

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

Project 4.32
October 1982
Document B

AN EXPLORATORY STUDY ON THE EVALUATION OF LEARNING OUTCOMES
AND LARGER IMPACT OF LITERACY, POST-LITERACY AND CONTINUING
EDUCATION PROGRAMMES IN DEVELOPING COUNTRIES

Planning Meeting

Hamburg, 11 - 15 October, 1982

TENTATIVE DAILY SCHEDULE



Sessions:

Morning Session: 09.00 - 12.30
Coffee Break: 10.30 - 11.00
Afternoon Session: 14.00 - 18.00
Coffee Break: 16.00 - 16.30

Daily Programme

* Monday, 11 October

Morning Session:

- (i) Welcome
- (ii) Introduction by participants
- (iii) Comments on the project
 - Background
 - Objectives
 - Major areas of study:
 - learner evaluation
 - curriculum evaluation
 - programme monitoring
 - impact evaluation
- (iv) Brief presentation (about 10 mins.) and discussion (about 20 mins.) of evaluation practices in each of the participating countries

Brazil	Indonesia	Niger
Cuba	Iraq	Sri Lanka
Ethiopia	Kenya	Tanzania
India	Mali	Thailand

Afternoon Session:

- (i) Brief presentation and discussion of evaluation practices in participating countries - continued

* Please hand over the note on the evaluation practices in your country to Mr D A Perera at 9.00 am on 11 October.

Tuesday, 12 October

Morning Session:

- (i) Brief presentation and discussion of evaluation practices in participating countries - continued
- (ii) Identification of illustrative elements of the major areas of study for the purpose of preparing an overall draft research design and conducting national case studies
- (iii) Preparation for group work

Afternoon Session:

- (i) Group work - preparation of a draft design for a descriptive-cum-analytical study of the following areas:

learner evaluation
curriculum evaluation
programme monitoring
impact evaluation

Wednesday, 13 October

Morning Session:

- (i) Discussion of draft designs

Afternoon Session:

- (i) Discussion of draft designs - continued
- (ii) Individual work / Free Time (from 15.30 h. onwards)

- DR. DAVE dinner

Thursday, 14 October

Morning/Afternoon Sessions:

- (i) Discussion of draft designs - continued
- (ii) Group work - revision of draft designs
- (iii) Discussion of amended designs

Friday, 15 October

Morning/Afternoon Sessions:

- (i) Discussion of amended designs - continued
- (ii) Preparation of a time schedule for further work
- (iii) Other organisational matters
- (iv) Conclusion of the meeting

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

Project 4.32
October 1982
Document A

AN EXPLORATORY STUDY ON THE EVALUATION OF LEARNING OUTCOMES
AND LARGER IMPACT OF LITERACY, POST-LITERACY AND CONTINUING
EDUCATION PROGRAMMES IN DEVELOPING COUNTRIES

Planning Meeting

Hamburg, 11 - 15 October, 1982

PROVISIONAL LIST OF PARTICIPANTS



Brazil

Miss Ana Maria Coutinho

Special Advisor to the President
Ministry of Education and Culture
Brazilian Literacy Movement Foundation/
MOBRAL
SCRLN 704/5 - Bloco H - Lojas 33/34
70730 Brasilia - DF

Cuba

Dr Caballero Carrera

c/o Comisión Nacional Cubana de la Unesco
Ave. Kohly No. 151, esq. 32
Nuevo Vedado
La Habana

Ethiopia

Mr Abebe Berhanu

c/o Ethiopian National Agency for Unesco
P O Box 2996
Addis Ababa

India

Mr R S Mathur

Deputy Director
Government of India
Ministry of Education and Culture
Directorate of Adult Education
34 Community Centre
Basantlok
Vasant Vihar
New Delhi 110057

Indonesia

Mr W P Napitupulu

Director General of Non-formal Education,
Youth and Sports
Ministry of Education Culture
Jalon Halimun 4
P O Box 3290
Jakarta

*not present
in Hamburg
- will join the team*

Iraq

Dr Abdul Aziz Al-Bassam

c/o The Arab Educational, Cultural
and Scientific Organization (ARLO)
113 Abu Nawas Street
Baghdad

Kenya

Mr David Macharia

Director
Ministry of Culture and Social Services
Department of Adult Education
P O Box 30117
Nairobi

Niger

Dr Kané Oumarou

Directeur
Centre de Formation des Cadres de
l'Alphabétisation (CFCA)
Ministère de l'Education Nationale
B P 525
Niamey

