way so as to make the respondent willing to co-operate.

Standardisation: the aim here is to make sure that each interviewer asks each question in exactly the same manner.

Coding: always try to make sure that the interviewers write down everything as the person says it so that the distortion effects of time do not interfere with his memory of the event.

Probing: a probe question is an additional question an interviewer asks in order to get further information. For example, if a question is, "Who comes to visit you in your house?" and the answer is, "My sister," a probe question could well be, "And anyone else?" A probe question should never suggest the answer. One should not have said, "But don't your grandparents come and visit you?" It is always important to make sure that the interviewers record all the probe questions that they use otherwise you will end up with a lot of answers without knowing to which questions they refer.

Practice: there is no substitute for getting your interviewers to practice not only on each other but also on the types of people that they will meet in the survey.

When finally choosing which interviewers are going to interview in what area one must keep in mind a number of points. Firstly, the content of the survey; if one is asking questions about child weaning it will obviously be better to have female interviewers. Secondly, if the respondents perceive the interviewer to be too different from them they might become suspicious and adjust their answers accordingly. As a general principle it is best if the interviewer interviews people from the same social and economic background as himself though he should not know the respondents personally.

## Interview design

We now move on to the design of the interviews that we are going to use to find out what the problems are. At this stage we are likely to want to use unstructured interviews and open questions rather than structured interviews and closed questions. In an unstructured interview one asks as few questions as possible and encourages the respondent to talk as much as possible in the belief that if this happens the person is more likely to provide a true picture of what he believes. In a structured interview all the questions are asked in a very precise order.

An example of an open question is "What do you do in your spare time?" while a closed version of this question would be "Which of the following things do you do in your spare time: church/books/listen to the radio, etc?" When we design questions for benchmark measurement we will be using structured interviews and closed questions. The advantage of using these techniques is first of all that they are easier to code. In the open question the interviewer has to write down the whole answer and then all of the answers have to be analysed to find common patterns while in the closed question all the interviewer has to do is to place a tick in the appropriate columns. We will come back to interview design when we consider benchmark measurement.

## Analysis of documents

At this stage also one will have to analyse documents which will mainly be reports from ministries. In this case it is important to check on the reliability and validity of the documents. Reliability relates to the consistency of the document. One has to make sure, for example, that the document does not say in one part that 20% of all farmers are illiterate and in another part that 40% of all farmers are. If such is the case the document is unreliable; one cannot place much faith in it and it will have little validity. Validity asks the question of whether something is measuring what it says it is. If the document says that all literate farmers are self-confident one should ask: how was this statement ascertained; what measuring techniques were used to measure self-confidence etc? We will again come back to these terms of reliability and validity as they are two of the more crucial concepts in measurement.

## Observation

The next possible stage is to carry out observations of communities to supplement the information one has gained from the interviews. There are two stages when one is likely to use an observation (sometimes known as process study or participant observation study). The first is appropriate at this stage when it will provide additional information about problems and the second in between the base line, interim and final surveys. Observation studies enable us firstly to supplement quantifiable information and secondly, enable us to obtain more qualitative information which cannot be obtained by quantifiable techniques. In any observation there are three questions to be asked:

what should be observed?

how should the observation be recorded?

what relationships should exist between observer and observed?

As regards the stocktaking survey, one will be interested in observing social, economic and cultural factors. As regards social factors, one will want to note information like the amount of drunkenness; in economic factors, the amount of poverty; in cultural, the amount of radios and books in use. An observer in this case ought to have a list of all those different points and to write in opposite each one what he sees. In the stage after the base line survey the observer has to observe the inter-relationships between these indicators. For example, suppose the final survey shows that participants drink less than they did before the campaign was introduced, one would want to know whether this was due to the fact that participants worked longer or due to the fact that they read more, listened to the radio more, that drink cost more, etc. The observation study will enable one to gain some understanding of this on-going inter-action between indicators.

There are two ways that an observer can inter-act with a community. He can either call occasionally or he can become a full-time member of the community: there are problems with both approaches. Firstly, if the observer calls on only some occasions it is likely that he will obtain only a biased sample of the community's behaviour. One way to get round this is to treat everything that the community does as the total population of events that one is interested in and then the aim of the observer is to sample this population and thus one can see the principles of sampling come back in to make sure that he obtains as accurate a sample as possible. The observer should come as frequently as possible, his visits should be on a random basis and he could well stratify what he observes for example, in terms of market place, household activities. If the observer lives in the community on a full-time basis then he has the problem of how identified with the community he should become. If he remains very much an observer then the community is likely to remain suspicious of him and limit what he is allowed to observe. If, on the other hand, he fully partakes of the community's activities there is a danger that, firstly, he will lose his impartiality, and secondly, that he will find himself allied to one faction within the village and thereby cut himself off from observing others.

In observation one is still concerned with problems of validity and reliability. To check validity it will be possible, in some cases, to refer to other findings. For example, if the observer finds that there is a decline in drunkenness then it should be possible to check police or court figures. As regards reliability, one can always send a second observer along with identical instructions and then compare the reports that both bring back.

## Planning the evaluation programme

Let us now presume that we have finished the stocktaking stage; we move to the planning stage. Our basic problem is to evaluate the effect of the independent variable, functional literacy, on a number of dependent variables or indicators. An independent variable is that which the experimentor introduces into the situation. The dependent variables or indicators are those things that we expect to change as a result of the introduction of the independent variable. The problem of research design is to make sure that any changes found in the dependent variables are due solely to the functional literacy campaign and not to some other factor.

## The defects of retrospective evaluation

Suppose that we had introduced a campaign into one village which aimed at functional literacy, improvement of agricultural production, nutrition and sanitation. If we measured the performance of the participants in the village at the end of three years and found that they achieved certain standards, could we then say that the campaign had been a success? The answer obviously is 'no' because we do not know how well the villagers would have done without the functional literacy campaign. Suppose then we examine another village which has had no functional literacy campaign to see how well the inhabitants in that village do on the test. Suppose that we found that they achieved lower standards, could we then say the functional literacy campaign had been a success? Again, the answer is 'no'. It might well have been the case that these differences existed before the campaign was introduced. Another possible explanation of these differences might have been due to what is called the Hawthorn effect. This refers to findings made in an industrial study in America where experimentors introduced a number of planned changes and found that production increased. They then found, much to their astonishment, that irrespective of what changes were introduced production still improved. The reason for this turned out to be that the workers were responding to the fact that an interest was being taken in them. Therefore, if one finds increased productivity in a village one must make sure that this is due to the functional literacy campaign and not to the fact that the villagers have become more interested because of the arrival of the operators. This particular design, involving only retrospective evaluation, cannot isolate what is due to the functional literacy and what might be due to other effects. Another difficulty of this design is that even if a difference is found and is suspected to be due to the functional literacy campaign it does not permit one to work out why the changes occurred.

## The need for before-and-after measurement and for control groups

It therefore appears that we need to have a before-and-after measurement, i.e. a base line and final survey. But suppose we take our one village again and measure people's performance before the campaign is introduced and at the end of three or five years. If we find an improvement can we say that this was due to the functional literacy, again the answer is 'no'. Firstly, we have not accounted for any Hawthorn effects; secondly, there is the problem that we cannot isolate what changes are due to the functional literacy campaign and what changes to other factors, like radio broadcasts, being contributed. What we require is a control group. A control group enables you to say: "This is what would have happened in the experimental village if the functional literacy campaign had not been introduced." That is we have an experimental village and a control village and we measure both villages on the indicators/dependent variables before we introduce the functional literacy campaign and we measure them both again at the end of the campaign; then we can say that any differences found between them are due to the functional literacy campaign. In fact it is not quite as simple as this because we have to make sure that the control and experimental villages are matched on all the appropriate variables. Suppose, for example, the experimental village was in tribe area A and the control village was in tribe area B, then even if one found that the experimental village increased their productivity it would be possible to argue that this would have happened anyway because tribe A was a more hard-working tribe. We have to match the experimental and the control villages on all the variables which might influence the functional literacy campaign. A second point is that we must make sure that during the campaign any factor which is introduced into either the control or the experimental village, for example, television, is introduced into both of them otherwise someone could always argue that it was the television rather than the functional literacy which produced any changes. It is important to select a control group at the community level of analysis, and the participant level of analysis.

## Comparative evaluation of different methods

The design we have sketched so far is a very simple experimental design. Suppose, however, we want to compare different types of functional literacy campaigns, for example, a campaign which used radio and one which used television. We would then have to have an additional experimental group. Now suppose that action research was going to be incorporated into the radio campaign and we wanted to evaluate its effect, then we would need another group so that when the feed-back from the action research was incorporated into the radio programme in one group this did not affect the other one. Thus we would have to have:

Experimental Radio Village One

Experimental Radio Village Two

Experimental Television Village

Control Village.

If the action research is fed back into the first village then the second radio village functions as a control group for the first one, because it shows what the effects of the radio programme would have been if one had not introduced the results of the action research. One also has to make sure that all of these groups are matched.

## Evaluation of effects on different groups

If we are interested in comparing the effects of the campaign on different groups of individuals we might use a factorial design. Suppose we wanted to compare the influence of age, religion and social class on the functional literacy campaign. Let us say that we were interested in comparing people who were aged over 40 and those under 40, people of religion A and religion B and people of middle and lower social class. We would now have what is called a 2 x 2 x 2 design. That is we have eight cells:

Cell 1 - People aged over 40
People of religion A
People of middle social class.

Cell 2 - People under 40
People of religion A
People of lower social class

Cell 3 - People over 40
People of religion B
People of middle social class

Cell 4 - People under 40
People of religion B
People of lower social class

Cell 5 - People over 40
People of religion A
People of middle social class

Cell 6 - People under 40
People of religion A
People of middle social class

Cell 7 - People over 40
People of religion B
People of lower social class

Cell 8 - People under 40
People of religion B
People of middle social class

With such an analysis we can work out the effects of each of the variables. Suppose that we had taken sample people aged over 40 and people aged under 40 and compared them. Now in this case it might well have been that all the over 40s came from religion A and all the under 40s from religion B and thus even though one thought one was getting the effects of age one might well have been obtaining the effects of religion. One can only separate out these effects by the use of a factorial design. In a factorial design one must make sure that one has sufficient people in each of the cells in order to be able to carry out the analysis. There are two ways of doing this. One can either make sure that the sample is stratified along this particular dimension or one can draw a quota sample either as part of the main sample or in addition to it. In a quota sample one would simply tell interviewers to find 20 people to fit into cell 1 or cell 2. Obviously, this method is not strictly random.

There are a number of other sampling techniques but one of the most relevant ones is what is called area or cluster sampling. Suppose that we were sampling households in a town, one way would be to make a list of all the households in the town and then draw a number of them at random. This would mean, however, that they would be scattered everywhere. What we can do is to put in an additional step where we divide the town into a number of areas or clusters of households and then draw a random sample of those. This means that all the households will be in the one area.

Benchmark surveys: the use of interviews

Now that we have planned the study, we are at the stage of designing benchmarks. There are a large number of psychological assessment techniques that we can use. Let us first consider interviews. We have already said that what we want here are structured interviews and closed questions. To design a closed question one normally asks it first as an open question, obtains all of the alternatives and then designs categories for them. If one simply sits in an office and tries to make up closed questions it might well be the case that one leaves out options that are very popular in some parts of the country. When one has designed a number of open questions one should then go and try them out or pretest or pilot them.

Piloting is a very important stage. In a pilot one takes a small but representative sample of the population. The pilot serves a number of functions:

a. to try out the sampling frame;
b. to discover the homogenity of the samples;
c. to try out the open questions;
d. to get some idea of the types of answers that one will obtain and how useful they might be;
e. to test the efficiency of the interviewers;
f. to try out and design coding frames;
g. to check that all the questions can be understood by all levels of respondents in all parts of the sample.

When one has the answers to the open questions one then designs a coding frame for each question remembering:

i. that any set of categories should be exhaustive, that is that all responses can be classified;
ii. that any set of categories should contain mutually exclusive items so that a given response cannot be placed into more than one category.

When one has closed all of the questions and, if there is time, one should again pilot these questions to make sure that they work.

There are many pitfalls in questionnaire design into which there is no time to go. For example, one should not use words like 'often', i.e. the question "Do you often use the library?" is a meaningless question.

Reliability and validity

We must now consider concepts of reliability and validity in more detail. To take reliability first: as we said before, reliability refers to the consistency of a measuring instrument. There are two basic sorts of reliability within a test, that is are all the items measuring the same thing and secondly, reliability over time, does the same test produce the same result when given to the same people on two different occasions? Some cases of unreliability are poorly designed scales, tests which do not cover completely the subject being measured or tests with ambiguous or unclear items. To measure reliability within a test one can use a split half reliability coefficient. Here one divides the test into two parts and correlates the scores obtained on each part. If the test is measuring the same thing throughout then the two halves should be highly correlative. One way of making sure that a scale or test is doing this is to carry out an item analysis where one correlates the score on each item with the total score obtained by the individuals and then one retains only those items which correlate highly with the total score. As far as test reliability over time is concerned we can administer the test on two separate occasions to the same group of people.

Let us move now to validity which asks the question of whether the test measures what it sets out to measure. There are a number of different ways of measuring validity. Firstly, there is what is called face validity - does the test look like it is measuring what it says it is? This is not a particularly useful test of validity as some tests, particularly those of personality, are so designed as not to appear to be measuring what they measure. Another type of validity is content validity - does the test cover all of the area that it is meant to be covering? This is useful, particularly for assessment tests, where one can see whether a test of arithmetic, for example, covers all the mathematics taught on a course. Another type is criterion or empirical validity. For example, if we wish to validate a mathematics test we might well expect that people who did best would also be those who did best at mathematics in schools. School performance would then be the external criterion. On factual questions, such as whether anybody has been in trouble with the police or not, one can often refer to records of police or courts as a criterion. Then there is construct validity, in this case one argues that if the test is measuring what it says it is then one would expect, for example, that it would correlate highly with other tests measuring the same thing, that people who did well on it would have a prior history of success in that field, and so on. Finally, there is predictive validity. In this case, one predicts some future behaviour of a person on the basis of his test score. One point in relation to attitude scales, even though many people use predictive validity to assess the validity of attitude scales, this seems to be inappropriate because it presupposes that there is a one-to-one correlation between attitudes and behaviour.

## Assessment Techniques

Let us very briefly look at a number of different assessment techniques:

### Achievement tests

On indicators, like productivity, literacy and so on, we need a test which can be applied both before and after the campaign is introduced so that we can see to what extent performance has increased. In designing such a test one must keep in mind the content of the course so that the test adequately reflects what the course teaches.

### Attitude scales

A lot of fuss is made about changing attitudes, but it is important to remember that one is not really interested in changing attitudes but in changing behaviour and that the relevance of measuring attitudes is only proportional to the extent to which they serve as predictors of behavioural changes.

There are many different types of attitude scales, the easiest to use probably being the Likert scale. In designing a Likert scale we obtain a whole series of things about a topic, give them to a representative sample of the population, calculate a score for each person, then take the top 25% and the lowest 25%, see which of the items differentiate best between these two groups and include them in the final score.

### Sociometric techniques

Sociometric techniques are useful for finding the pattern of social inter-action in a community. A sociometric question normally asks one to select two or three other people on some criteria, for example, "Which two people in the village are your closest friends?" By giving this question to all of the people within a community at the beginning and end of a campaign one could then see if there were any differences in the pattern of inter-action within the community.

## Interest scales

We might want to measure interests either as a predictor of success or to see if a person's interests changed as a result of a functional literacy campaign. There are two different types of interest questions. One, which establishes the areas in which a person is interested and a second one which is concerned with the amount of interest a person has in a particular area. This, of course, leads straight on to the whole problem of

## Motivation

Motivation is not a simple concept. Firstly one wants to know how much a person wants to do something but if you know the quantity of motivation this does not mean that you know the causes of it. That is, a man might want to produce more because:

a. he wants to earn more money;
b. because everyone else is;
c. because his wife wants him to;
d. because he wants to please the functional literacy operators, and so on.

It is important to possess this additional information, firstly so that one can adjust the appeal of the campaign to the various needs of different sections of the community, secondly, so that we can differentially adjust the pattern of rewards within the campaign and thirdly, so that we can gain some insight into why people drop out of the campaign.

## DISCUSSION ARISING FROM DR. WEINER'S ADDRESS

### Reliability/validity of interview schedules

If structured interviews were reliable a consistent pattern of answers ought to be produced, despite time lags in application, provided that the units sampled were matched. If interviewers in different districts produced widely different returns the explanation may be either that there are real differences between the groups sampled or personal differences in the interviewers themselves. It was regarded as important that checking procedures on the interviewers should be built into questionnaires. Alternatively follow-up interviews might be carried out by other interviewers. Particular problems arise regarding supplementary probe questions. Care should be taken by interviewers to write in to the questionnaires all supplementary questions.