Tanzania

Mr Frank Elinewinga

Deputy Director
National Mwanza Literacy Centre
P O Box 1141
Mwanza

Thailand

Mr Nished Suntornpithug

Deputy Director-General
Non-Formal Education Department
Ministry of Education
Bangkok 10300

UIE Staff

+Dr Ravindra H Dave (Director)
Mr Paul Fisk
Professor Gottfried Hausmann
Dr Gert Loose
Mr Eizo Nagasaki
Dr Adama Ouane
Mr Daya A Perera
Mr Friedhelm Zanter

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany



Project 4.32
October 1982
Document A

AN EXPLORATORY STUDY ON THE EVALUATION OF LEARNING OUTCOMES
AND LARGER IMPACT OF LITERACY, POST-LITERACY AND CONTINUING
EDUCATION PROGRAMMES IN DEVELOPING COUNTRIES

Planning Meeting

Hamburg, 11 - 15 October, 1982

REVISED LIST OF PARTICIPANTS

Brazil

Miss Ana Maria Coutinho

Special Advisor to the President
Ministry of Education and Culture
Brazilian Literacy Movement Foundation/
MOBRAL
SCRLN 704/5 - Bloco H - Lojas 33/34
70730 Brasilia - DF

Cuba

Dr Hermes Caballero Carreras

Asesor del Vicepresidente
Ministerio de Educación
Obispo 160
La Habana

Ethiopia

Mr Abebe Berhanu

Head, Research and Training Service Division
Educational Mass Media Department
Ministry of Education
P O Box 3025
Addis Ababa

India

Mr R S Mathur

Deputy Director
Government of India
Ministry of Education and Culture
Directorate of Adult Education
34 Community Centre
Basantlok
Vasant Vihar
New Delhi 110057

Iraq

Dr Abdul Aziz Al-Bassam

Member of the Iraq Academy
c/o The Arab Educational, Cultural
and Scientific Organization (ARLO)
113 Abu Nawas Street
Baghdad

Kenya

Mr David Macharia

Director
Department of Adult Education
Ministry of Culture and Social Services
P O Box 30117
Nairobi

Niger

Dr Kané Oumarou

Directeur
Centre de Formation des Cadres de
l'Alphabétisation (CFCA)
Ministère de l'Education Nationale
B P 525
Niamey

Tanzania

Mr Frank Elinewinga

Deputy Director
National Literacy Centre
P O Box 1141
Mwanza

Thailand

Dr Nished Suntornpithug

Deputy Director-General
Non-Formal Education Department
Ministry of Education
Bangkok 10300

UIE Staff

Dr Ravindra H Dave (Director)
Mr Paul Fisk
Professor Gottfried Hausmann
Mr Eizo Nagasaki
Dr Adama Ouane
Mr Daya A Perera
Mr Friedhelm Zanter

UNESCO-INSTITUT FÜR PÄDAGOGIK
UNESCO INSTITUTE FOR EDUCATION
INSTITUT DE L'UNESCO POUR L'EDUCATION

12

Miss Ana Maria Coutinho
Advisor to the President
Fundacao Movimento Brasileiro de
Alfabetizacao (MOBRAL)
Ministério da Educacao e Cultura
Rua Voluntários da Pátria 53
22.270 Rio de Janeiro



FELDBRUNNENSTRASSE 58
2000 HAMBURG 13 F.R.G.

Brazil

YOUR REF.

YOUR LETTER OF

OUR REF. RHD/qs
PRG/4.32

DATE 30 July 1982

Dear Miss Coutinho,

1. I am happy to inform you that the Unesco Institute for Education (UIE) is planning to launch an exploratory study on the Evaluation of Learning Outcomes and Larger Impact of Literacy, Post-literacy and Continuing Education Programmes in Developing Countries (PRG 4.32). The purpose of this letter is to invite you to participate in this research project.
2. As you may be aware, UIE has been working for the past one decade on the concept of lifelong education and its implications for curriculum development, evaluation and teacher education including both formal and nonformal systems of learning. In order to apply the principles of lifelong learning to some of the educational problems of developing countries, UIE launched a research project on the Development of Learning Strategies for Post-literacy and Continuing Education of Neo-literates (both adults and out-of-school children). A research-based international orientation seminar was organised on the above subject in 1981. The participants of this seminar unanimously expressed their view that much help was needed in the area of monitoring and evaluation of literacy and post-literacy programmes, and that UIE should undertake a project on exploring and disseminating techniques of evaluating learning outcomes, larger impact, etc. Thus, following this and other similar observations UIE has decided to mount an exploratory study at the outset.
3. It has been proposed that the study may include:
 - (i) Learner evaluation,
 - (ii) Programme monitoring, and
 - (iii) Impact evaluation.