### Motivation analysis

Doubts were expressed over the various meanings of this phrase, the uses of which ranged from an individual's internal psychological state to the material incentives driving individuals or groups, for example the need for literacy. In attempting an analysis of such a concept it was regarded as important that motivation should be broken down into individual compartments, for example financial motivation, social motivation, etc. Some doubts were also expressed on the usefulness of motivation study in connection with the evaluation of functional literacy programmes, although it was pointed out that an explanation of drop-out from many schemes was that motivation on the part of the participants had been assumed rather than proved. Motivation was regarded as a researchable problem and could be tackled by the careful study of differences in attitudes between drop-outs from schemes and those who remained in them. In addition, motivation study might be useful in selling the project to particular participants if the factors which would attract participants to the scheme were clearly isolated and publicised. In addition, such studies had a value in underlining the fact that all problems were multi-causal, multi-dimensional. To isolate a single factor and attribute to it sole responsibility for 'solving' a problem was an out-dated notion.

### The need for statistical skills

With the acceptance of the contribution of many factors to, say, development, the importance in a research and evaluation team of competent statisticians was underlined. Statisticians were necessary to isolate predominant factors in particular problems

## Socio-metric techniques

Similar problems were seen with regard to the analysis of attitudes and behaviour. There were socio-metric techniques which would analyse inter-personal relationships within a group. The role for such techniques in the evaluation of functional literacy might be to prove a hypothesis that those who progress in functional literacy programmes will change in their social relationships within their group. Similarly, in the planning of a functional literacy programme, such techniques might be useful in locating structures and factors within a particular group or groups at which a functional literacy programme might be aimed. Within an intensive and selective strategy it was important to locate the innovators within a group and aim the programme at these. The value of these techniques in planning lay in the fact that they produced quantifiable results fairly quickly, more quickly than anthropological studies of communities in depth.

## Anthropological studies

Ideally much value was seen, however, in anthropological studies which gave a more intimate understanding of the local community and culture. The problem was that such studies were expensive and time-consuming. Socio-metric techniques had therefore a particular value in short-circuiting such procedures, although some concern was felt that these techniques might produce, in a project planned solely as a result of these, a lack of sensitivity to the totality of a culture. Consideration might be given to the building in to the preparatory research phase of a proposed functional literacy programme of an element of anthropological studies. Doubts, however, were expressed as to the practicability of these measures, however valuable, both from the point of view of the shortage of personnel and of funds.

## CHAPTER 9

## THE EVALUATION OF RURAL EXTENSION PROGRAMMES

by

DR. J. DI FRANCO

### The role of extension and the contribution of literacy

Extension is a part of the general educational process. Most of the basic concepts of education are fundamental to the work of extension. The main differences between extension and formal education centre on the fact that participation in extension programmes is essentially voluntary. Participants are, in the main, adults with a wide range of needs and interests. Curricula are flexible and in the main problem-oriented.

The role of extension is, therefore, to tackle, through out-of-school education programmes, particular educational problems centring largely on the motivation of adults in rural areas towards greater agricultural productivity and the provision of the necessary technical skills. The problem of widespread adult illiteracy in many countries is a fundamental obstacle inhibiting the attainment of this general end. The teaching of literacy, enabling the reading process to begin, opens their horizon and develops a questioning attitude on the part of new literates, producing a dissatisfaction with their conditions of living. It is this dissatisfaction that extension programmes should exploit; helping the rural population through educational means to improve their skills and knowledge.

### Mistrust of evaluation and means of overcoming it

One of the main problems facing educators generally and particularly those involved in extension programmes is the fact that their objectives are often not fully understood by those responsible for the overall administration and supervision of such programmes. Evaluation of extension is a particularly uncertain area. Many extensionists are reluctant to expose their programmes to evaluation for fear of incurring criticism which might reflect upon themselves. There is a need to popularise the idea of evaluation as a device available to those engaged in extension which will indicate the state of their programmes. Evaluation procedures enable those engaged in programmes to measure the capability of their own organisation and of the group with which they are concerned. If constructive extension programmes are to be devised, regular evaluation must be a fundamental element.

The traditional view of extension programmes envisages a sequential arrangement, the organisation of the programme being followed by the operational stage which in turn is followed by evaluation. This is an over-simplified and unrealistic view of the contribution that evaluation can make to an extension programme. Evaluation procedures detached from the programme as a whole are merely academic exercises. Only when the purpose of evaluation is to bring about practical improvements in the programme is it of real importance. The aim of evaluation should be to make the programme more meaningful. It is important to maintain a distinction between regular evaluation arrangements within a programme and evaluation studies designed to assess the overall impact of a programme. The requirement to assess impact is often politically motivated. The near certainty that operational programmes would not be able to prove impact to the degree suggested in programme objectives (which are in any case often determined by political as opposed to technical reasons) mean that operational extensionists often fight shy of the need to evaluate.

This general situation is apparent in the case of the various extension services which were launched in Latin American countries during the early 1950s. After approximately ten years of operation, that is in the early 1960s, it was decided that a series of evaluations should be carried out. The original intention had been to set up a series of studies which would analyse impact. It was, however, decided eventually that such studies would only prove the obvious, that the impact of extension programmes over the decade of operation had not lived up to expectations. Such studies would, therefore, only contribute to the curtailment of extension programmes. It was instead decided to institute evaluation arrangements which would seek to locate why programmes had not been able to achieve the original expectations and which would suggest how these programmes might be improved. The role of evaluation, therefore, was to strengthen, not to undermine, extension.

## The evaluation of organisational aspects of extension

In these arrangements particular attention was given to a detailed analysis of the organisational aspects of extension programmes. One feature in a particular country, which was discovered over the ten year period of operations, was that the average time of extension officers serving in one particular location was 13 months. It was quite clear that although the organisation was 10 years old the programme was one year old. No one officer had completed a full cycle of work at any one location. Another organisational feature which emerged in this analysis, which contributed towards the failure of these programmes to achieve their targets, was the location of extension offices, scattered throughout the country at varying distances from the centre, and the policy of the department of using favourable postings as a system of rewards for long or efficient service. This meant inevitably that the more experienced and effective extension officers tended to be posted near the capital city to the neglect of outlying areas. These 'best' officers in effect received direct supervision; the poorer or newest officers received little or no supervision. A solution which was suggested by this evaluation was that outlying offices should be closed and that the programme should aim at achieving successes in limited situations, proving the effectiveness of the general extension formula, and thereafter bringing the scope of extension programmes outwards as the organisation developed the capacity to expand. Another alternative was to have a higher pay scale for distant posts to attract the better qualified officers who needed less supervision.

This experience underlined the need for continuous evaluation. Evaluation should be an activity built into the total programme, affecting the organisational side as well as the operational side of the programme. Evaluation should act as a means for strengthening the hand of those in charge of programmes and not a threat from outside.

On the operational side, it was important to realise that if the quality of instructors was low, then high quality results could not be expected. This is particularly important in functional literacy programmes, many of which use elementary school teachers with low academic qualifications. Regular evaluation would bring to the attention of administrators that they could not expect high levels of literacy as a result.

Other features revealed, by continuing evaluation illustrated, for example, the attitudes of experts. Often, experts, whether national or international, are not able to operate successfully at the project level because of their failure to descend to the level of the participants themselves. It was important to realise that the effective functioning of the programme is fundamental and that there is nothing degrading, as far as experts are concerned, in getting down to participant level if, by doing so, the programme works more effectively. If communication between participants and experts is not achieved, then the impact of an extension programme could not be expected to be good. Individuals with high levels of academic training should be able to communicate with all levels of participants. Unfortunately too often professors tend to talk for other professors, but in extension programmes it is imperative that the rural community level be kept in mind.

Whether programme objectives were imposed on the operators for political or social reasons, the importance of evaluation arrangements as a mechanism for defending the operators was further underlined. If objectives are unrealistic, (or too idealistic), it is important that the extensionists themselves should clarify the objectives in realistic terms before attempting to evaluate the results. Extensionists should take the initiative in evaluation exercises. They should not wait until they are told to do so. Evaluation is a fundamental and continuing part of extension programmes. Its role is to determine the degree of success of programmes but more so, it should be to enable immediate location and consequent rectification of faults and weaknesses which would result in constructive changes in organisation and programming, so that ultimately there would be no doubt that a positive impact might be achieved. As for as costs and benefits are concerned, provided that it could be proved that an extension programme was having a positive impact, questions of costs and benefits would be less likely to arise. Such questions only tended to arise when a positive impact could not be proved.

Evaluation should be seen as the programme operator's controlling device, used for the improvement of his programme and a defence mechanism for his programme against over-ambitious expectations from outside. Secondly, for the extensionist, evaluation should be a matter of personal belief in a device which was immediately useful. It should not be an uncomfortable procedure which had to be tolerated because theory said so. The educational process called extension includes a continuous and built in evaluation procedure.

## CHAPTER 10

## VISITS TO THE OVERSEAS VISUAL AIDS CENTRE((OVAC) AND THE CENTRE FOR EDUCATIONAL TELEVISION OVERSEAS (CETO) AND ADDRESS BY MR. J. B. BOWERS ON ACTION RESEARCH IN THE PRODUCTION OF MEDIA FOR FUNCTIONAL LITERACY

The importance of active media for functional literacy was a recurrent theme of the Workshop. Accordingly, the Workshop programme provided the opportunity for participants to pay working visits to two Organisations which are directly concerned with training in the preparation and use of various educational media in all areas of education.

At OVAC, a joint session was organised with participants in OVAC's current course for senior students from overseas to hear and discuss a lecture by Mr. J. B. Bowers on Action Research in the Production of Media for Functional Literacy. Participants were also able to look at OVAC's exhibition of audio-visual aids and to discuss with OVAC staff the uses of audio-visual media in functional literacy programmes. The CETO visit was similar in purpose; the work of this organisation was discussed with particular reference to the use of television in functional literacy and the role of action research and evaluation. Participants were given a live demonstration of the production of an educational television programme in a studio especially equipped with low-cost apparatus.

## ACTION RESEARCH IN THE PRODUCTION OF MEDIA FOR FUNCTIONAL LITERACY

by

Mr. J. B. Bowers

### Definition of terms

It may be useful to distinguish between three types of research activity, which should not, however, be sharply differentiated:

- 'fundamental' or 'basic' research, intended to test hypotheses and to establish general principles, (e.g. research into perception of images by illiterate adults);

- 'action research' (or 'action-oriented research') - systematic study and investigation carried out as part of an operational programme, either by the operational staff or by research workers, in the latter case the results being fed back immediately to the operational staff to help them to improve their day-to-day work;
- 'evaluation' or measurement of the effectiveness of a programme or of methods or media used in it.

It is with the second of these that this talk is concerned. The term 'media' will be used in a general sense to mean all kinds of instruments ('hardware' and 'software') used for the communication of ideas, information and skills to 'target audiences'. This will include books and other written and printed materials, audio-visual aids and educational radio and television programmes.

The importance of media in functional literacy

Media are particularly important in functional literacy for a number of reasons. In formal education in the schools, the curriculum is generally well established, reinforced by well tried textbooks and visual aids and related to an examination system; in adult education, however, the content is potentially unlimited and needs to be defined. To borrow a metaphor from the food industry the raw material must be processed and packaged before it can be delivered to the consumers. 'Processing' involves action research to investigate the needs, problems, wishes and capacities of the consumers and to select the 'message' accordingly; 'packaging' involves embodying the message in appropriate media. Some media will be more appropriate to convey the skills of reading and writing and calculation, others to communicate useful knowledge or to assist in developing skills. In all cases, however, functional media must be adapted to the needs of each particular cultural, linguistic and occupational group.

Production practices

Media for informal adult education and extension have been, and are being, produced in many parts of the world by a variety of teams, units and individuals. Procedures vary widely. Sometimes the teacher, extension worker or instructor is expected, and perhaps trained, to produce his own visual aids; sometimes a specialist is commissioned to write manuals, and an artist or photographer to illustrate them; sometimes a film unit or cameraman is sent into the farm, factory or training centre to produce an instructional film. Some preparatory planning and investigation inevitably takes place; specialists are consulted, projects visited, books studied, but in most cases production relies largely on the experience and intuition of the media makers. It is essentially 'one-way' communication. Writers sit in offices and write; artists sit in studios and draw; broadcasters sit behind microphones and talk; and the results are projected upon the 'target audience'. Attempts are sometimes made to pre-test media, or to improve radio programmes by 'feedback' from listening groups, but those procedures are seldom scientifically planned as an integral part of the production programme. No wonder the messages are so often confused or misleading and the media defective in communicating them. The situation is naturally worse where there is a multiplicity of languages and a high level of illiteracy.

## More professional production methods: a plea for action research

The movement to functional literacy is beginning to change this situation, for it demands functional media and these cannot be provided by haphazard and intuitive production methods. In many functional literacy projects media are now being planned by combined teams of 'content specialists' (in agriculture, crafts and industrial processes, nutrition, health and other fields) and 'educators' or 'communicators'. In some cases a serious attempt is made to involve the illiterates themselves in the process, by various methods of ensuring 'feed-back', but this is still not very systematic. The object of this talk, therefore, is to put the case for more professional production based on more systematic action research. How can this be achieved? First, any important functional literacy or informal adult education or extension programme should be supported by a professional media production service, with one or more production teams operating from an adequately equipped centre. Secondly, the various media from books to visual aids, from films to recorded radio programmes, should be regarded as inter-dependent - as components in a learning system - and should, as far as practicable be planned and produced together, by co-ordinated team work, in the same centre. Thirdly, the media production service should incorporate action research into its operations.

## Who should carry out action research?

Action research can, of course, be done on a modest scale by the ordinary staff - educators, writers, artists - in the media team, especially if they can obtain some training and guidance in social survey methods, and even this will certainly help them to improve the quality of their production. If, however, it is to be really serious and systematic, and to make a major contribution to the improvement of media, it is a full-time job for a full-time member of the media production service. An action research worker should, then, take his place in the team with the educators, writers, artists, radio programme and film producers. He should be either an educator with some training in the social sciences and experience of social survey methods, or a social scientist with some experience in education. Unfortunately such people are not easy to find in a developing country. One way of securing their services may be to appeal to a University to assign to the media production staff one or more adaptable graduate students, with supervision from the University Social Science Department. Another possibility may be to seek international technical assistance to send in the right person. In a major functional literacy project he might be assigned part-time or full-time from the evaluation and research unit to the media team. However, in view of the heavy task of evaluation falling on this unit, it seems unlikely that a qualified person would be liberated for duty with the media service. Hence, in view of the importance attached to systematic and continuous action research in media production and the fact that it differs in its aims and techniques from programme evaluation, a strong plea is made for the inclusion of a full-time action research worker in the media service. He should be responsible entirely to the head of the media service but have good working relations with the evaluation unit.

Although this may seem an extravagant suggestion, it is less extravagant in cost/benefit terms to employ an action research worker than to produce media which fail to communicate their message or, worse still, communicate invalid or inapplicable messages. It is also less extravagant to employ a trained research worker than to engage the time of writers, artists and other production specialists in action research, for which they are not fully qualified.

How then should action research operate in the production of media for functional literacy and what will be the task of the action research worker in the production team?

## Problem survey

In functional literacy programmes the communication of ideas, useful knowledge and skills is a major aim and literacy is a means to it. But what ideas, knowledge and skills? This may be determined by a survey of the society and environment, sometimes called a 'problem survey', since it is largely designed to discover the problems which handicap development in the area. This problem survey is a form of action research, since it feeds back information to the programme planners. It is part of the basic-survey-stocktaking process. It will help the 'planners' to determine the aims of the programme and amongst other things to decide what media they require and for what purposes. Incidentally it is to be hoped that the head of the media service is included in the project planning group. This initial problem survey will not, however, be carried out by the action research worker in the media team, since the media team, or teams, will not come into action until their service receives a request for media.

When such a request is made, discussions between the project staff and the media team should take place to determine: what 'messages' are to be communicated (i.e. what content) by what media, to what target audiences.

Where the content or 'topic' to be covered is of a technical nature, which is usually the case, the project staff should assign to the media team one or more 'topic specialists' to provide technical advice. These specialists may be university people, members of government technical services, skilled workers in industry, craftsmen or experienced field workers. As specialists in their subjects they generally have a higher academic education and, sad to say, are often out of touch with the population of less educated and even illiterate adults, who are likely to be the target audience for the educational media. These critical remarks may not, of course, be true where the media team is working with the staff of an extension service or with field workers with real local experience. Even so, the production of media generally involves adapting the ideas of the specialists to the abilities of the target audience to understand, absorb and apply them. It is for this reason, and at the preliminary planning stage, that action research intervenes in the production process with a systematic study of the subject matter, which may be conveniently called 'Content Study' or 'Topic Study'. (1)

---

(1) It is sometimes called a 'Problem Study' which should be distinguished from a 'Problem Survey'. The latter is a survey to discover and identify problems, the former is a study of one single problem in depth.

## Topic Study

The purpose of topic study, in jargon terms, is to 'validate the message' before it is incorporated in the media - to make sure that:

- any facts stated are correct;
- any ideas put across are sound;
- any action proposed:
    - can be carried out by the target audience:
        - with the skills they possess or can acquire,
        - in the conditions in which they live (e.g. climate, communications, housing, working conditions),
        - with the funds and equipment, capital and resources, that they have or can obtain:
    - will have the desired results in the given situation;
    - does not seriously conflict with social customs or taboos or personal beliefs and attitudes;
    - does not tell people what they already know.