At its exploratory stage, a series of case studies should be carried out including current practices and innovative procedures concerning

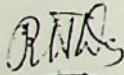
../2

one or more of the above three dimensions of monitoring and evaluation of literacy, post-literacy and continuing education. The current practices and innovative efforts should be examined in several countries in order to identify strengths and weaknesses, and to evolve suitable techniques, instruments and procedures relevant to the local conditions and resources. This will be done in some 8 to 10 countries including yours.

4. You have done a considerable amount of work in this area and therefore, as stated above, we wish to invite you to participate in this important project and enable us to share your experiences with other colleagues working in different parts of the world. The study will chiefly include (i) a planning meeting of all participants concerned, (ii) a national level case study to be carried out by you in your country after the planning meeting, and (iii) a review meeting to discuss national level reports and develop an international synthesis as well as possible publication. The working language of the project will be English. I request you to let me know as early as possible, and latest by 25 August, 1982 whether you will be able to participate in this project, including the planning and review meetings.
5. Assuming that your reply will be positive, I extend an invitation to you to attend the planning meeting to be held at UIE, Hamburg from 11 to 15 October 1982. The cost of your air travel from your place to Hamburg and back (economy class) and a per diem of DM 153,00 for the duration of the meeting will be covered by UIE according to its rules. On the completion of the case study a token remuneration of DM 750,00 will also be paid to you by UIE.
6. On hearing from you, I shall send further information and materials concerning the project and the planning meeting.

Looking forward to your kind co-operation and a positive reply at your earliest,

Yours sincerely,



Ravindra H Dave
Director

P.S. We very much wish that you participate in this project of international significance yourself. If, however, you are not in a position to do so on account of totally unavoidable circumstances, may I request you to nominate another person from your organisation who has sufficient experience in this field and in preparing an effective case study.

UNESCO-INSTITUT FÜR PÄDAGOGIK
UNESCO INSTITUTE FOR EDUCATION
INSTITUT DE L'UNESCO POUR L'EDUCATION

MOBRAL 17
PROTOCOLO
Nº - 2534/12
DATA - 14-09
BRASILIA - DF



Miss Ana Maria Coutinho
Advisor to the President
Fundacao Movimento Brasileiro de
Alfabetizacao (MOBRAL)
Ministério da Educacao e Cultura
Rua Voluntários da Pátria 53
22.270 Rio de Janeiro
Brazil

FELDBRUNNENSTRASSE 58
2000 HAMBURG 13 F.R.G.

YOUR REF.

YOUR LETTER OF

OUR REF.

DATE

RHD/rk
PRG 4.32

September 6, 1982

Dear Miss Coutinho,

Many thanks for your letter of August 16, 1982 accepting our invitation for your participation in an exploratory study on the Evaluation of Learning Outcomes and Larger Impact of Literacy, Post-literacy and Continuing Education Programme in Developing Countries (PRG 4.32). We are extremely happy that you will be able to participate in this important project.

As regards the preparation for this programme, I request you to collect information and materials in respect of the following three aspects of monitoring and evaluation as mentioned in my letter of invitation:

- (i) Learner evaluation,
- (ii) Programme monitoring, and
- (iii) Impact evaluation.

This may be done on the basis of the work done by MOBRAL and other similar agencies working in this field. Further information in this regard will also be sent to you in due course.

Travel arrangements have now been made for you as follows:

9.10.	Rio de Janeiro	19.40	LH 507
	(International)		
10.10	Frankfurt	10.50	
	Frankfurt .	12.45	LH 765
	Hamburg	13.50	

.... 2

<u>15.10</u>	Hamburg	19.30	LH 023
	Frankfurt	20.35	
	Frankfurt	22.00	LH 506
16.10	Rio de Janeiro	5.45	

You are requested to contact the Lufhansa Office in Rio de Janeiro to obtain the ticket.

A single room with shower has been reserved for you in the following hotel from 10. to 15. October 1982 (5 nights):

Baseler Hospiz
Esplanade 11
2000 Hamburg 36
Tel. 34 19 21

With best wishes, *and looking forward to seeing you in October,*

Yours sincerely,



Ravindra H Dave
Director

UNESCO-INSTITUT FÜR PÄDAGOGIK
UNESCO INSTITUTE FOR EDUCATION
INSTITUT DE L'UNESCO POUR L'EDUCATION



Miss Ana Maria Coutinho
Advisor to the President
Fundacao Movimento Brasileiro de
Alfabetizacao (MOBRAL)
Ministério da Educacao e Cultura
Rua Voluntários da Pátria 53
22.270 Rio de Janeiro - Brazil

FELDBRUNNENSTRASSE 58
2000 HAMBURG 13 F.R.G.

YOUR REF.

YOUR LETTER OF

OUR REF. RHD/LS/4.32

DATE 20 September 1982

Dear Miss Coutinho,

I hope you have received by now the following eight reference documents sent to you a few weeks ago by airmail:

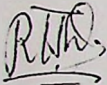
- | | |
|-------------------|---|
| <u>Document 1</u> | A Built-in System of Evaluation for Reform Projects and Programmes in Education
by Ravindra H Dave |
| <u>Document 2</u> | Reform Evaluation
by E Ayotunde Yoloye |
| <u>Document 3</u> | A Self Evaluation of the National Literacy Programme in Botswana - Second Outline
by Ulla Kann |
| <u>Document 4</u> | Appraisal Studies of the Adult Education Programme in Bihar, Gujarat, Maharashtra, Rajasthan and Tamil Nadu
by RS Mathur / Prem Chand |
| <u>Document 5</u> | Feedback Action - The Essence of Monitoring
by RS Mathur |
| <u>Document 6</u> | Learner Evaluation Report of National Workshop on Monitoring and Evaluation (June, 10-15, 1982) organised by Directorate of Adult Education, Ministry of Education and Culture, Govt. of India, New Delhi |
| <u>Document 7</u> | Conceptualizing Evaluation for Lifelong Education
by R Skager |
| <u>Document 8</u> | Outline of the global evaluation of literacy within the OACV (Ground-nut and Crops operation), Kita, Mali 1978 |

The travel and hotel arrangements have also been made for you according to my letter of 6 September, 1982.

In order to obtain some preliminary information about the present position of evaluation of literacy and post-literacy programmes in your country, we wish to request you to prepare a brief note of about 5 to 8 pages. Suggested guidelines for preparing such a note and other pertinent details are given in a document enclosed herewith. Kindly see that your note is handed over to Mr D A Perera at UIE on Monday 11 October at 9.00 am as stated in the enclosure. Also please bring the relevant evaluation reports, samples of tests and certificates, and any other such materials with you as stated in my letter of 6 September, 1982, and hand them over to Mr Perera as well, which will be displayed in our conference room and will be used as reference in the course of our planning meeting. The first session of the meeting will begin at 9.15 am on 11 October.

Looking forward to seeing you at UIE,

Yours sincerely,



Ravindra H Dave
Director

Encl.

Project 4.32
October 1982

SUGGESTED GUIDELINES FOR PREPARING A NOTE ON THE
EVALUATION OF LITERACY AND POST-LITERACY PROGRAMMES

The guidelines given below are meant for preparing a brief note of about 5 to 7 pages on the position of (a) learner evaluation, (b) curriculum evaluation, (c) monitoring of programme implementation, and (d) evaluation of larger impact and long range effects of literacy and post-literacy programmes in your country. The note will primarily be used for exchanging information at the beginning of the planning meeting being organised from 11 - 15 October 1982 at UIE, Hamburg in connection with the research project on evaluation (PRG 4.32).

Please bring one typed copy to the meeting and hand it over to Mr D A Perera at UIE on Monday 11 October at 9.00 am for reproducing the required number of copies. We are going to distribute and discuss your note right at the initial stage of our planning meeting.

Sections

Estimated length in
A/4 size pages
single-spaced type

1. Introduction

1 page

- Population of the country
- Percentage of literacy of male, female and total
- Present position of literacy programmes
- Position of primary education (basic education) including number of years, enrolment ratio and drop-out rates. (This is important in the context of a two-pronged approach of universalization of primary education, and adult literacy and post-literacy programmes to eradicate illiteracy)
- Provision for educational television, radio, correspondence programmes, etc., if any, for continuing education at the primary and secondary stages of learning
- Is there a separate unit for evaluation work? If so, please indicate its organisation and number of persons working in it.