It is the task of the media team to assemble the relevant knowledge about the topic, to translate it into language and visual symbols that can be understood and accepted by the target audience and applied to their daily lives. Topic study is a technique of action research which helps them to do this more efficiently. It involves, on the one hand, extracting knowledge from the topic specialists, from books and other documentation, and from direct observation, and, on the other hand, finding out what the target audience do, know, think and believe about the topic.

How then does this differ from the ordinary preparation and planning that goes into traditional media production? In the first place, it is more systematic and aims to discover valid and objective data rather than relying on general experience, subjective opinions and intuition. In the second place, it involves 'feedback' from the target audience, even before production begins.

Topic Study involves a series of procedures:

- interviews, correspondence and discussions with the topic specialists;
- study of documents (reports, plans, projects, books on the topic, media, including those produced in other countries and other programmes, films, recorded broadcasts, and any material which will provide useful information and knowledge);

- field study (visits to projects, for example, factories, firms, farms etc. where activities relevant to the topic are going on and can be inspected);
- 'opinion study'; - by interview; - by discussion groups.

Opinion study is a vital element that is seldom included in the production of media. Its purpose is to determine what the target audience do, know, think and believe about the topic, before telling them what they should do and know.

These procedures are not necessarily carried out in the order stated and may overlap or interlock. For example, it may be necessary to have further interviews and discussions with topic specialists after the opinion study has revealed the current practices and opinions of the target audience. Again field studies may have to be undertaken at several points, as the team needs to inform itself more fully on what is being done.

If these procedures are systematically and seriously carried out they involve full-time work for a competent action-research worker. The planning and execution of opinion studies for example, requires some knowledge of social survey techniques, sampling procedures, interviewing and the selection and briefing of interviewers and the analysis and presentation of data, which only a trained person can be expected to possess.

### Job analysis and performance analysis

If the production team is preparing training media for a vocational training project, action research to validate content may take a rather different form. Instead of the topic study described above, research may rather be needed:

- to analyse the skills required in specific jobs (both manual skills and other skills, such as speaking and understanding the working language, reading, writing and calculation, the interpretation of drawings and diagrams);
- to discover the skills and abilities already possessed by the trainees and their level of performance, if they are already on the job.

Eventually the training programme must bridge the gap between the second and the first.

## Pre-testing

The next point at which action research enters the production programme is after the media have been drafted, the texts written and the first sketches or photographs made for illustrations or visual media. This is the point at which the draft media should be pre-tested. Pre-testing is designed to ensure that the media will communicate their message as effectively as possible to the people for whom they are intended - the 'target audience'. Too often pre-testing has consisted of asking a few people to read the text or look at the picture and tell the writer or artist what they think of it. Even this slapdash procedure may discover faults and deficiences which can be remedied before the text is made final or the pictures reproduced, but how much more effective is systematic pre-testing conducted by the more scientific procedures of action research.

The aims of pre-testing must be clearly defined in advance. At the simplest they are to detect and remedy any defects in the draft media and to improve their communication; but it may also be desired - and will certainly be useful - to pinpoint defects, for example to discover:

- what pictures or parts of a picture cannot be identified and understood or are misunderstood;
- why people find them defective;
- how serious are the defects (e.g. what percentage of people make wrong responses), what are the reasons for failures (e.g. a sick child looks well and happy).

The techniques of pre-testing naturally vary with the different media. The testing of written material can be done in at least three ways. The first is to ask individual members of the 'test audience' (the sample of typical readers) to read the text aloud; this often reveals words on which they stumble or which they do not understand, sentences that are too long or too difficult. The interviewer carrying out the test can then mark the points at which the reader hesitates or makes mistakes. The second procedure is to give the complete text to a number of readers, asking them to return as soon as they have read it, at which point they are given a simple test or prepared examination to detect how far they have understood it. The third is to give it to a sample of people to read and then to bring them together to discuss it. If the discussion is recorded on tape the resulting dialogue may suggest changes in style and vocabulary and even provide 'live' phrases and local words that can be incorporated into the text. All three methods yield valuable results and may be applied to the same text. Attempts to discover whether the readers like the text (e.g. the story, or the ideas conveyed) are interesting but by no means easy to carry out objectively. Merely asking the question: "Do you like this?" generally produces the answer that seems likely to please the questioner.

The testing of visual aids (such as flannelgraphs, flip charts, film-strips) involves essentially using the visual aid, with its accompanying talk, lecture or commentary, on the test audience and then questioning individual members of the audience on their comprehension of each picture. Motion picture films are difficult to pre-test. Until a film is complete it does not really convey its message. Once it is complete it is difficult and expensive to make changes, if the results of a pre-test are negative. It may, however, be worth making a film-strip on the theme of the film and pre-testing this. Radio programmes, if recorded on tape, can be played to test audiences, who are then interviewed or tested individually to discover how far they understood the programme.

## Sampling for Action Research

Naturally the difficulty and complexity of pre-testing, as also of topic study, will differ greatly with the nature of the target audience. Where this is large and heterogeneous, more complex sampling is needed to ensure tha the 'test audience' (the sample of persons involved in the test) is representative of all the component elements of the target audience. Fortunately the selective strategy for functional literacy embodies the idea that programmes (and media) should be specific to particular cultural and occupational groups - homogeneous target audiences, so that sampling for test audiences should be much simpler than for more general and extensive adult education programmes.

## Interviewing

Opinion study (as a component of topic study) and pre-testing will make extensive use of interviewing and the action research worker may need to engage and train a number of part-time interviewers and to have some funds for their travel when they come in for training, briefing or reporting to the media centre. Even if he employs trained teachers or extension workers, they must be retrained as interviewers, above all they must learn to reverse their role as teachers to listen to and record other people's ideas and not to inject their own.

## 'Feedback' and reporting

The term 'feedback' signifies a response from the receiver of a message to the communicator or sender. But it involves more than this; and the added element is called 'control', meaning that the response from the receiving end actually controls or shapes the sender's behaviour. It is this element of control which makes feedback so important in communication and it is the function of the action research worker to ensure not only that the ideas and responses of the target audience reach the production staff, but that they do so in time and in digestable form, so that the production staff can make use of the knowledge they obtain to modify and improve the media during production - or - but this is rare - to confirm that the media are faultless and can be finalised, reproduced in quantity and used as they are.

If speed and promptness are necessary to the action research worker (even at the cost of rigorous research design), so is tact. The writers, artists and photographers who may have worked devotedly to produce texts and pictures are not likely to accept very easily reports of failure to communicate or negative responses from the target audience. So the action research worker must explain and support his findings, try not to appear as a censorious critic, and where possible take the writers and artists with him occasionally to observe the opinion study and testing sessions in the field and listen to the responses themselves.

## Evaluation or 'post-testing' of media

The evaluation of completed media is a procedure which does not entirely fit the definition of action research since the feedback from this type of evaluation can hardly modify or control the production process. Once the media are published and distributed in quantity a negative evaluation can do little more than cause disappointment and possibly provide some data for the improvement of similar media in the future. Whilst therefore the evaluation of completed media may help the users of the media to select and employ them more efficiently, it is less valuable than pre-testing to the producers.

## A plea for action research in media production

The integration of action research with the production of educational media is a comparatively new and difficult enterprise. The techniques have seldom been systematically applied by a media production team. They undoubtedly involve new expertise, additional time and additional cost, but the limited experience so far gained indicates that the benefits and improvements will greatly outweigh the costs and difficulties.

## CHAPTER 11

## VISITS TO THE CITY AND GUILDS OF LONDON INSTITUTE, THE NATIONAL COAL BOARD AND PRESSED STEEL FISHER LTD.

The Workshop made three visits to organisations concerned either with research into evaluation and testing in the industrial training field or in the application of evaluation procedures to particular industrial training situations.

In the first category, the City and Guilds of London Institute was visited. The Workshop was addressed by Mr. D. Wheatley of the City and Guilds who outlined the work of his organisation and emphasised particular developments in which the City and Guilds is involved in the preparation of trade tests for operatives and craftsmen within particular industries. An opportunity was also provided for an open discussion between Workshop participants and various City and Guilds personnel engaged in this work.

At the National Coal Board, the Workshop was addressed by Mr. A. Mitchell on the evaluation of training programmes for mineworkers of varying levels of skill and of programmes for instructor training within the mining industry.

At Pressed Steel Fisher, the Workshop was addressed by Mr. G. Smith, Chief Training Superintendent with that organisation, who outlined an experiment in the evaluation of operative training in the motor industry which had brought about a complete reorganisation of this level of training.

### EVALUATION AND TESTING WITHIN BRITISH INDUSTRY

Address by Mr. D. Wheatley
City and Guilds of London Institute

The work of the City and Guilds of London Institute, an organisation which had many working contacts with countries outside Britain, was described. City and Guilds was founded by private initiative over 90 years ago to provide education for those working in industry. Its activities marked the real beginnings of technical education in Britain. Originally, this was achieved by building schools and colleges. This side of its activities had lapsed due to the high cost and to the involvement of the Central Government in this area of education. One of the City and Guilds' original activities was to set examinations in vocational skills. This aspect of its work had grown steadily over the years since its foundation. The

Institute itself published syllabi to which classes organised by a variety of other agencies chose to work. With the growth of State education, City and Guilds was absorbed into the formal structure of technical education although it was not nationalised and remained an autonomous body. New courses were being continually launched in order to meet the advertised needs of commerce and industry.

Initially, a large proportion of the Institute's courses were concerned with educating people to become craftsmen. Subsequently, a category of semi-skilled workers emerged for which training courses had to be devised. It was generally felt that technical training for this category, the operatives, was still lagging behind that of craftsmen. There was a particular need for appropriate courses and examinations. At this level problems of literacy arise. The type of person trained at operative level was not capable of lengthy essay-type answers to formal examination questions. In many cases they do not enter employment through apprenticeship schemes. In addition, employers were often reluctant to send them to technical colleges. This was particularly the case with girls for whom employers did not foresee a lengthy working life. A wide range of examinations had therefore emerged under the Institute's aegis, covering almost all industries and involving approximately half a million students in its courses. In each group of subjects an advisory committee existed comprised of representatives of industry, of the technical teachers and of various Departments of Education in the United Kingdom.

The types of examination varied widely according to the particular industries and the levels of the students. In form, they varied from the traditional essay-type to the most modern multiple-choice type. In addition, practical oral tests on the job were common. These were very valuable at operative level. Through the determination of examination content, the Institute was able to influence teaching methods and, particularly with the introduction of practical examinations foc 1 on special industrial needs, was increasingly able to secure the support of employers.

Examination techniques for persons with low levels of literacy were reviewed. It was important to make it clear to the trainees exactly what the examination was seeking of them. Questions should therefore be highly structured and require the minimum verbal contribution by the student. Great use should be made of sketches and drawings. Recently there had been particular developments in practical testing, requiring an actual performance on the part of the student within the particular area of skills for which he was being trained.

There were disadvantages in this. If it occurred in a technical college, then the facilities of the college would necessarily limit what might be done. For example, practical tests in mining were not feasible. Also, despite the fact that such tests, at the end of practical training, might be fairly lengthy, it was impossible to give a comprehensive coverage of all the work in all the skills which a trained person must be able to offer. 'End tests' were necessarily therefore a sampling of the skills sought. To tackle this problem 'phase tests' were being increasingly developed. Training programmes were organised in set compartments. Tests were conducted upon the completion of each of the components of the training programme, enabling a comprehensive assessment of the skills of the student in all aspects of the programme to be achieved. In addition, such 'phase tests' revealed deficiencies in the programme to the teacher during the course and allowed him to take steps to remedy defects in the programme.

A further trend was the increasing involvement of instructors themselves in evaluation. Traditionally the United Kingdom was wedded to external assessment. Teacher involvement was, however, growing with the increasing professional skills on the part of teachers. There were now many more professionally trained teachers in technical education than had been the case a decade ago. These teachers had already had a training in assessment techniques and it was important that an outlet for these assessment skills should be provided. A further development was the use of specific work projects within industry. Trainees would be required to carry out projects which had the approval of both their technical colleges and the industry to which they were attached, thus providing a course exactly suited to the interests of the trainee, and providing an indication of his real capacity in a specialised area of his training.

The experience of the Institute had shown that all or some of these elements can be built into an examination formula composed of varying components, so that the right things could be tested for a particular vocational purpose and so that both students and teachers throughout the training programme would be aiming in the most rewarding directions.

In Britain the pattern of industry was being greatly altered by the increasing involvement of Industrial Training Boards, which the various industries were now, by law, required to set up. In particular, several of these industries put substantial effort into the detailed analysis of the skills required for certain jobs, and this detailed job analysis was becoming increasingly used as the basis for the phasing of instruction and testing. A clear distinction had emerged between educational knowledge and industrial skill and procedures were being devised which would test each of these items separately. It was important to be absolutely clear on what was being tested and what were the aims of training. The questions to which a training programme should be subject were:

a. whether the trainee, on completion of the programme, could do a task;
b. whether he could do it to the standard of skill expected of the industry;
c. whether he could do it within a time which was economic for the industry and with an economic use of materials.

The essence of modern training programmes was speed and effectiveness.

## DISCUSSION ARISING FROM MR. WHEATLEY'S ADDRESS

### Curriculum development

In answer to a question on the influence of 'feed back' on curriculum development, Mr. Wheatley reported that curricula were being subjected to intensive pressures towards change. There was, however, a feeling in many quarters that over-frequent change was dangerous. The Institute's established arrangement was that every three years a syllabus should be reprinted and might, at this stage, be amended, whilst every five years there was a major inspection of the syllabus. Problems and queries were fed into the advisory committees on on each subject by the teachers, the technical colleges and the Industrial Training Boards, so that a continual reformulation of syllabi, in accordance with modern demands, was possible.

### Tests in languages other than English

In general, the Institute confined its activities to the English language, although a small pilot project had been completed in Hong Kong for training plumbers. Whilst the Institute's syllabi were copyright, its policy was the promotion of technical education generally and the reproduction of the Institute's syllabi in other languages would normally be approved. There were, however, dangers of over-faithful reproduction of syllabi which were designed essentially for United Kingdom conditions.

### Agricultural courses

The Institute provided courses in response to demand in Tropical Agriculture and in Tropical Agricultural Engineering.

This area was, however, one in which there was a maximum necessity for flexibility in view of the diversity of conditions within the developing countries. Adaptation of syllabi, particularly in this area, was very necessary, if it were not as yet possible for syllabi to be developed within the particular countries to meet local requirements as specifically as possible.

## Assistance with technical education outside Britain

The Institute had a lengthy record of such activities and was willing to make its services available to those concerned in the work of evaluation and examination in developing countries. Already the Institute was actively co-operating with the West and East African Examinations Council, particularly in the training of examiners.

## Level of entrant for courses for operatives

The Institute itself imposed no minimum levels of education or attainment for entry to training courses for operatives. Admission to courses was dealt with by technical colleges themselves. In many cases it had to be recognised that industry must recruit personnel whose academic level was low for certain categories of employment. At times this might verge on illiteracy. Whilst the Institute sought to accommodate trainees of such academic levels by maximising practical aspects of training and minimising the need for verbal skills, in practice minimum literacy was required. The Institute had had requests for persons taking tests to be allowed the assistance of persons who would write on behalf of the person examined, at his dictation. This had not, however, been accepted. In general the Institute's aim was to test a person on job performance. At the lowest level, certificates were being devised which would indicate not whether a person had passed or failed, but simply that he had taken a certain course of training and that he had achieved certain standards on this course.

## Follow-up

The Institute itself was not able to conduct follow-up procedures. The specialist subject committee provided professional feedback on training programmes. In addition, the various Industrial Training Boards were setting up their own research programmes and the Department of Employment and Productivity had now established its own manpower and research unit. For the present, therefore, follow-up to training programmes was on an informal basis. In the future it could be expected to become much more formalised and rigorous although follow-up would not be carried out by the City and Guilds itself.

# THE EVALUATION OF TRAINING PROGRAMMES FOR MINEWORKERS AND MINING INSTRUCTORS IN THE COAL INDUSTRY IN BRITAIN

Address by Mr. A. Mitchell
National Coal Board

The National Coal Board has a total of 400,000 employees and a training budget of £7 million per annum, a massive industrial training operation. Training within the National Coal Board was divided into two categories:

a. staff training for administrative and management positions;
b. industrial training of craftsmen and operatives.

Attention would be concentrated on the second of these, where the major effort in the evaluation of training programmes was being made.