2. Learner Evaluation

A. Literacy stage

2 pages

- What is evaluated (3 Rs, any other objectives?)
- Periodical evaluation (including self-evaluation): How is it carried out, how often, how are the results used?
- Formal final evaluation, if any. How is it done, who does it, what is the pass percentage, etc.
- Any certificate given? If so, what is the value of the certificate?

B. Post-literacy stage

- Is there any formal evaluation of learners at this stage? If so, please give details. Is this evaluation connected with the primary school certificate in any manner?

3. Curriculum Evaluation

1 page

- Are literacy textbooks, charts, learning aids, post-literacy curriculum and materials, teaching and learning processes, etc. evaluated (a) before they are put to use on a large scale (pre-testing), and (b) after they are

Sections

Estimated length in
A/4 size-pages
single-spaced type

put to use on a large scale? If so,
please give details specially in
respect of formative evaluation.

4. Progress Monitoring

1 page

- If you have either a systematic or an informal system of monitoring the effective implementation of literacy and post-literacy programmes, please describe it by stating:

- What functions are monitored?
(Administrative, financial, pedagogical - specify)
- At what levels - ranging from the national to the local - are different functions monitored?
- What procedures and tools are used?
- How are corrective measures taken to improve effectiveness?
- Any other information on monitoring.

5. Impact Evaluation

1 page

- Whether the immediate or long-range impact of literacy and post-literacy programmes on the individual, family and community has been evaluated so far. If yes, please give information about

- criteria or indicators used
- procedures adopted
- one or two examples of findings
- any other significant matters.

If it is not done systematically, please describe any informal ways of impact evaluation carried out so far and results obtained.

6. Research on Evaluation

1 page

- Please indicate whether any research on tools, techniques and other aspects of evaluation has been carried out. If so, who carried it out and to what effect? One or two concrete examples may be given.

GROUP WORK



1. Composition of Groups

Group A: Dr Abdul Aziz Al-Bassam

Mr Abebe Berhanu

Mr Nixed Suntornpithug

Mr R S Mathur

Mr Frank Elinewinga

Dr Adama Ouane

Mr Eizo Nagasaki

Dr R H Dave

Group B: Dr Kané Oumarou

Dr Caballero Carrera

Miss Ana Maria Coutinho

Mr David Macharia

Prof Gottfried Hausmann

Mr D A Perera

Mr Friedhelm Zanter

Dr R H Dave

2. Meeting Places and Topics

Group A

Meeting place: Large Conference Room

Topics: *Introduction*
Learner Evaluation / Curriculum Evaluation

Group B

Meeting place: Small Conference Room

Topics: Programme Monitoring / Impact Evaluation

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 1



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOMES AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

A Built-in System of Evaluation for
Reform Projects and Programmes
in Education

by

Ravindra H Dave

A BUILT-IN SYSTEM OF EVALUATION FOR REFORM PROJECTS AND PROGRAMMES IN EDUCATION

RAVINDRA H. DAVE

Abstract – In a conventional system of planning and implementation of educational reforms, evaluation generally appears only at the end of the process. While such a summative evaluation is necessary, the author argues for a more effective use of evaluation procedures for efficient planning, implementation and assimilation of reform projects. In order to use monitoring and evaluation procedures for contributing substantially towards the optimization of immediate outcomes and long-range effects of any reform project or programme, it is essential to follow a built-in, comprehensive and development-oriented approach. Such an evaluation procedure should take into account both extra-educational factors and intra-educational variables that influence the processes and outcomes of an educational reform. An EIPOL grid which combines five major dimensions of a broad-based evaluation system with different steps of a project cycle provides a basic operational framework for designing and adopting a more functional system of reform evaluation.

Résumé – Dans les systèmes conventionnels de planification et mise en oeuvre de réforme de l'éducation, l'évaluation se trouve généralement à la fin du processus. Bien qu'une telle évaluation sommaire soit nécessaire, l'auteur soutient qu'il faudrait aussi utiliser des procédés plus efficaces pour la planification, la réalisation et l'assimilation de projets de réforme. S'il est souhaitable que les procédés de surveillance et d'évaluation contribuent effectivement à l'optimisation de résultats immédiats et des effets à long terme de tout projet ou programme de réforme, une approche compréhensive, orientée vers le développement et intégrée dans le processus même s'impose. Un tel procédé devrait tenir compte aussi bien des facteurs relevant d'autres domaines que des variables à l'intérieur du domaine de l'éducation qui influent sur les processus et résultats d'une réforme. Le schéma EIPOL présenté combine cinq dimensions essentielles d'un système étendu d'évaluation avec les différentes étapes d'un cycle de projet et offre un cadre de base pour élaborer et adopter un système plus fonctionnel d'évaluation de réformes.