Despite the admitted difficulties surrounding the evaluation of industrial training, the National Coal Board accepted the necessity for this. Evaluation was focused at two levels:

a. youth training through a dual system of apprenticeship aimed at producing both 'craftsmen' for the coal industry, mainly electrical and mechanical engineers, and 'operatives', fully trained mine workers. Programmes lasted for 3½ to 4 years. At each stage in the training programmes, practical tests of knowledge and skill were carried out. There were two principal intermediate tests leading to a final test. To become a fully qualified craftsman or mine worker every entrant in the industry must pass his final test. Because of statutory requirements, the coal industry now employed only fully trained personnel;

b. instructor training - carried out in the main by persons who had themselves worked at the specialised tasks which they now sought to communicate. Again training programmes were compartmentalised and a rigorous test was applied at each interval in the training programme.

Evaluation was a constant aspect of the training programmes both for the benefit of the trainee himself and for the guidance of the instructor.

The evaluation of the overall effect of training on the industry was especially difficult. Programmes reached across many levels of skill and attainment and the industry itself was very large. For several years, however, the coal industry had shown productivity gains of 8 - 9% per annum. During this time the technological basis of the extraction of coal had changed substantially and training programmes had been devised to meet these new technological requirements. Whilst, therefore, the recent history of the industry revealed substantial productivity gains, it would be very difficult to isolate the contribution which training made to these gains. Some research had been carried out within the industry at the level of the work unit - the colliery. A typical colliery had been selected with average productivity results. A complete modern system of training had been applied from management level to worker level. The effects were not as yet precisely known. However, since November 1968, improvement in output of over 30% had been recorded. In addition, there had been a substantial reduction of labour turnover and a change in the attitudes of the miners towards new techniques and methods of working. Again, however, these results could not be categorically attributed to training alone, although training had made an obvious contribution.

With regard to the training of instructors, evaluation procedures were directed at the man on the job, the individual instructor, and the effectiveness of his work. Training was carried out both in purpose-built training centres fully supported by the National Coal Board, and directly on the job. Instructors divided into two types:

a. craft instructors who were fully qualified within the particular craft involved;

b. operative instructors who had themselves mastered particular skills at the coal face and who carried out at least part of the training directly on the job.

As far as the training of operative instructors was concerned, the training programme was essentially involved with knowledge about the particular function of the operative, his skill in performing his task and his attitude towards his task.

## Level of trainees

There were some comparisons possible with the educational level of the labour force in developing countries. Generally, entrants into the coal industry could be expected to have had at least minimal education. Some entrants were still illiterate, although this had been reduced to approximately 1%. When a prospective trainee enters the industry, he is given a simple test of his academic ability based on multiple-choice questions. The main purpose of this test was to select potential craft apprentices for whom a minimal academic standard was required.

In training programmes for craftsmen, it was possible to apply a national standard test. With regard to operative training, such a test was not practical, since collieries differed in their nature, one from the other, and even within collieries there might be substantial differences in mining conditions. No standard test could be devised which would successfully evaluate the training of an individual operative. It had, therefore, been found necessary to tailor-make tests. These had to be devised locally by the local instructors, using their intimate knowledge of actual working conditions. It was therefore necessary to train potential instructors in testing techniques.

In instructor training, the first priority was placed upon the clearest definition of the objectives of training, the instructors, having themselves a detailed knowledge of the components of the various jobs, were required to have an absolute understanding of

a. what a trained operative must be able to do;
b. how well he must be able to perform his task;
c. under what working conditions he would be required to perform it.

Having achieved clarity of objectives, training schemes for instructors could be devised to meet those objectives. In those schemes the widest range of audio-visual techniques were used, from the conventional blackboard to closed-circuit television.

## Testing methods

It was important to regard testing procedures not as unpleasant hurdles. They were essentially

a. devices by which progress could be measured within a particular training programme;
b. devices for measuring the effect of the training experience at the level of the work unit.

They could and were also used as a measure of achievement. The trainee operative or craftsman could complete his course within 3½ years, depending upon his capacity for succeeding in the final qualifying test. It was important, however, to disregard the traditional concept of pre-testing and post-testing. Within training programmes in the coal industry testing was a means of measuring gradual progress all along

ERIC Full Text Provided by ERIC

the line, providing an answer to the fundamental questions

a. whether the trainee is learning what is sought of him;
b. whether the instructor is performing his task properly.

The shaping and reshaping of training programmes, therefore, would depend on evaluation of the ways in which these questions were answered. At a time of rapidly changing mining technology, and at a time when the nature of the industry could be affected by new legislation, training programmes had to be flexible. The continued use of testing procedures within training meant that trainees - instructors, craftsmen, operatives - become very used to testing techniques. Testing becomes a normal and essential part of the training process.

The increasing use of programmed learning within training was in turn having an impact on evaluation. The National Coal Board had set up a project at Fordell, based on providing a four-year programmed training for craftsmen. Entrants to the programme were school leavers starting from the same academic level. They were programmed carefully through the practical use of particular tools and techniques. The programme was self-pacing and self-testing. The trainee could only proceed once he achieved the required standards of skill and knowledge in the particular techniques with the particular tools. Again the trainee works to explicit standards with explicit objectives. Throughout his training programme he keeps a log which provides a record of his achievements and provides an evaluation, at any stage, both of his own performance and that of the group as a whole.

Whilst no precise data were available which would indicate the impact of training on productivity, some indications were emerging particularly with regard to recent experiments in the development of team training of adult miners, which was relevant to the productivity issue. The team approach was essentially concerned with groups of adults on the job, allied to discussions and conferences away from the job. The Coal Board was sufficiently confident in the value of this approach that it was prepared to extract teams from the coal face, and accept complete loss of productivity for this team, in order that training conferences could be held in which the miners themselves would discuss with the trainers and the mining executives their problems and their observations on this new approach. Some evaluation of these teams, on their return to work, had shown that the trained and fully functioning team could make up, within 7-10 days, the output lost by being absent from the coal face for conference purposes for one day. This, however, was a new field regarding which precise data were not yet fully available.

# AN EXPERIMENT IN EVALUATION AND OPERATIVE TRAINING IN THE BRITISH MOTOR INDUSTRY

Address by Mr. G. Smith
Chief Training Superintendent, Pressed Steel Fisher Ltd.

Following a conducted tour of the Pressed Steel Vehicle Body Assembly Plant and training facilities, participants were addressed by Mr. G. Smith, Chief Training Superintendent of Pressed Steel Fisher Ltd., on Pressed Steel's training programme and on the evolution and modification of this programme, particularly as regards operative training, following a major evaluation exercise carried out at the Pressed Steel factory at Linwood in Scotland.

In this instance, Pressed Steel had established a new factory in an area where the traditional industries were of a different nature as far as the working habits of the employees were concerned. It was necessary to accustom workers to the discipline of working on a moving production line and to a pattern of work which, although repetitious, required a consistent level of skill. Because of the fact that there was previously no established motor industry in the area, it was not possible to use traditional training methods to build up the numbers of semi-skilled operatives needed to get the plant quickly into production. The traditional training method - to place the trainee on the shop floor alongside a trained man over a lengthy period and gradually to allow him to follow the techniques that he had observed - could be applied, because of shortage of trainers, only on one of the two daily shifts. To train operatives for the second shift it was necessary to set up for the first time a formal operative training programme. Systematic analysis of the skills and techniques required of an operative was carried out and a training programme worked out, based on the transference of these skills, to workers who were drawn from a wide range of previous employment.

At the same time the opportunity was taken to carry out an evaluation of the formal and the informal methods of training. The results of this evaluation, carried out over several weeks after the newly trained operatives began to work independently, indicated that those operatives who had gone through the formal training programme showed consistently better results in terms of productivity and in attitude towards their job. The former had been measured in terms of units produced, the latter in terms of lateness, absenteeism, wastage, etc.

As a result of this evaluation, operative training had been organised on a formal basis in all other Pressed Steel plants throughout the country. Other branches of the motor industry in Britain had expressed a particular interest in those developments and in several cases had gone over to similar training arrangements. Whilst it was felt that this experiment showed without question the value of formal as against informal operative training, it was not maintained that all of the benefits revealed had arisen from the particular training programme itself. It was conceded that in this instance a Hawthorn effect might be discerned, particularly as the labour force had problems centring on attitudes to alien ways of working. Any programme, therefore, which displayed a special concern with the workers and a willingness to try and understand their problems might have been expected to yield some beneficial results. The training programme in itself, therefore, should not be viewed in too optimistic light.

Consideration was given to the possibility of broader schemes which would study the society carefully and locate the particular social problems before a particular training programme was launched. Although the value of such broader research work was not to be dismissed lightly, it was felt that in the particular industrial context faced by Pressed Steel when setting up the Linwood factory, with the government pressing for immediate action to relieve particular employment problems, such activities were likely to be regarded as rather luxurious. There was as yet no conviction on the part of government and industry of the value of such general preparatory work, despite efforts that were being made to increase understanding of such matters.

# CHAPTER 12

## REPORTS OF WORKING GROUPS AND DISCUSSION

In order to study particular aspects of evaluation in greater depth, the Workshop divided into three working groups and visiting specialists and full-time staff were assigned to them. The themes chosen for consideration were:

Group A - Policy, Planning and Organisation in the Evaluation of Functional Literacy;

Group B - Methods and Techniques Used in the Evaluation of Functional Literacy;

Group C - Cost Benefit Analysis in the Evaluation of Functional Literacy.

The reports of the working groups follow, together with lists of the members who participated in each of them.

An opportunity was subsequently provided for the rapporteurs elected by the three working groups to present their reports verbally to a full session of the Workshop. An account of the discussion of each report follows the text.

### Report of Group A

### Policy, Planning and Organisation in the Evaluation of Functional Literacy

Members:
Dr. J. di Franco
Prof. Aurisela Alvarez Medina
Mrs. Omaina Ahmed el Nahry
Dr. Karol Noya
Porol Sitonggang
Abdulkadir Suleiman
Dr. Gopal S. Vidyarthy

#### The meaning of evaluation

In any experimental literacy project, evaluation should be a natural phase of the project, because experimentation cannot be done without evaluation.

Evaluation implies testing, examining the methods, techniques and approaches used, the organisational pattern applied and the planning involved to measure the quantity and quality of the achievements of projects so as to produce scientific data about the usefulness of money spent, efforts used, time allocated, hypotheses tested, etc.

The Group distinguished between:

- Evaluation: measurement activities carried out scientifically.
- Assessment: a more estimative form of evaluation, not necessarily realized in a scientific way throughout.
- Action Research: systematic study and investigation carried out as part of the operational programme, either by the operational staff or by the evaluational unit. In the latter case the results are fed back immediately to the operational staff to help them to improve their day-to-day work.

Basic differences between evaluation and assessment might arise from the presence of controls in evaluation and the absence of them in assessment, rigorously scientific sampling in evaluation and a less rigorous method in assessment etc.

Evaluation is defined in the Unesco Manual (p.3) as follows:

"Evaluation aims at measuring, wherever possible in quantitative terms, and according to well defined criteria, the major direct and indirect effects of a certain activity, taking into account its objectives as established before its inception. The purpose is not only to determine the nature of this activity in its social, economical or cultural context, but also to derive guide lines for its future planning, as well as that of new projects of a similar nature. Evaluation is an on-going activity which should be planned together with the project itself, comparing the situation before the project is started with that at different stages of its implementation and on its completion, endeavouring to explain its effects and implications from beginning to end."

Arising from the foregoing definition the Group would stress the following objectives:

1. The long-term purpose of evaluation should be to collect data and analyse real situations and the

possibilities and obstacles they offer as regards functional literacy, so as to facilitate the efforts of many institutions, agencies, government organisations and individuals all over the world concerned with the planning and implimentation of of functional literacy projects, particularly in the nineteen seventies.

2. To measure as scientifically as possible the economical, educational and social impact of projects, both on an individual and at group level.
3. The study of the effectiveness of pedagogical methods, techniques, teaching aids and media, alternative approaches in administrative and organisational structures.
4. To study the cost/benefit ratio of functional literacy projects.

Plan of work and calendar of evaluation

The Group feels that structuring activities in time-sequence is essential to the realisation of the evaluation tasks.

The three general phases, i.e. a base-line survey, intermittent measurements during the operational phase and a terminal survey are well suited to divide the tasks according to the specific requirements of all the work concerned.

In some projects the operational phase will be divided into a first stage when operations are established and a second one when functional literacy programmes are extended to other areas or to the existing national literacy programmes.

Evaluation activities should be structured accordingly. For instance, evaluation in the first operational stage should aim at having concrete results with a practical demonstrable value ready before the second stage of operations starts.

The base-line survey should be completed during the preparatory phase of the project, though there might occur some overlapping with the first stage of the operational phase with regard to studies which are intended to provide bench-mark data.

In the base-line survey, emphasis should be given to collecting these bench-mark data at the participant and community levels; some work units should be included, but organisations such as factories, co-operatives or trade unions might prove to be a difficult field of study because of reluctance to contribute facts by the leaders or members of these organisations. A possible structure of evaluation activities of the base-line type during the preparatory phase might consist of:

- an initial phase for establishing the evaluation unit and gathering background data;
- regional problem surveys;
- a national survey of data - a stage in which the operational team engaged in the functional literacy project should plan with the evaluation unit activities of the operational phase in general and of its first year in particular;
- technical feasibility surveys, coupling specific programme activities to socio-economic problem situations; (for instance, in a certain area a number of typical problem situations should be identified and briefly described according to their characteristics. In some of them different functional literacy activities should be established, coupling, for instance, pattern No.1 of functional literacy with situation 'A', pattern No.2 of functional literacy with situation 'B' etc. A general hypothesis would be to test whether the operational activities would make a significant contribution to the solution of the problems in the environment in which they are placed.)
- a stage, partly overlapping with the beginning of operational activities, during which bench-mark studies are being carried on.

Due time should be given to the drafting of the research reports after data have been gathered and processed. It should be stressed that the actual writing of scientific reports involves considerable effort and time.

As regards the six types of evaluation activities related to the operational phase (ref. Unesco Manual p.53), the Group accepts the importance given to follow-up studies of participants and to process studies, particularly of small rural communities. The Group feels, however, that cost/benefit analysis, although important, should only form part of wider studies on organisation and administration; organisational analysis and continuing observation of the project should be made, particularly with regard to such aspects as the structure of authority both internally and in relation to co-operating institutions, the supervision of activities and training and and the professionalisation of personnel. The Group believes that cost/benefit analysis will be extremely difficult to carry out and that some model of this type of research should be pre-tested before establishing general rules on this point.

Operational research in educational methods and media is of the highest practical importance, particularly for the benefit of existing national literacy organisations.

The Group feels that it might be wise to make sample studies of each of several cohorts of learners in view of the often high rate of drop-out, of probable changes in programmes and of the relatively short duration of functional literacy projects.

### Organisational arrangements for the evaluation of functional literacy

In order to implement successfully the strategy of experimental functional literacy projects in member countries, the Group felt that some specific measures should be taken to provide a sound administrative and organisational structure for evaluation.

### The National Evaluation Unit

Administrative structures for evaluation will be required at the national and sub-project level.

At the national level a unit needs to be set up, comprised of the international and national experts assigned under the project. This unit may be called the National Evaluation Unit. It should ideally have the following personnel structure:

1. Director, or officer in charge of evaluation (full-time)
2. Professional staff:
    a. Economist
    b. Sociologist
    c. Psychologist
    d. Statistician

The Unit should have adequate support from short-term consultants in specialised areas and have at its disposal for field work sufficient supporting personnel according to activities and needs.

In order to establish the unit the member countries will make the necessary financial arrangements as regards salaries, honoraria, transportation and other expenses while international aid should provide the necessary equipment, fellowships and international experts.

The Group considered this matter in detail. Apart from various research activities, the Evaluation Unit should actively concern itself with the practical training of all personnel involved, particularly part-time sub-project level personnel. It should also make adequate arrangements for the training of its staff in other institutions.

The Unit may function as an integral part of the whole project. Experiences in some experimental projects were taken into account and it was felt that though the Unit may be a part of the project, a special direct link should be established with an outside organisation which would lay down the norms for and examine the results of evaluation. A university or a research organisation may be identified to provide external guidance and supervision in evaluation.

Alternatively the National Evaluation Unit may be located outside the project and be completely independent. However, for a successful working of the evaluation programme, a working committee consisting of members of the operational staff as well as of the evaluation unit may be set up for mutual consultation and communication of feed back.

The Group stressed that wherever a National Evaluation Unit may be created, it would be essential that it should work very closely with the operational staff in a relationship characterised by team spirit and mutual assistance.

As evaluation is of an inter-disciplinary nature, it is essential that the members of the National Evaluation Unit have specialized experience and competence in related fields of research, for instance agricultural economics, the economics of education, the study of levels of living, group dynamics, etc.

Role of universities and research institutions

Even though the National Evaluation Unit should be staffed with qualified and experienced specialists, yet it will be of more significance if the available resources of the universities and other public and private research institutions could be involved in its activities.

This could be done by paying honoraria to consultants from these institutions, or by assigning parts of the evaluation activities directly to them and making provision for this in the budget of the project.

The university link can offer valuable services, for instance in framing research designs, preparing tools, rewriting investigation data, etc. The universities should be encouraged to assign postgraduate theses on subjects related to functional literacy. Universities could also help in fact-finding and data-processing, particularly coding and tabulation.

Whenever arrangements for such activities are still inadequate, the member-country should explore the possibilities of receiving technical assistance from foreign sources, both international and bilateral. Member countries should also identify suitable national personnel and provide specialised training for them possibly under international programmes.

These arrangements should be made before a programme enters into its operational phase. In those countries where activities have started already, those arrangements should be strengthened.

## Action research

As regards the place of action research, when the National Evaluation Unit is completely independent of the project, a separate action research unit should form part of the project. Where the Evaluation Unit is part of the project some of its members may be assigned to do this type of research in particular.

## Sub-project evaluation units

The Group felt that in order to carry out evaluation at sub-project level, a full-time staff should be created in each geographical area of operation of the project, or in a group of sub-projects. These units should consist of one full-time evaluator, supported by a few full-time assistants and adequate numbers of part-time field workers. These should function under the direct control of the National Evaluation Unit which should specify their functions such as:

a. Identifying the personnel resources in the sub-project area for evaluation work.

b. Arrangements for their selection, training and placement.

c. Implementing guidelines provided by the National Evaluation Unit and adapting the research instruments to regional and local conditions.

Adequate and timely budget provision should be made to ensure an effective working of these regional units.

Arrangements for institutional co-operation and co-ordination

The Group suggested setting up a National Evaluation Committee, which may consist of the Director of the National Evaluation Unit, representatives of the operational units and other related institutions, such as universities, government departments, private industry and private foundations, to perform the following functions:

a. To develop an integrated inter-disciplinary approach to evaluation and research by offering advice and guidance on related aspects of evaluation.

b. To ensure active and timely assistance from specialists at national/sub-project level, for any of the aspects of evaluation.

c. To check periodically the progress and problems related to evaluation.

d. To secure inter-organisational and inter-national co-operation for future tasks in evaluation.

For the effective functioning of the National Evaluation Committee, it is necessary that it should consist of members of high standing, having various areas of expertise. This Committee should meet regularly, at least once a month in the preparatory phase and once every three months in the operational stage.

In most projects the organisation of a separate national-level evaluation committee may not be necessary in view of existing national co-ordination committees. If so the functions of the National Evaluation Committee could be performed by the co-ordination committee in which case the latter committee may co-opt certain additional members in order to form a sub-committee for evaluation.

## Discussion Arising from Report of Group A

### Staffing

Whilst it was agreed that, if evaluation of a fully scientific nature was to be carried out of functional literacy programmes within the life of the project, a wide range of professional expertise would have to be present within the evaluation unit, the particular difficulties faced by many developing countries in obtaining specialists with the skills and experience stipulated in the report of Group A would mean that comprehensive staffing of the scale suggested was unlikely to be obtained. In one particular programme, despite the fact that a full professional staff had been provided for the evaluation unit, all the activities stipulated in the Unesco Manual could not be carried out. Whilst universities might be able to assist the work of evaluation with certain services, this could not always be relied upon and in particular it would be necessary to pay for such services. Project budgets were not always sufficiently large to permit the recruitment of assistance of this nature from outside. In view of the limited availability of specialist evaluation staff in many countries, it might be necessary to cut back on the work allocated to evaluation units in the Unesco Manual and that in particular this work was unlikely to be achieved within the time put forward in the Manual.

The application of cost-effectiveness principles, to the work of an evaluation unit, would imply that a full-time team of experts in all disciplines required for evaluation was unlikely to be available. It could be expected that various specialities would be combined. The Group report had indicated an ideal staffing. In fact what was more likely was that a smaller evaluation unit would be maintained, consisting of the minimum number necessary for useful evaluation activities, although not necessarily on the scale suggested in the Unesco Manual. The method with which the work was approached was perhaps more important than the particular staffing.

### Training

Three levels of staff would be directly involved in evaluation activities, that is, graduates directing the evaluation programme, supervisors of field activities and interviewers. In addition, the evaluation unit should be involved in the orientation of project personnel so as to create a general consciousness in the project of the significance of evaluation and of the possible assistance which evaluation could provide for the operational aspects of the programme.

With regard to the training of evaluation specialists at the graduate level, it was admitted that Unesco, whilst anxious to provide fellowships which would enable suitably qualified personnel to receive a particular training for evaluation, had as yet limited success in this. Universities throughout the world offered courses in a variety of specialised fields of concern to evaluation. In particular, courses in social research methodology would be of substantial value. The specific requirements of graduate evaluators were, however, for integrated courses drawing upon a variety of specialisms. Such courses were unlikely to be provided in universities as at present organised. It was therefore recommended that Unesco should approach selected universities, in particular those in countries where functional literacy projects were in progress, with a request that integrated courses suitable for the training of evaluators should be provided. Such courses might be available to personnel from many countries who were interested in working within evaluation units at the professional level.

In view of the difficulty of providing specialised training for professional evaluation personnel, it was suggested that Unesco should give consideration to the involvement of professional evaluators in high-level conferences at which the specialised aspects of evaluation might be discussed.

## Organisation of functional literacy

The suggestions of the working group were again regarded as ideal. It was felt that the organisational side should be kept as simple as possible.

The possibility of the evaluation unit being located entirely outside the operation of the project was discussed. The report of the working group took up no particular position on the location of the evaluation unit, merely mentioning the possibility of its location either inside or outside the overall project. What was regarded as of supreme importance was that the evaluation unit should be able to present a completely independent view. It was agreed, however, that if the unit was externally organised its activities should be co-ordinated with the operational programme.

## Report of Group B

## Methods and Techniques Used in the Evaluation of Functional Literacy

Members: Hassanin Abd. el Moaty Hassanin
Khalafalla Mohed. Idris
Emmanuel P. R. Mbakile
Kamol Sudaprasert
Mrs. Beatriz Lidia Vazques Fuller
Dr. R. Wiener

### Research design

In the first working session the group decided to concentrate on problems of research design and to take as an example a problem that existed in one of their countries - how to evaluate radio functional literacy broadcasts. The problem centred on the fact that there were three radio stations in the country that people could listen to, only one of which, radio B, carried functional literacy broadcasts. specifically aimed at people attending functional literacy classes and were, in fact, part of the class programme. The programme was on the air two hours a day only two days a week. One suggestion was that one could take a random sample of the population in any one area and compare those listening during the times of the functional literacy broadcasts to radio B, to those listening to radios A and C. An objection that was raised to this plan was that the sample might not include anybody who actually listened to radio B. A second objection was that there was no need to have people who both listened to radios A and C as one was really concerned only with comparing those who listened to radio B with those who listened to other radio programmes. A much tighter research design was suggested. This involved taking a random selection of functional literacy classes which listened to the radio programme - these were the experimental groups. There also was a need for a number of control groups. One of these would consist of people who attended functional literacy classes which did not have the radio broadcasts and this would show the effect of the radio broadcasts on the functional literacy programme. A second control group would consist of a matched group of people who listened to radio B but did not attend functional literacy classes, which would show the effect of radio B on its own. A third control group would consist of a matched group of people not attending classes who listened either to radio A or C which would assess their effect in relation to radio B and finally one could have a matched control group of people who did not listen to any radio programmes. This would enable one to assess the effect of either radios A or C or radio B. In all cases one would need a base line study.

A further complexity is that one would also have to control the standard of teaching. In matching teachers one cannot only consider qualifications and years of experience because good teaching depends primarily on the relationship between the teacher and the pupils and that is what one ought to be measuring.

As a general point on evaluating teaching, it was agreed that one must carefully define the aims of teaching. The results of teaching could be measured in terms of academic success, individual maturity, socialisation, political awareness etc. and that success in any one sphere was not necessarily related to success in any other.

## Collecting base line data

The second session concentrated on different techniques for collecting base line information. Even though the effects of the functional literacy campaign were unlikely to be directly observable at national level, there was still a need to look at documents at this level so that one could obtain relevant information for sample design. When looking at documents it is essential that one should check not only their internal consistency but also that any two documents about the same area or topic report the same findings.

At community, household and to some extent regional level, interviews and observation techniques are most useful. The Group agreed that one ought to be very flexible in the use of techniques and that different areas of study such as literacy, productivity, political and social relations and so on, needed different types of investigation. The group also agreed that it was possible to measure the same indicator by a number of different techniques. It considered, as an example, the problem of how one measured a change in the amount of fertilizer that farmers in a community used. It was pointed out that one could get figures from the local store showing how many bags of fertilizer had been sold; this would involve interview and document analysis. However, the number of bags sold would not indicate how many of them were actually used or which went to farmers taking part in the functional literacy campaign. A second technique would be to sample the farmers who were taking part in the functional literacy programme and those who were not and to interview them. Again, one member of the Group pointed out that farmers sometimes used fertilizer for purposes other than that for which it was intended. A third technique was then suggested - that of observing a sample of farmers at the time they ought to be using the fertilizer to see if they actually were. The different methods obviously have different degrees of precision but the one chosen will depend on the time, cost and quality of the staff available.

The Group then considered observational techniques and argued that the aim in observation ought to be to devise as many objective measurement techniques as possible. This was for four reasons:

1. so that one can quantifiably and statistically compare differences found at base line, interim and final survey stages;
2. so that observations carried out in different communities are comparable;
3. so that a new observer coming in at either the interim or final survey stages could carry out the same analysis as that done at the base line stage;
4. so that one can check the reliability of the data gathered.

There are a number of such objective techniques that an observer can use:

a. A check list: by the use of a properly designed and a piloted check list an observer can quantify such things as:

    1. amount of fertilizer sold;
    2. number of books read in a family;
    3. amount of money a family spends on drink, and so on.

b. Socio-metric techniques.

c. A map of the area showing location of schools, churches, functional literacy households, etc.

d. Charts showing the pattern of interaction between a sample of functional literacy households and a sample of non-functional literacy control households. This, for example, would be useful if one had hypothesised that households taking part in the functional literacy programme would increase the amount of interaction they had with each other and decrease the amount they had with non-participant families.

If an observer only comes on an occasional basis then he should consider the population to be studied as the total activities carried out by the community he is observing. As this is a very heterogeneous population it should then be stratified along the main dimensions that the observer hypothesises would influence the indicators he is measuring, for example, the dry versus the rainy season; the market day versus the day of rest, versus the rest of the week; day versus night time activities, etc. Thus the observer might visit a community at both day and night for three days covering a market day, day of rest and another day in both the dry and rainy seasons. He could also randomly sample the households taking part in the functional literacy campaign and only observe a small number of these and compare them to a matched group of control households. To test the reliability of these findings there should be a subsequent visit, ideally carried out by another observer.

At the participant level many instruments can be used to collect relevant information. These include: attitude scales, achievement tests, questionnaires, motivation tests, opinion questions and others.

Attitude scales

One might define an attitude as a degree of positive or negative effect associated with psychological objects. By a psychological object one means any symbol, phrase, slogan, person, institution, ideal or idea toward which people can differ with respect to positive or negative feeling. In looking at attitude scales the Group concentrated on Likert scales. The difficulty of constructing such a scale was gone into. It was pointed out that one first has to find a large number of representative statements and questions about the area to be studied. Those then have to be designed as specific attitude scale items and tried out on a small sample representative of the final population to which the final scale is to be given. An item analysis then has to be carried out in which the items which discriminate best between the high scorers and low scorers are retained. Many members of the Group argued that not only was this a rather long and costly procedure but also that it required a qualified social scientist to construct such a scale. One suggested alternative was simply to use only one question instead of a whole scale. This, of course, strictly means that one is measuring rather unreliably an opinion rather than a more basic attitude. For example, instead of a multi-question scale about literacy one could have one question such as:

"How important is it to be literate?"

| | | | | |
|---|---|---|---|---|
| Very Important | Important | Not Sure | Not Important | Not Important at all |

Another sample technique is a paired choice of question, for example:

"Is it more important to be literate or earn money?"

## Motivational tests

One member of the Group reported his experiences in using an adaptation of the Atkinson McClelland Need Achievement Pictorial Test. Another member suggested that there were serious doubts about the reliability and validity of such a measurement. It was suggested that this was because firstly there was a difficulty as regards both unreliability and the design of a comprehensive and meaningful coding frame in analysing responses to projected pictures. A second difficulty was in being able to design a series of pictures which would adequately represent a total range of options open to the subjects. A third reason was that the ability to identify with figures in the pictures is dependent on the past experiences of subjects which might well vary from area to area. A fourth difficulty is that they are difficult tests to administer and need highly trained interviewers. A fifth reason is that the answers are, to some extent, dependent upon literacy and that any changes found in between base line and interim measurements might well be due to differences in literacy rather than to a change in the needs of the subjects through their participation in the programme.

It was then suggested that one could use more objective questions to get at motivation. One could, for example, ask farmers to estimate how much they would like to produce per acre, e.g.

"How many tons would you like to produce per acre?"

20 tons _____ 40 tons _____ 60 tons _____ 80 tons _____

Why? ________________________________

Such a question followed by a "why" question would enable one to obtain some easily quantifiable measurement of motivation and some idea of the explanation behind it.

## Interviewing

The Group briefly looked at the problem of interview design and one member pointed out that in his country one of the main difficulties was that the people in the rural districts spoke a different vernacular to that of the university students used as interviewers. The problem of finding good interviewers was difficult because the majority of university students did not want to leave the urban areas to go into the rural districts. An alternative suggested was to use interviewers from other ministries but these did not seem to be available in this case. Another suggestion to get round the language problem was to use the teachers in the rural areas. There were two problems here, however, the first of these was that as they were directly involved in the situation one could not guarantee their objectivity. A further suggestion was that then one ought to build in a large number of reliability checks. It was pointed out that if one used the interviewers from the town and was thus forced to use interpreters then there were two additional potential sources of unreliability; during translation of the question and during the translation of the answer.

Another point made about the ordering of questions within an interview schedule was that questions subjects might be reluctant to answer should be placed at the end.

## Sampling, matching and research design

The Group then went on to discuss further problems of sampling, matching and research design. Two members presented problems that they had had. In the first country the problem arose from the fact that the evaluation team consisted of only the Unesco expert and one national. The functional literacy campaign was introduced into three areas in the country and within each area to all of the villages. The problem was how to produce an evaluation scheme which would be manageable for such a small unit?

Through discussion it turned out that two areas were very similar, e.g. the people came from the same tribe and lived in agricultural communities. It was therefore suggested that one should evaluate only one of these areas and for convenience should select the one closest to the evaluation unit.

This left two areas. Both of these areas were homogeneous areas with the main difference in one being that some villages were close to a main road and some further away. It was then suggested that in one area one should select one village on a random basis and in the second area one should stratify the villages in terms of those next to the main road and those not, and select at random one from each. For each of the villages chosen it was necessary to include a control community from outside the area as similar as possible to each of the experimental villages.

Finally, as regards individual participants, it was argued that within each functional literacy class one could draw a random sample of participants and match these with the people in the villages not taking part in the campaign.

In the second study, the member was comparing two villages. In one of these the experimental village, a functional literacy programme was being carried out and in the other, the control village, a traditional literacy campaign was in operation. It thus became theoretically possible to compare people in the experimental village who took part in the functional literacy programme (Group 1) with a control group of people in the experimental village who did not (Group 2). One could also compare these with people in the control village who took part in the traditional literacy campaign (Group 3) and those who did not (Group 4). Ideally this would have enable one to find out:

a. the effect of the functional literacy programme by comparing Groups 1 and 2;

b. the Hawthorn effects of the functional literacy programme by comparing Groups 2 and 4;

c. functional literacy versus traditional literacy by comparing Groups 1 and 3.

However, in the study, the families chosen in each of the four groups had not been matched and thus such comparisons were not possible. Another difficulty was that only three families had been chosen in each group and this meant that if any family withdrew from the campaign then it would be impossible to obtain any results. Irrespective of this, the numbers involved were too small for any statistically significant analysis and any results could only be used as being indicative of trends and thus used to design hypotheses for future testing.