Zusammenfassung – In herkömmlichen Planungs- und Durchführungssystemen für Bildungsreformen steht Evaluation gewöhnlich am Ende des Prozesses. Obgleich solche summarische Evaluation notwendig ist, plädiert der Verfasser für Evaluationsverfahren, die eine wirksamere Planung, Durchführung und Assimilation von Reformprojekten ermöglichen. Wenn Überwachungs- und Evaluationsverfahren einen wesentlichen Beitrag zur Optimisierung der unmittelbaren Resultate und Langzeit-Wirkungen eines Reformprojekts oder -programms leisten sollen, muß ein in den Reformprozeß eingebauter, umfassender und entwicklungsorientierter Ansatz benutzt werden. Ein solches Evaluationsverfahren sollte sowohl außerhalb des Bildungsbereichs liegende Faktoren als auch bildungsinterne Variablen berücksichtigen, wel-

International Review of Education – Internationale Zeitschrift für Erziehungswissenschaft – Revue Internationale de Pédagogie XXVI (1980), 475-482. All rights reserved. Copyright © 1980 by Unesco Institute for Education, Hamburg and Martinus Nijhoff Publishers bv, The Hague.

che die Prozesse und Ergebnisse einer Bildungsreform beeinflussen. Ein EIPOL Raster, in dem fünf wesentliche Dimensionen eines breit angelegten Evaluationssystems mit den verschiedenen Stufen eines Projekts verbunden sind, bietet einen operationellen Grundrahmen für die Gestaltung und Einführung eines funktionelleren Reform-Evaluationssystems.

Conventional Evaluation of Educational Reforms – A Threat

In recent years, most developing countries have undertaken significant programmes and projects of renovating and reforming their educational systems. Primary education is undergoing rapid expansion, and additional educational facilities are being created on a massive scale in many countries in order to meet social demand for education. Secondary education is being drastically overhauled by introducing diversification of courses into technical, commercial, agricultural, home economics, and academic streams. Corresponding reforms are being introduced in teacher education, curriculum and examination systems. Alternative delivery systems like distance learning and correspondence-cum-residential courses at the school as well as university stages, are being attempted in an increasing number of countries. Non-formal patterns of learning arrangements as part of lifelong education are also being evolved for improving the production skills of farmers and factory workers or for eliminating illiteracy and general backwardness among the population of urban slums and rural areas. These reforms also include basic changes in educational policy, administration, educational structure, and related fields. Similar changes, though of a somewhat different character and order, are discernible in developed countries as well.

These reforms are usually mounted by undertaking planned programmes and time-bound projects in order to attain certain goals and targets. Most of these developments consume large amounts of money, time and other resources and hence their effectiveness and impact are matters of concern for policy-makers, planners and administrators. Consequently, what happens in a typical situation is that at the end of the period of implementation of a particular project, some sort of an evaluation is attempted. Typically, some high-level official or a public body makes the proposal to carry out an evaluation before undertaking new projects. Such an evaluation exercise, coming only at the end of the implementation stage, frequently proves to be a threat to the workers directly involved in a particular programme or project of educational change, and often generates a negative attitude among all concerned. This is because such an end-of-project evaluation does not help in improving the outcomes of that project and usually results in a fault-finding type of post-mortem exercise. The reform evaluation that is carried out at the end of a project or programme does not help in identifying the barriers and

bottlenecks that have arisen from time to time during the implementation period, in time to apply corrective measures and improve results. The end-of-project type of evaluation therefore implies a rather restrictive use of evaluation and does not become an effective tool for educational development and change at all stages of reform management.

Rejection of a Restrictive Model

The model that is implicit in the kind of typical and commonly practised evaluation procedure mentioned above is cyclic and sequential in nature as shown in Figure 1.