It was pointed out that when choosing control families to be matched with experimental families this could be done either:

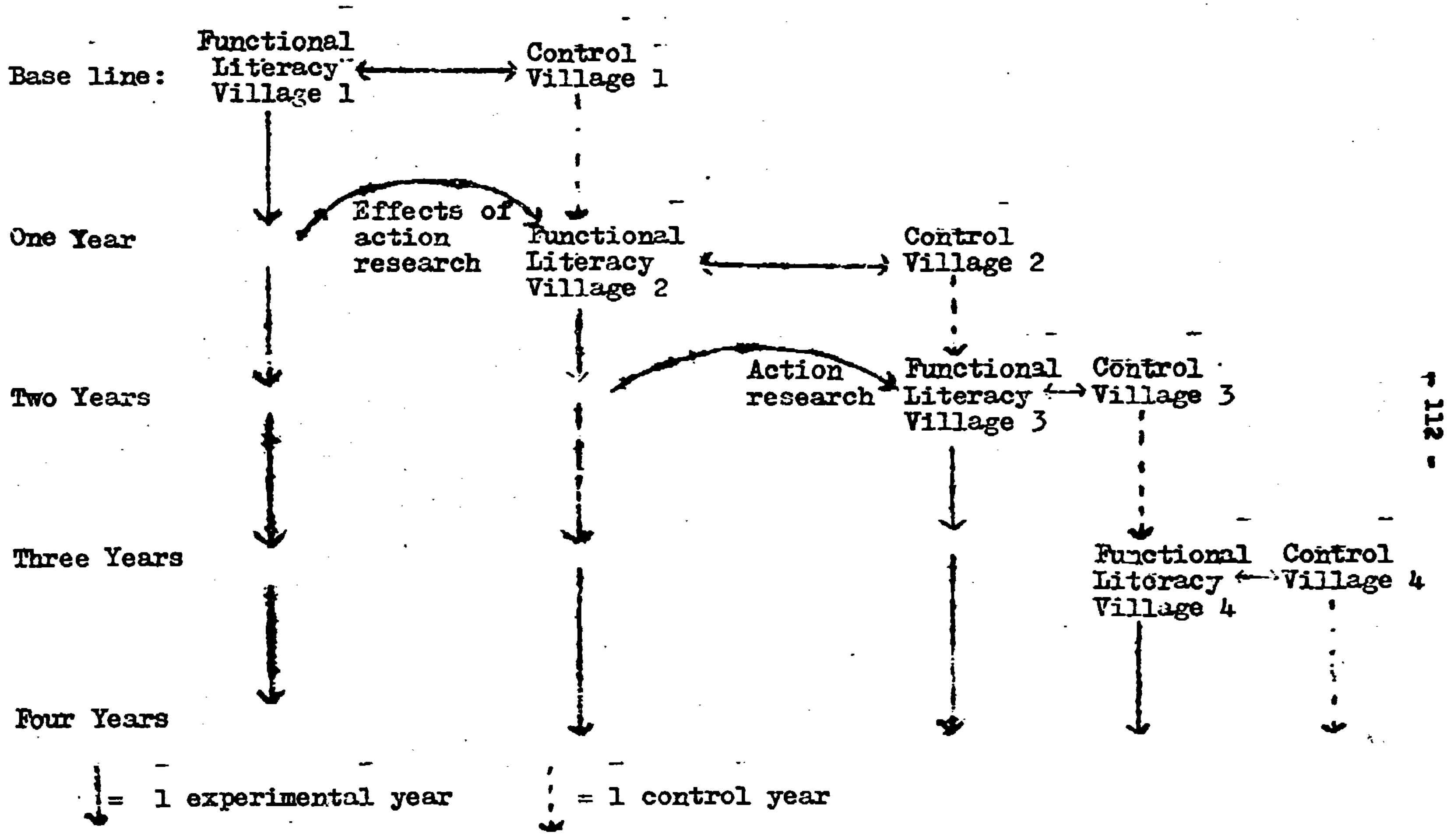
a. by choosing a family from the control village such that it matched the experimental family on a number of important dimensions, or

b. by selecting at random as many control as experimental families from the control population of families.

Whereas the first method would obviously produce a much tighter matching, it did mean that the control group was not necessarily representative of all families in that population who had not participated in the functional literacy campaign. As sometimes one wanted to be able to treat the control sample as being a representative sample of the control population, in some instances it might be necessary to have two control groups, one of whom would be matched on precise variables with the experimental group and a second one selected at random.

Another difficulty reported by at least two members of the Group and mentioned in the Manual, was the problem of control communities or households feeling deprived at not being given the same opportunities as the experimental group. In one case the control village threatened to march over to the experimental village and take away all the functional literacy materials. A number of suggestions were made as to how to get over this problem:

1. to choose a control village that was some distance away. Many difficulties would result if one did this, however. Firstly, the control village would have to be far enough away so that the inhabitants of the two communities did not meet even at the market place; secondly, there would have to be no publicity about the campaign in the mass media; and thirdly, if the control village was too far away it would probably lose in comparability with the experimental village.
2. A second solution was to give the control village a traditional literacy campaign so that they did not feel they were being neglected. Again, this would only work as long as the inhabitants of the control village did not perceive that the experimental village was receiving additional returns from its campaign.
3. Another suggestion put forward was that the control village should be promised that the functional literacy programme would be introduced next into their area once it had proved to be successful. It was, however, pointed out that the control village was unlikely to wait five years if it saw that the campaign was being successful.
4. A fourth suggestion is outlined on the next page.


Base line:
Functional Literacy Village 1
Control Village 1
One Year
Effects of action research
Functional Literacy Village 2
Control Village 2
Two Years
Action research
Functional Literacy Village 3
Control Village 3
Three Years
Functional Literacy Village 4
Control Village 4
Four Years
= 1 experimental year
= 1 control year

As can be seen, the plan requires that at the base line study there is one experimental village and one control village, that after the end of the first year the control village becomes a second experimental village and that a new control village is introduced. This has two advantages:

a. Firstly, the results of any action research that is being carried out during the first year can be fed back into the design of the campaign which can then start afresh in the second experimental village. This, therefore, means that experimental village 1 then serves as a control group for experimental village 2 saying, in fact, this is what would have happened if one had not fed back in the [illegible] of the action research. It also means, of course, that the action research does not interfere with the evaluation.

b. The second advantage is that no control village ever remains a control village for longer than one year because at the end of that year it becomes the next experimental village. This process can obviously be repeated through the five years of a project as long as all the control groups are matched beforehand.

Another member of the Group wanted to know how it was possible to match individuals on intelligence. It was suggested by other members that intelligence was a very vague and at least a multi-factorial concept which was difficult enough to measure with refined techniques on highly literate subjects but virtually impossible to measure in totality with illiterate ones. One suggestion was that one should isolate salient factors of intelligence which might be measureable, such as mechanical aptitude. Overall, however, the Group considered that intelligence was too complex to be considered as a matching variable.

It was also suggested that there ought to be flexibility in the timing of the interim study, for example, if it was suspected that a new government was about to be elected which might cut back on the funds for evaluation, or if operators were about to pollute the research by introducing the results of action research studies then interim studies should be timed to take place before this contamination occurred.

The Manual suggested that experiments should be carried out to test different aspects of a functional literacy programme. One member wanted a list of such experiments but it was pointed out that any single part of such a programme was open to experimentation. It was then suggested that even though the Manual seemed to be recommending that the same experiments ought to be carried out in all the projects, this might well be a waste of time. If only a limited number of experiments could be carried out then it might be better to have each evaluation unit carrying out a different series of experiments. In order to check that the results of these experiments were 'culture free', one would then only have to test the results out in one other country with a markedly different cultural orientation. If this programme were to be adopted then ideally Unesco ought to draw up a list of all the experiments that it would like to see done and then to suggest to different units that they might like to carry these experiments out within their projects, dependent upon the size of their projects and the range of specialist staff they had available.

The Group briefly considered the follow-up study which they defined as being something which was carried out after the five year evaluation period had ended. It was suggested that such a study would have two main justifications. The first of these is that literacy, to be an enduring characteristic, needs to be practised and that even though 80% of the participants might have achieved a certain standard in functional literacy after five years this might not last once the campaign has ended. A second reason is that some productivity changes will not be expected to show up for some years after a campaign has finished. Thirdly, only with follow-up procedures can one isolate what was due to the Hawthorn effect of the experimental condition.

Finally the Group discussed data processing and analysis. The first thing to do when data is received at a processing centre is to check its reliability and consistency. The next step will be to design coding frames for the questions which have not as yet been closed. A reliability check on the coders ought also to be built in here. The next step is to tabulate the results. This can be done either manually or by machine. Members of the Group felt that not enough attention was paid to training people in data processing and that more time, personnel and money should be put into this part of a programme.

A lot of the data at this stage will be in terms of frequencies and as such groups can be compared either in terms of percentages (not if the numbers are too small) using a test such as Zubin's nomographs (see A. N. Oppenheim) to measure significance, or by the use of the $x^2$. When the individual participants have been rated or scored in some quantifiable terms it will then be possible to use some form of non-parametric statistic (see Seigl's book on non-parametric statistics).

For more complex statistical analysis such as multivariate or factor-analytical a statistician needs to be employed.

One member pointed out that in a study in which he was involved five out of thirty people had changed their reading habits after a year and he wanted to know what to do with this data. It was pointed out firstly that there were non-parametric tests which one could use even on data as simple as this but that a change in such a small number would be unlikely to be significant and that any differences found would only be a source of hypothesis at this stage. He indicated that he was looking at these five cases to see if there was anything that all five had in common and it was pointed out by another member that the results would be more meaningful if one also took five matched subjects where no change had taken place.

Finally, the difficulty was pointed out of writing a report which was both statistically complex enough for social scientists but simple enough for non-social scientists in the various ministries to understand. Two solutions were suggested: firstly, that separate reports should be written and secondly that the statistical tables could be added as appendices. The Group as a whole agreed that the second alternative was preferable.

In conclusion, the Group agreed to ask Unesco for:

a. more clear-cut objectives concerning methodology in functional literacy projects;

b. a priority list of experiments that Unesco wished to be done in functional literacy projects;

c. the systematic training of evaluation teams at either international level or through the evaluation expert appointed by Unesco;

d. distribution of a bulletin on functional literacy which includes evaluation techniques;

e. some recommendations on evaluation research to the member countries.

## Discussion Arising from Report of Group B

### Research design

Doubts were expressed on the practicability of a simultaneous design for experimental and control groups, particularly in rural areas. In these areas there were special problems of matching. The before and after design might be more appropriate in assessing the impact of functional literacy.

### Sampling

The selection of the sample within the experimental groups was considered in the model put forward by the working group. Samples had been selected by using demographic data, particularly on the occupation of the family, on the size of the family etc. This amounted to a stratified sample. Matching problems were again foreseen, with particular difficulty in matching families. In this connection the choice of an appropriate method was seen as dependent on the particular social situation, especially where the culture studied was homogeneous or otherwise.

### Dependent variables for measurement

It was regarded as important that if, in the report of Group B, a model was to be laid down of evaluation techniques and their application, then it should be made absolutely clear what were the dependent variables which were sought to be measured in order that it should be clearly realised what were intended to be evaluated.

### Social science techniques and the objectives of the Experimental World Programme

The Group considered how testing and measurement techniques used in technical and vocational training, fitted into the overall process of measuring the impact of functional literacy on development, development in this instance being seen in both terms of the individual human being and in terms of the society of which he was part. In the first place, it was felt that the aims of evaluation would have to be expressed in realistic terms before any useful evaluation exercises could be carried out. The problem would have to be broken into researchable components and evaluation activities directed at those separate areas. Ultimately, however, a total assessment of all of those activities would have to be arrived at which might vindicate the general hypotheses that functional literacy was a positive factor in development.

Whilst it was agreed that this was the general purpose of the Experimental World Programme, it was at the same time contended that functional literacy, as a causal factor, can best be isolated at the individual and the family level. Tests existed which would indicate, as far as participants in functional literacy programmes were concerned, gains in knowledge, changes in attitude, disposition towards the adoption of new techniques and the skills attained through exposure to functional literacy programmes. A variety of such tests might be necessary, their application might have to be adapted to various situations, for example the work place or the home. The principle, however, was the same, that is the isolation of the effect of functional literacy on the participants. Such techniques might give positive results in measuring simple concrete gains at the level of the individual and the local community.

When, however, the objective was to capitalise the effect of functional literacy on national development, then the exercise was transformed. It was suggested that at this level functional literacy was not a factor which might be validly isolated, but should be seen as part of a comprehensive exercise in social engineering. The Workshop agreed that the techniques discussed by Group B would enable the impact of the amalgam of literacy and technical and vocational training (functional literacy) to be satisfactorily evaluated at the individual level and at the level of the work unit. Beyond this level, that is at regional and national level, the exercise was not regarded as a serious possibility within the limited time suggested in the Unesco Manual and with the limited staff and funds at the disposal of most projects within the Experimental World Programme.

# Report of Group C

## Cost Benefit Analysis in the Evaluation of Functional Literacy

Members: Mohammad Mehdi Behkish
Abraha Ghermazion
Dr. Mrs. Sumati Mulay
Michael Ostey
Prof. Ayo Ogunsheye

### The need for cost benefit analysis

The projects for functional literacy require substantial amounts of funds. International assistance has been provided for these projects by UNDP and UNESCO. In countries where the projects are being undertaken, national governments have had to make substantial provision in their budgets for them. Moreover, sometimes groups, such as industrialists, trade unions, farmers organisations etc., and the participants themselves are contributing financially or otherwise.

This poses a question - when and where there is a problem of scarcity of resources, which might be used in several different ways for achieving the objective of economic growth, how should cost benefit analysis be applied to provide sound evidence for arriving at decisions concerning allocation of scarce resources and funds? In particular, how should such an exercise be applied in the evaluation of functional literacy programmes?

Before we proceed further towards the problem of cost benefit analysis it would be useful to explain the concepts involved in such a study.

### Definition and concept of functional literacy

Functional literacy is defined as a training programme in which the literacy component is built in to enable a person to acquire through a literacy programme essential knowledge and skill to increase his working efficiency and productivity.

According to this concept vocational and technical training whether in industry or agriculture, is linked up with literacy. This entails a new approach to almost every aspect of literacy work.

## Costs

There are two concepts of costs:

a. Accountancy Cost - the actual monetary expenditure incurred in the programme;

b. Opportunity Cost - the value which would have been gained if the resources were expended for the next best use. For example, the opportunity cost of time given for the literacy project is the amount of money that would have been earned if the time was used in some other way considered next best.

The Committee decided to emphasise the second concept of cost. It also stressed that in comparing literacy projects it is essential that the same concept of cost should be used.

## Benefits

The benefit of a particular project means the changes it produced. In a cost benefit analysis study, the task is to establish the results of old and new training programmes or the situation before and after a new training programme to cost the two different programmes or the new programme and then to quantify and evaluate the change in results.

## Analysis of cost benefit

For the study of cost benefit analysis it is essential to study the components of costs and of benefits. However, it must be stressed here that there is no suggestion that the list of items included in each category is exhaustive, or that the lists will necessarily apply without modification in every country.

## Cost components to be measured

### 1. Cost of initiating the programme

The initiation of training programmes involves cost. The following are some of the most important cost items which are to be measured while initiating the programme.

a. Fees of consultant who is not the member of the staff.
b. Opportunity cost of the services of members of staff who are engaged in the programme.
c. Cost of travelling, stationery, subsistence, etc.

2. Cost of administering the programme

This consists of the staff outlays on the training programme.

a. Salary of training officers.
b. Opportunity cost of voluntary workers.
c. Salary of secretarial staff.
d. Office equipment such as desks, typewriters - here the market price of new equipment is measured and the opportunity cost of the old is to be considered.

3. Cost of recruiting trainees

Just as the initiation of the programme needs expenditure, the task of recruiting trainees also involves some cost Elements in this process are:

a. Publicity.
b. Search and selection.
c. Radio announcements.
d. Visits and travelling of personnel.
e. Subsistence.
f. Aptitude tests - if used.

4. Cost of fixed training capital

This cost is concerned with:

a. The buildings.
b. Training schools.
c. Lecture halls.

In the case of new buildings the opportunity cost could be quite large and this may be much more than the actual cost of the building. The rental is included. Opportunity cost should be considered if the programme is carried on in existing accommodation.

5. The cost of working training capital

The main reason for distinguishing this item from the above cost is that it is a variable cost. However, the line between fixed and variable costs is not always easy to draw. It includes:

a. Cost of supplies and materials.
b. Cost of media.

The opportunity cost of the above applies if existing supplies are used but where new supplies are purchased the market price applies.

6. Cost of providing instruction

This cost concerns the training of instructors.

7. Cost of giving instruction

This is a variable cost and therefore distinguishable from cost 6.

a. The salary of the instructors or
b. Part-time instructor's fees.
c. The opportunity cost of the part-time instructor who is already on the staff.
d. Opportunity cost of the instructor who is honorary.

8. Cost of the wages of trainees - not cost of trainee output value.

This involves the loss value of productive work done by the trainee during the training period. Of course this cost does not normally arise in the situation where the trainees participate in the programme during their spare or leisure time.

Methods of measuring benefits

1. Value of economic objectives - the benefits are measured in terms of the various objectives of the programme. Certain characteristics are to be observed for setting up the economic objectives in terms of benefits.

a. They should be decided in advance before the project is launched.
b. They should be immediately valuable in economic terms.
c. The benefits to be measured should be realisable within the life of the project.
d. The level of achievement [illegible] it be low enough to be clearly attributable [illegible] o training process.

Examples of these economic [illegible] tives can be given as follows:

a. Higher yields in a crop such as wheat, paddy, tobacco.
b. Reduction of crop damage due to a given pest.
c. Reduction of wastage on a given production line.
d. Reduction of stock held in a production process.
e. Reduction of man-power of a given agricultural operation.

The following will often not be appropriate to be measured:

a. Total output of the firm, factory or mill.
b. Total agricultural production.

## 2. Value of economic spillover benefits

This includes the value of any increased productivity that was not planned for in the given project. In other words, the side effects coming out of the project - they are:

a. Value of any general increase in production.
b. Reduction of labour turnover.

## 3. Behavioural objectives

The behavioural objectives are to be considered in absence of or in addition to economic objectives.

a. Increase in knowledge.
b. Acquisition of new skill.
c. Adoption of new technology.

## 4. Non-monetary objectives

These are the benefits which are not immediately or directly measurable in economic terms.

a. Changes in attitude.
b. Changes in motivation.
c. Changes in values.
d. Changes in social systems, e.g.
   i. in communications patterns.
   ii. in person to person and group relations.
   iii. in stratifications.
   iv. in social participation.
   v. in media participation.

5. Experience gained is also one of the main benefits.

## Methodology used for measuring the cost benefit analysis

In cost benefit analysis we are confronted with different training situations as well as with training versus non-training situations. Hence the analysis must necessarily seek to quantify the before and after situation.

The cost classification discussed above facilitates a systematic monitoring of changes in training costs brought about by the introduction of a new training scheme. The costs involved would be measured against the various benefit variables discussed above. The difference between the before and after measurement of costs and benefits may indicate the cost benefit ratio.

## Conclusion

We return finally to the original question how to proceed with cost benefit analyses? Cost benefit analysis raises several problems. The most difficult problem is to set objectives which are realisable and measurable within the period of a functional literacy project.*

The valuation of certain non-monetary items of benefits may present a difficult problem. Here the valuation may not be perfect, but certainly may indicate the orders of magnitude. Moreover the inadequate records of many projects make it extremely difficult to measure the cost. Many times it may become a problem of deciding the opportunity cost. But as the cost benefit analysis is of great importance, it may be suggested here that adequate procedures for budgeting and book-keeping should be established from the beginning with a view to facilitating cost analysis.

Consideration may also be given to action research. In action research the cost of the projects communicable benefits can be measured and checked from time to time which will give the picture of cost involved and the benefits attained.

---

* The objectives having been set the next task is to achieve a level of costs which are sufficiently low to make the project worthwhile.

## Discussion Arising from Report of Group C

In considering the utility of cost benefit analysis applied to functional literacy programmes, the need was stressed for the establishment of measurable objectives. The essential utility of such techniques was seen in the ability which they had to indicate through a close analysis of costs related to benefits, whether a project was worth while.

### Costs of experimental phase

Regarding projects within the Experimental World Programme, it was agreed that there was a need to distinguish between the costs incurred by the experimental nature of projects, (e.g. costs of research activities; comparative experiments in using different media) and costs which would be incurred after the experimental phase when the programme was generalised. Experimental projects necessarily incurred higher costs. Higher spending might be justified at this stage in order to discover how to do something more cheaply in future. Several components of cost in the experimental phase would not recur. In addition, on the benefits side at the post-experimental phase, since much larger numbers would be involved, unit costs could be expected to be reduced. Some concern, however, was felt that many Experimental World Programme projects were of such size as to raise doubts as to their experimental nature.

### Systems analysis

The possibility was considered of analysing the efficiency of different types of organisations in carrying out functional literacy programmes. For example, governments might not be the most effective means of promoting such programmes. The need was seen for the application of systems analysis and management study to various forms of organisations; this could be an important factor in the reduction of costs.

### Costing of long-term changes

Particular problems centred on the costing of residual changes as a result of functional literacy programmes. Long-term changes in the attitude, for example of a rural population, towards improved agricultural practices could not be readily expressed in monetary terms, although there could be expected to be a direct effect on productivity. With regard to structures and capital investment, established in the operation of a functional literacy project, the need was seen to consider that many of these structures could be expected to have a life extending beyond that of the project itself and should be costed at the end of the project period and the costs deducted from that of the project. Particular benefits for women were seen arising from functional literacy projects. Whilst it was agreed that no special methodology could be applied to any particular section of society, it was agreed that specialised spill-over effects could be detected in particular parts of the social structure.

The basic principles of action research applied also to cost benefit studies. The interim cost benefit study within a programme designed to improve the operation of the programme could be seen in terms of action research.

### Comments by Mr. M. J. Oatey

Mr. Oatey, who worked with Group C as technical advisor, could not be present at the discussion of the Group's report in the full session, he was therefore invited to send written comments which are reproduced below.

### Measurement of economic benefits

"The most difficult problem is to set objectives which are realisable and measurable within the period of a functional literacy project."

This problem, raised in the report of Group C, will be discussed with particular reference to evaluation at the 'work unit' level, which is considered to be the most feasible. The practical usefulness of the approach has yet to be tested, since even in advanced countries the economic evaluation of industrial training with respect to the firm is virtually non-existent.

### Explicit and implicit benefits

The explicit-implicit distinction is a continuum. The extent to which any benefit deviates from the following four rules, the less explicit and the more implicit it becomes.

Rule 1 : It should be specifiable ex ante (both qualitatively and quantitatively), to an acceptable level of uncertainty.

explicit example: an increase in average operator production from 75 - 80 pieces a day:

implicit example: a possible reduction in labour turnover (unless previous studies have shown the extent to which this can be expected to occur).

7

The 'acceptable level of uncertainty' should correspond with the uncertainty associated with any business activities and investment decisions, and will vary with attitudes to risk, capital available, local practice, etc. It should be emphasised that common business ventures sometimes have surprisingly high levels of uncertainty (e.g. advertising, launching a new product, estimating the costs of a new production process, research and development).

Rule 2 : It should be directly evaluable in economic terms.

explicit example: a 4% reduction in tobacco leaf wastage in the production of cigars:

implicit example: a 90% score on a given Work Motivation Scale (unless previous controlled experiments have shown how much, if anything, production rises following such a motivation score).

The motivation benefit may comply with Rule 1, but can be very difficult to translate into immediate economic terms. On the other hand, the reduction in wastage can be immediately evaluated in terms of the reduction in leaf used to make a given number of cigars.

Rule 3 : It should be realisable within an acceptable time horizon.

explicit example: the ability to select and use fertilizers and insecticides appropriate to given crops (following a 3 year part-time training course):

implicit example: the effect on a trainee's children, whereby they will be better educated and more productive in 12 years time.

These examples assume a project of 5 years.

As will be emphasised below, an implicit benefit is not necessarily unimportant - just more difficult to deal with.

Rule 4 : The benefit should be set at a low enough level such that it can be clearly attributable to the training.

explicit example: increase in output of one production line over one week:

implicit example: increase in output (or profit) of a whole factory over one year.

The point here is to reduce as far as possible the number of variables that can influence the realisation of a particular benefit. The ideal is a stable situation where records have been kept for some time before training, thus making a 'before and after' comparison relatively simple.

Characteristics of Explicit and Implicit Benefits

1. Implicit benefits can, in principle, be rendered explicit by appropriate field studies and controlled experiments. Sometimes this will be relatively easy, other times quite impossible - at least within the time available. Experience will play a great part in rendering benefits more explicit. (Controlled experiments, though valuable in appropriate circumstances, can be complex, time-consuming, unreliable and non-generalisable, and should be embarked on with great caution.)

2. Implicit benefits are not necessarily small or unimportant. If they are considred to be significant, consideration should be given to rendering them explicit (see 1 above). Alternatively they may be measured or estimated with due weight to the uncertainty associated with them (see Rule 1). In any case they will rightly enter the subjective component of the final decision on the undertaking of a project.

3. Explicit benefits are obviously preferable to implicit ones because of ease of measurement. A positive search should be made of projects that are likely to give rise to important explicit benefits. If none can be found, a decision must be taken as to whether the time and resources are available to deal with the more implicit benefits, or whether cost benefit analyses should be attempted at all.

4. A project must be evaluated to confirm that the explicit benefits have been realised (unless there is a very high level of certainty), and to identify and measure as many implicit benefits that have occurred as possible.

5. To justify a project economically, all the benefits do not have to be accounted for, but only so many as will show that the benefits will (or could) exceed the costs. One should therefore start with the most explicit benefits, and only attempt to account for the more implicit ones if the necessity arises ( or if the time and resources to do so are available).

## Application to Literacy Projects

The economic benefits of simple literacy will tend to be rather implicit. However, functional literacy, because of the involvement of vocational and industrial training, is in principle capable of providing projects with reasonable explicit benefits. These will tend to be at the work-unit level: increasing output or decreasing costs within the firm or agricultural unit (see examples of explicit benefits given earlier). Benefits to the community and the economy will be mainly implicit - because techniques of measurement are not available, or because time and resources cannot cope with the complexities of the problem.

In any one country, a positive search must be made for projects with explicit benefits, or with important implicit benefits that could be made more explicit by field studies or by experience. There must be a willingness to recognise that some projects will not be amenable to cost benefit analyses, and possibly that no really suitable ones can be found in a certain country.

It is not easy to see to what extent, if any, the results of projects at the work-unit level will be generalisable to other projects within, or between countries. This could be a second major factor in the selection of projects, and warrants further investigation. However, projects should at least give an indication of the orders of magnitude of the economic benefits that might be expected. This of itself would be very useful, since typically industrial training has not been subjected to economic analysis (even in developed countries), and it is not unreasonable to believe that very substantial returns to initial investments in training can be realised.

## "Benefits which may be measured"

1. Economic objectives. Group C clearly advocates the selection of projects with explicit benefits. But of course more implicit ones may have to be included if they are considered significant enough.

2. Economic spill-over.benefits. This is an attempt to use relatively explicit benefits to capture and reflect the more implicit ones. Thus any general increase in productivity or reduction of labour turnover (other than that specifically designed for) might reflect implicit benefits such as increases in motivation, morale, attitudes to work, safety standards and efficiency of communications.

Initially, at least, benefits such as general increases in productivity will tend to be somewhat implicit in the sense that they cannot be predicted and can only be determined ex posta ( see Rule 1 ). However, field studies and/or experience may render them more predictable.

3. Behavioural objectives. These can be useful when (a) it is difficult to measure the final economic outcome, and (b) when the final economic outcome depends on many other factors as well as the training input. For example, the effect of the correct use of fertilizers on crop production: the behavioural objective would be the adoption of the correct methods, while the economic objective would be an increase in the value of the crops produced. The behavioural objective will often be much easier to deal with.

The applicability of the approach will depend on the technological knowledge in the field concerned, i.e. the relationship between the adoption of a technique and its subsequent effects on output under various conditions: for example, that under certain climatic and soil conditions, the correct use of a fertilizer will increase a certain crop production by x%. Given that the behavioural objectives have been achieved, the uncertainty now depends on the other conditions (e.g. weather) and again a decision must be made as to the acceptable level of risk. The approach does seem to be applicable to agriculture, if the technology exists.

4. Non-monetary objectives. These are clearly implicit, though not necessarily small or unimportant. In passing, it can be noted that almost any possible effect of an educational project could conceivably have some (implicit) economic benefit - care must be taken not to get bogged down in attempting to account for too many.

5. Experience of course is an invaluable method of rendering benefits more explicit.

## Cost-effectiveness analysis

When the objectives of a project have been specified, the next exercise is to determine the most economic method of achieving them. This is cost-effectiveness analysis: the determination of the least cost combination of instructional methods and media that can attain the given educational objective. In a literacy project there will be many possible combinations: television, films, radio + slide projectors, radio + correspondance courses, slide-tape presentations, tape recorders + workbooks, live instructors + handouts, or any combination of those.

It is not immediately obvious, for example, how the live instructor compares with other media. To be sure the superficially trained 12th grader may be the cheapest, but perhaps at the expense of not fully achieving the anticipated benefits on which the whole economic justification of the project rests. And the extra training the instructor may require to bring him up to the required level of effectiveness may render him less economic than other media systems.

The design for efficient instructional systems demands a knowledge of educational psychology, media technology (including relative costs) and a fairly intimate knowledge of the characteristics of the population being taught. It cannot be expected that the same system will be appropriate for different populations or countries.

Cost-effectiveness analysis can be used to investigate one important explicit benefit of the (simple) literacy component of functional literacy - that the simple literacy will reduce the costs of vocational training. Literate persons may tend to learn quicker, or require less expensive instructional systems; for example cheaper written expositions (manuals, workbooks, pamphlets) may replace more expensive aural expositions by radio, tape recorder or live human instructor.

## Cost benefit analyses and decision making

Analyses are only concerned with the economic component of a decision - the return on the investment compared to other investments open to the government department, firm, community or country. However, most educational decisions will have non-economic, social and political components, and the decision maker must decide on the weights to be given to these various components. Presumably there are worthwhile social benefits of literacy that will have costs without explicit economic benefits.

If projects can be found where the explicit economic benefits exceed the costs, there is no problem. If, however, there is a 'deficit' the decision maker must determine if it can be compensated for by his subjective assessment of the worth of the non-economic benefits plus his, again subjective, assessment of the value of the implicit economic benefits. In any event, he will have more information on which to base his decisions than he did before.

Suggested action

1. An inquiry, as a preliminary to possible cost benefit analyses, to determine:
    a. what projects can be found that could have explicit benefits, or benefits that could be rendered explicit.
    b. what generalisable results might be expected from such projects.
2. An inquiry, as a preliminary to possible cost-effectiveness analyses, to determine what instructional media systems might be used, and to give an indication of the relative costs.

## CHAPTER 13

# THE UNESCO MANUAL ON THE EVALUATION OF EXPERIMENTAL LITERACY PROJECTS

The Workshop devoted a full day to consideration of the Unesco Manual on the Evaluation of Experimental Literacy Projects. This was greatly assisted by an introductory address on the Manual given by Dr. Saksena of the Unesco Department of out-of-school Education.

## ADDRESS BY DR. SAKSENA

The evaluation exercise in the field of functional literacy drew upon lengthy experience in two broad areas of educational activity - the school system with its long tradition of testing and measurement: and agricultural extension which had also evolved its own procedures aimed at primarily measuring the amount of new agricultural knowledge absorbed by farmers, and subsequently at measuring social and behavioural patterns related to the attitudes of rural dwellers to change. As the areas with which those procedures were concerned widened, so measurement techniques became more sophisticated and the exercise itself more difficult.

The aim of the Unesco Manual on the Evaluation of Experimental Literacy was to assist those operating in the field to prepare their own local evaluation design. It would be necessary for manuals to be devised to meet every particular local situation. The Unesco Manual aimed to present ideas which might assist in preparing a local manual and to suggest areas where thinking should be applied in the preparation of such a manual.

The Manual was not to be regarded as a job description for an evaluation specialist. It sought to supply ideas for the entire sphere of evaluation activities throughout a project as well as providing practical guidelines for the specialist.

The distinction between action-oriented research and evaluation was admitted. In many developing countries there was an understandable frustration with long-term measurement procedures which necessarily had a delayed usefulness. The new view of continuing evaluation linked to feed-back which would modify a programme was attractive. There had, however, been some doubts on the part of purist professional evaluators who felt that the effect of this type of research would be to distort the strictly scientific exercise of evaluation. At the other extreme, there was a view that scientific evaluation was an unnecessary luxury which could well be dispensed with and might be usefully replaced by an action research unit which would feed in relevant data arrived at through fairly scientific procedures which would have the immediate effect of bringing about improvements in the operation of programmes. The Manual takes up an intermediate position, suggesting that it was necessary:

a. initially for background data to be collected systematically for subsequent evaluation purposes and for programme planning;

b. to apply techniques of social research in order to locate those areas in societies where change through functional literacy could be expected;

c. to evolve measurement tools in order to locate where and how changes were occurring.

These techniques might usefully serve both for overall evaluation and for action-oriented research.

It was hoped that the Workshop would give a guide to Unesco on the role of evaluation within functional literacy programmes. Having decided on the role, then the final decisions on design could be taken in the light of the situation within a particular country.

## DISCUSSION ARISING FROM DR. SAKSENA'S ADDRESS

The Workshop distinguished between three areas of activity that were dealt with in the Manual:

1. the evaluation of the effect of functional literacy on development;

2. special studies, such as the evaluation of specific methods and the comparative evaluation of alternative methods and media;

3. action-oriented research to provide data for project planning and to improve on-going programmes: e.g. feasibility studies, pre-testing media etc.

It was agreed that the first of these should have the highest priority but that all were important. This priority was justified by the requirements of the financial sponsors of projects, whether within the Experimental World Programme or outside it. The financiers, bankers and administrators, who had provided the funds which enabled the projects to be launched, were understandably interested in receiving information on whether their investment was achieving its ends. In this case it had to be proved whether functional literacy was indeed a positive factor in development. This proof could only be obtained through the application of systematic social research methodology. Secondly, there was growing dissatisfaction with the personal and impressionistic assessments which had hitherto often posed as evaluation. The techniques of social science research could obviate the use of such imprecise methods.

The Workshop endorsed the Manual's definition of the term 'evaluation' as distinct from 'inspection','assessment' or'appraisal'. It was suggested, however, that a clearer distinction might be made between evaluation and action research (action-oriented research), the former aiming "to measure....the major direct and indirect effects of a certain activity," the latter to feed back data to the operational staff to help them to improve their day-to-day operations.

There was then an interesting discussion as to whether action-oriented research should be entrusted to a separate small action research unit within the operational staff or should be a function of the evaluation unit. This ended in agreement that, although the aims of action research and evaluation might be recognisably different, both activities should generally be entrusted to a single unit, which might more appropriately be called the Research and Evaluation Unit of the project.

It was agreed that evaluation might most usefully focus upon the individual participant and the work unit. The Workshop expressed doubts about the possibility of evaluating the effects of functional literacy on a community of any considerable size and ruled out the possibility of evaluation at the national level. The need for background socio-economic studies at the national level was, therefore, called in question. It was explained that this particular exercise was not of high priority, that its use was merely to provide an understanding of the background to the project and that it would be done in the beginning when pressures on the evaluation unit might be expected to be less. It was pointed out that in many cases the kind of information required in such a study was already likely to be available in other documents. The exercise had no value for measurement.

The Workshop then discussed the work load of the evaluation unit, as described in the manual. There was a strong feeling that if evaluators, working in small teams with tight budgets, were required to tackle all recommended activities, then disillusion might result. The main criticism of the manual was that, in setting out such a comprehensive view of evaluation, special studies and action research, its suggestions might appear to be over-ambitious and difficult to carry out with the staff and budget normally available to a functional literacy project. It seemed desirable, therefore, to distinguish more clearly what were the absolute essentials for effective evaluation of functional literacy at the levels of the individual participant and the work unit or small local community. In particular it was considered unrealistic to expect that within the 'preparatory period' - a year from the signing of the Agreement for the project - the following procedures could be completed: the recruitment and placement of evaluation staff, background studies, problem surveys, project planning, technical feasibility studies, the definition of objectives in measurable terms, the selection of indicators, the determination of samples and controls and the base-line surveys. It was explained that in fact the base-line survey could not be started until participants were selected, that is until the operational period had begun. It was agreed that this should be clarified in the Manual and, in particular, that the diagram on page 5(b) of the Manual needed to be redrawn accordingly.

In recognition of what was even so an overwhelming workload, it was suggested that more time should be demanded for these preparatory activities, despite external pressures for quick results to be realised and demonstrated; otherwise project operations would begin before the completion of these procedures with resultant adverse effects on the work of evaluation.

It was explained that the one year preparatory period had been laid down by agreement between Unesco and UNDP but that the activities for the evaluation unit enumerated for the first year were not to be regarded as definitive. Many of these activities were designed to enable evaluation personnel to collect background data whilst gathering experience and understanding of the project area. It should be noted that the evaluation specialist was generally one of the earliest appointments in any project team. In any case, firm decisions on what should be the actual activities of the evaluation unit would be taken at the level of the project.

With regard to cost benefit analysis, it was considered doubtful whether such procedures could be applied rigorously in a functional literacy project. Governments, however, wanted to know the kinds of economic returns they were receiving from their investment in functional literacy. It was, therefore, suggested that a simplified approach to cost benefit analysis should be tried which would give a descriptive account of costs and known benefits. In particular, it was important to try to compare the costs and benefits of integrated functional literacy programmes with those of traditional literacy programmes.

In conclusion, it was recognised that whatever detailed criticisms were made of the Manual, these were due, in large part to the fact, in itself a merit, that it was so comprehensive and adaptable. There was indeed general approval of its content and full agreement as to its outstanding value as a general guide to those concerned with evaluation and research in functional literacy.

# APPENDIX A

## LIST OF PARTICIPANTS AND STAFF

### Participants

Name: MOHAMMAD MEHDI BEHKISH

Home Address:
Vaziri Hotel,
Isfahan,
Iran.
Tel. 2992

Official Address:
P.O. Box 103,
Isfahan,
Iran.
Tel. 5745

Present Post: Evaluation Specialist - Work Oriented Adult Literacy Pilot Project.

Functions: Research and evaluations.

Specialisation: Economics.

--------------------

Name: ABRAHA GHERMAZION

Home Address:
Unesco Functional Literacy Project,
Box 60006,
Addis Ababa,
Ethiopia,
Tel. 45081 or 45553

Present Post: Evaluation Counterpart.

Functions: Plan, organise and administer functions relevant to evaluation with the aid and co-operation of Dr. B. N. Singh.

Specialisation: Psychology and Elementary Education.

Name: HASSANIN ABD. EL MOATY HASSANIN

Home Address: 2 Mohamed Talat Street,
Agouza,
Cairo,
U.A.R.

Official Address: Ministry of Education,
Adult Education Department,
Cairo,
U.A.R.
Tel. Cairo 24722

Present Post: Member of Project and Foreign Affairs Section.

Functions: External Relations for the Project in the Field.

Specialisation: Audio-Visual Aids.

--------------------

Name: KHALAFALLA MOHED. IDRIS

Home Address: Getaina,
Khartoum,
Sudan.

Official Address: Work Oriented Adult Literacy Project,
Ministry of Education,
Khartoum,
Sudan.

Present Post: Counterpart of Unesco Evaluation Specialist.

Functions: Taking part in the Evaluation in Work Oriented Adult Literacy Projects.

Specialisation: Economics and Social Studies.

Name: EMMANUEL P.R. MBAKILE

Home Address:
Institute of Adult Education,
University College Dar es Salaam,
P.O. Box 20679,
Dar es Salaam,
Tanzania.
Tel. Dar es Salaam 20377

Official Address:
UNESCO/UNDP Work Oriented Adult Literacy Pilot Project,
Lake Regions,
P.O. Box 1141,
Mwanza,
Tanzania.
Tel. Mwanza 2004

Present Post/Functions: Assistant Research Fellow at the Institute of Adult Education, Dar es Salaam University College and National Counterpart to Unesco Evaluation Specialist. Work Oriented Adult Literacy Pilot Project, Lake Regions, Tanzania.

Specialisation: Sociology.

--------------------

Name: PROF. AURISELA ALVAREZ MEDINA

Home Address:
Residencias Tamanaco No.18,
Avenida Principal El Llanito,
Petare,
Caracas,
Venezuela.
Tel. 341187

Official Address:
(Eduplan) Oficina de Planeamiento Integral de la Educacion,
Ministerio de Educacion,
Caracas,
Venexuela.
Tel. 515917 - 515926 ext. 55

Present Post: Head of the Department of Documentation and Information (Eduplan)

Functions: Study of documents - establishment of planning priorities - participation on projects, etc.

Specialisation: Philosophy and Science of Education

Name: DR. MRS. SUMATI MULAY

Home Address: 1 Mahader Road,
New Delhi 1,
India.
Tel. 386470

Official Address: Department of Adult Education,
National Council of Educational Research and Training,
37 Friends Colony,
Mathura Road,
New Delhi 14.
Tel. 74511

Present Post: Reader (Assistant Director)

Functions: Social Research and Teaching

Specialisation: Sociology and Social Research

--------------------

Name: MRS. OMAIMA AHMED AL NAHRY

Home Address: 188 El-Nil Street,
Agouza,
Cairo,
U.A.R.
Tel. Cairo 801337, Alexandria 22406 (Summer)

Official Address: Ministry of Education,
Adult Education Department,
Cairo,
U.A.R.
Tel. Cairo 24722

Present Post: Chief of Projects and Foreign Affairs Section

Functions: External Relations for the Project in the Field.

Specialisation: Primary Education - Adult Education

Name: PROFESSOR AYO OGUNSHEYE

Home Address:
University of Ibadan,
Ibadan,
Nigeria.
Tel. Ibadan 21051 Ext. 432

Official Address:
Department of Adult Education,
University of Ibadan.
Ibadan,
Nigeria.
Tel. Ibadan 21051 Ext. 548

Present Post: Professor of Adult Education, Head of Department and Director, Institute of Adult Education, University of Ibadan.

Functions: Professing adult education, heading the Department and Institute, teaching international trade. Research.

Specialisation: Adult Education - International Trade.

--------------------

Name: FAREL SITANGGANG

Home Address:
Djl. Rawabilal III no.6 (persodjo)
Tebet-Barat,
Jakarta,
Indonesia.

Official Address:
Direktorat Pendidikan Masjarakat,
Dep. P. dan K.,
Djl. Salemba Raya No.18,
Djakarta,
Indonesia.
Tel. 81483

Present Post: Chief, Educational Planning and Development Section, Directorate of Community Education, The Ministry of Education and Culture.

Functions: Planning work on Adult Education. Adult Education Development and Improvement.

Specialisation: Public Administration - Publication.

Name: KAMOL SUDAPRASERT

Home Address:
40/9 Soi Swasdi,
Prachasongkroh Road,
Payathai District,
Bangkok 10,
Thailand.

Official Address:
Department of Elementary and Adult Education,
Ministry of Education,
Bangkok,
Thailand.
Tel. 816400

Present Post: Educational Supervisor: Head of Educational Measurement and Evaluation Service.

Functions: Evaluation and Research concerning elementary and adult education.

Specialisation: Research and Measurement.

---

Name: ABDULKADIR SULAIMAN

Home Address:
11/6/20 Al-Mansour,
Baghdad,
Irak.
Tel. 510264

Official Address:
Specialist Inspector,
Ministry of Education,
Baghdad,
Irak.
Tel. 25010

Present Post: Specialist Inspector

Functions: Supervising and inspecting Teachers' Training Colleges and Institutes.

Specialisation: Education and Psychology

Name: MRS. BEATRIZ LIDIA VAZQUES FULLER

Home Address: Calle Venezuela No. 24-48,
Quito,
Ecuador.
Tel. 236544

Official Address: Proyecto Piloto de Alfabetizacion Funcional,
Ministerio de Educacion Publica,
Quito,
Ecuador.
Tel. 219431.

Present Post: Technical Education Programmer.

Functions: Counterpart to the Expert on Evaluation responsible for the Research and Evaluation Unit.

Specialisation: Social Services, Psychological Education, Fundamental Education

---

Name: DR. GOPAL S. VIDYARTHY

Home Address: DI/92 Chanakyapuri,
New Delhi 21,
India.
Tel. 79446

Official Address: Director of Extension,
Ministry of Food, Agriculture, C.D. and Coop.,
Department of Agriculture,
Government of India.
Tel. 384180

Present Post: Director of Extension

Functions:
1. Formulate programmes of Extension and Training at central level.
2. Implement Farmers Training and Education Programme.
3. Programmes of Home Science, Youth and Personnel.

Specialisation: Extension Education - Rural Sociology - Economic Development - Training and Research.

## Staff and Lecturers

### Full-time Staff

MR. JOHN BOWERS,
Workshop Director,
University of London, Institute of Education,
Malet Street,
London, W.C.1.

MR. J. DI FRANCO,
Agricultural Extension Specialist,
Rural Institutions Division,
F.A.O.
Rome.

DR. K. NEYS,
Evaluation Specialist,
Venezuela Functional Literacy Project,
Apartado 1969,
Caracas,
Venezuela.

MR. A. W. WOOD,
Education Research Officer,
Commonwealth Secretariat,
Marlborough House,
Pall Mall,
London, S.W.1.

### Lecturers and Consultants

PROFESSOR M. BLAUG,
Head of Research Unit in the Economics of Education,
University of London Institute of Education,
45 Russell Square,
London W.C.1.

MR. J. C. CAIRNS,
Director,
Literacy Division,
Department of Out-of-School Education,
UNESCO,
Paris, 7e.

MR. A. B. EDINGTON,
Assoc. Director of Information and Research,
CETO,
Nuffield Lodge,
Regents Park,
London, N.W.1.

DR. RICHARD JOLLY,
Senior Research Fellow,
Institute of Development Studies,
University of Sussex,
Stanmer,
Brighton,
Sussex.

MR. A. A. MITCHELL,
Head of Industrial Instructional Staff,
National Coal Board,
Graham House,
Longbenton,
Newcastle-upon-Tyne,

MR. M. OATEY,
Personnel Research Officer,
B.E.A.,
Douglas House,
Ruislip,
Middlesex.

DR. H. P. SAKSENA,
Programme Specialist,
Department of Out-of-School Education,
UNESCO,
Paris, 7e.

MR. G. A. SMITH,
Training Superintendent,
Pressed Steel Fisher Ltd.,
Cowley,
Oxford.

MR. D. E. WHEATLEY,
Director of Further Education and Training,
City and Guilds of London Institute,
76 Portland Place,
London, W.1.

DR. R. WIENER,
Research Officer,
London School of Economics and Political Science,
University of London,
Skepper House,
13 Endsleigh Street,
London, W.C.1.

## APPENDIX B

## WORKSHOP PROGRAMME

### Sunday 3rd August

18.30 - 19.15 - Evening meal at John Adams Hall
20.00 - Reception at John Adams Hall. Election of Workshop Steering Committee.
21.30 - Steering Committee meeting to consider draft programme.

### Monday 4th August

09.30 - 11.00 - Opening address and discussion. Functional Literacy and Evaluation: Aims of the Workshop (John Bowers).
11.30 - 12.30 - Administrative Arrangements and Programme (Freddie Wood).
14.30 - 16.00 - Experiences and Problems in the Evaluation of Functional Literacy (Karel Neys).
16.30 - Meeting of Steering Committee.

### Tuesday 5th August

09.30 - 11.00 - Experiences and Problems in Evaluation - statements by participants.
11.30 - 12.30 - Evaluation of Rural Extension Programmes (J. di Franco).
14.30 - 16.00 - Statements by participants (Contd.)
16.30 - 17.30 - Discussion of statements.

### Wednesday 6th August

09.30 - 11.00 - Evaluation of Functional Literacy: An Economist's View (Prof. M. Blaug).
11.30 - 12.00 - Discussion
14.00 - 16.00 - Seminar with Prof. Blaug

### Thursday 7th August

09.30 - 11.00 - Experiences and Problems in Evaluation - statements by participants.
11.30 - 12.30 - Techniques of Social Study for Evaluation (Dr. R. Wiener).
14.30 - 16.00 - Statements by participants.
16.30 - 17.30 - Meeting of Steering Committee to form working groups for second week's programme.

### Friday 8th August

09.30 - 10.50 - Combined meeting with participants in training course on audio-visual media at Overseas Visual Aids Centre (OVAC), Tavistock House South, Tavistock Square, W.C.1. (talk by J. Bowers and discussion on Action Research and Media Production).

11.15 - Continuing visit to OVAC exhibition and discussion with OVAC staff on media for functional literacy.

14.30 - 16.00 - Formulation of following week's programme for working groups.

16.30 - 17.30 - Meeting of groups for preparatory discussion of activities for following week.

### Saturday 9th August

- Morning free.

14.15 - Tour of London.

### Sunday 10th August

Boat trip to Richmond and Hampton Court. Lunch with John Bowers at "Cornerways", Petersham.

### Monday 11th August

10.00 - 11.00 - The Role of Cost Analysis in the Evaluation of Functional Literacy (Dr. R. Jolly).

11.30 - 12.30 - Discussion with Dr. Jolly.

14.30 - 16.00 - Working groups.

16.30 - 17.30 - Resumed.

### Tuesday 12th August

09.30 - 11.00 - Unesco's Programme in Functional Literacy and Evaluation (mr. J. C. Cairns).

11.30 - 12.30 - Working groups.

14.00 - 16.00 - Working groups.

16.30 - 17.30 - Seminar with Mr. Cairns.

**Wednesday 13th August**

10.30 - 12.30 - Visit to City & Guilds of London Institute, Examinations, Testing and Development Unit.
14.30 - 16.00 - Talk by Mr. A. Mitchell of the National Coal Board on Evaluation of Low-Level Vocational Training in the Coal Industry and discussion.
16.30 - 17.30 - Discussion continued.

**Thursday 14th August**

09.30 - 11.00 - Working groups.
11.30 - 12.30 - Full session. Interim progress reports from working groups.
14.30 - Visit to Centre for Educational Television Overseas (CETO).

**Friday 15th August**

Visit to Oxford
11.00 - Pressed Steel Fisher Ltd. - Operative Training and Evaluation in the Motor Industry.
15.00 - Tour of Oxford.
Overnight at Elms Hotel, Iffley.

**Saturday 16th August**

In and around Oxford.
16.00 - Return to London.

**Sunday 17th August**

Free.

**Monday 18th August**

09.30 - 11.00 - Working groups finalise reports.
11.30 - 12.30 - Working groups finalise reports.
14.30 - 16.00 - Working group 'A' presents its report.
16.30 - 17.30 - Discussion of report of working group 'A'

**Tuesday 19th August**

09.30 - 11.00 - Working group 'B' presents its report.
11.30 - 12.30 - Discussion of report of working group 'B'.
14.30 - 16.00 - Working group 'C' presents its report.
16.30 - 17.30 - Discussion of report of working group 'C'.

**Wednesday 20th August**

09.30 - 11.00 - Consideration of Unesco Manual on Evaluation of Experimental Liter Literacy Projects.
11.30 - 12.30 - Continued.
14.30 - 16.00 - Continued.
16.30 - 17.30 - Continued.

**Thursday 21st August**

09.30 - 11.00 - Oral presentation of provisional report of Rapporteur (Freddie Wood).
11.30 - 12.30 - Discussion of report.
14.30 - 16.00 - Evaluation of workshop by participants and staff.
16.30 - Closing session.

**Friday 22nd August**

Participants disperse.

ERIC Clearinghouse
OCT 27 1970
on Adult Education