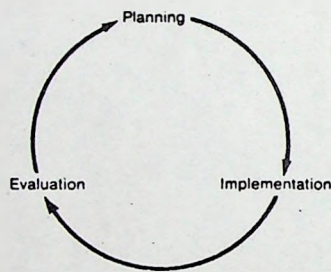


Figure 1. *Cyclic and Sequential Relationship between Planning, Implementation and Evaluation (Restrictive Model)*

The assumptions underlying such a relationship between planning, implementation and evaluation are worth noticing. Evaluation occurs at the end of the cycle, and hence it has only a measurement value for a particular project and not an improvement value for that project to a satisfactory degree. In this type of sequence, evaluation is not a built-in process in the entire project or programme and does not become an effective tool for improvement, change and optimisation of outcomes throughout the procedures of planning and implementation. It is therefore necessary to reject this model on account of its restrictive relationship between planning, implementation and evaluation, and work out a new relationship which should permit the use of evaluation procedures in a more effective and dynamic manner at all stages of a project cycle. Certainly, a summative and comprehensive type of evaluation, as is implicit in the restrictive model, is necessary at the end of any project, but this is not enough. What is needed is to develop a built-in system of evaluation in order to use evaluation more fully, more positively, and at all steps in the process of educational change and development.

A Built-in, Comprehensive and Development-oriented System of Evaluation

For a development-oriented evaluation, it is necessary to build the mechanisms of appraisal, feedback, diagnosis and remedial action at all stages of a project or programme including pre-planning, planning, implementation and assimilation stages. In this model, evaluation does not figure at the end of a project but it cuts across all phases of a project cycle as shown in a rather simplified manner in Figure 2.

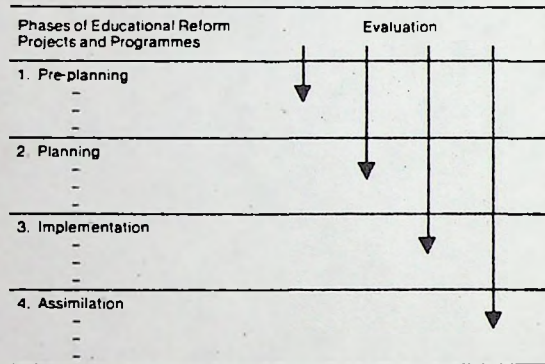


Figure 2. Relationship between Planning, Implementation and Evaluation in a Development-oriented Approach to Evaluation

It may be noted that apart from the pre-planning, planning and implementation phases which are often taken into account in developing a project or programme of educational change, a stage of assimilation has been added. This stage is considered important because any new development in education requires more time to settle down and mature, after the implementation phase is formally completed, on account of the nature of the process of education, amongst other factors. The change has to be assimilated with other aspects of the education system and, in some cases, with the larger social system. Furthermore, an educational programme is often expected to produce long-range effects besides immediate outcomes. Sometimes the programme degenerates and fades away for want of certain precautionary and reinforcement measures. It therefore follows that evaluation should be extended to the assimilation stage in order to find out what happens to a project or programme after a formal implementation stage is over and whether it has produced or failed to produce any long-range impact.

When evaluation of reform projects and development programmes is

viewed as a built-in system, its scope is not confined to the evaluation of immediate outcomes only as often happens when evaluation is introduced at the end of the implementation stage. In fact, the scope of a built-in system of project evaluation should be extended far enough to allow for a number of extra-educational factors besides intra-educational variables that are connected with larger and long-range effects as well as with immediate outcomes. It thus includes the evaluation of learners, institutions, curricula, etc. but goes much beyond them.

For any educational programme to become reasonably effective, it is necessary to make a careful appraisal of the environmental setting in which it is to function. Appraisal of an environmental setting should first of all take into account social, economic, political and other such factors that often condition and control educational outcomes. The success of an educational programme or project often lies to a considerable extent in the extra-educational domain. Moreover, the situation obtaining within the existing educational system itself is equally important for success and hence a good diagnostic appraisal of the educational domain in its historical and current perspectives becomes necessary. Based on these analyses, an assessment of learning needs and priorities as well as of resources and potentialities should be carried out at the pre-planning and other stages of a project, according to the levels at which a project is expected to operate at different times. Similarly, a systematic evaluation of material and non-material inputs as well as of administrative and pedagogical processes should be carried out by a progress-monitoring system, formative evaluation and other devices. Thus, a built-in evaluation system should have wider scope than just outcome evaluation and should include the following dimensions:

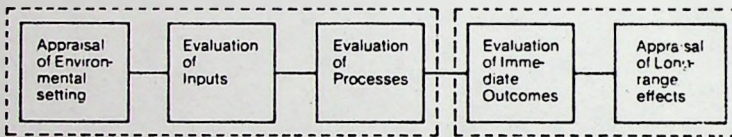
1. Appraisal of the Environmental Setting
 - 1.1 Diagnostic analysis of the historical and current situation in the socio-economic, political, educational and other domains
 - 1.2 Assessment of needs and priorities
 - 1.3 Appraisal of resources and potentialities
2. Evaluation of Inputs
 - 2.1 Material inputs
 - 2.2 Non-material inputs
3. Evaluation of Processes
 - 3.1 Management processes
 - 3.2 Pedagogical processes
4. Evaluation of Immediate Outcomes
 - 4.1 Intermediary outcomes
 - 4.2 Learning outcomes

5. Appraisal of Long-range Effects

5.1 On the educational domain

5.2 On socio-economic and other domains of development.

These five major dimensions of evaluation (abbreviated as EIPOL model) are inter-related in many ways. Figure 3 indicates their inter-relationships. Some of these relationships are self-evident or can be supported by available evaluation studies, whereas others need further exploration. The relationship can also be seen in different and alternative ways depending on a particular focus within these stages. For example, the immediate as well as long-range effects can be coupled together with a view to evaluating the role of environmental, input and process variables as follows:



An EIPOL Grid

These five dimensions or phases of evaluation are also inter-related in multiple ways with the four major phases of a project cycle, namely,

1. Pre-planning
2. Planning
3. Implementation
4. Assimilation.

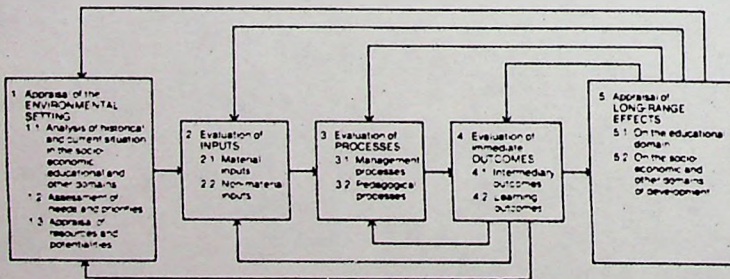


Figure 3. Stages of Built-in Evaluation in Educational Projects and Programmes: EIPOL Model.

In order to develop an instrument for planning a built-in system of evaluation of reform projects and programmes in education, these two sets of phases can be combined into an *EIPOL grid* as shown in Figure 4. Such a grid is useful for various purposes. Different cells in the grid indicate different elements of evaluation. For example, assessment of needs, priorities and resources belong to cell 1.1. Evaluation of learners' achievement belongs to cell 2.4 at the formative stage of curriculum development and to cell 3.4 at the periodical testing and summative stages. Likewise, cell 4.5 points to the need for tracer studies and other 'follow-up' evaluations in order to obtain some insight into the larger and long-range impact of a particular project after the formal implementation stage is over. Cells 3.2, 3.3 and 3.4 are of particular significance for progress monitoring of reform implementation. Thus, the grid can be used as a kind of mental tool for planning a comprehensive and built-in evaluation system, and for other similar purposes.

It may be added that such an evaluation system can be applied in a harmonious manner at different levels of project activity, such as local, district, provincial (regional) and national, depending on the nature and structure of the project. The stages of pre-planning, planning, implementation and assimilation will also assume different meanings and functions at different levels at a particular point in time. For instance, a programme that is at the implementation stage in the context of national level action, may, in effect, initiate a planning stage in a given school before the local level implementation stage is commenced. A built-in evaluation system should take this factor into account. In fact, a third axis representing the *level* can be added to the two-dimensional grid presented in Figure 4, but for simplicity's sake it has been omitted.

Stages of built-in evaluation system Stages of a Project Cycle	1 Appraisal of ENVIRONMENTAL SETTING	2 Evaluation of INPUTS	3 Evaluation of PROCESSES	4 Evaluation of IMMEDIATE OUTCOMES	5 Evaluation of LONG-RANGE EFFECTS
1 Pre-planning	1.1	1.2	1.3	1.4	1.5
2 Planning	2.1	2.2	2.3	2.4	2.5
3 Implementation	3.1	3.2	3.3	3.4	3.5
4 Assimilation	4.1	4.2	4.3	4.4	4.5

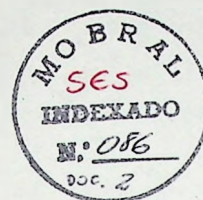
Figure 4. *EIPOL Grid for planning a Built-in System of Evaluation at Different Stages of a Project Cycle.*

When planning a built-in evaluation system, it is necessary to pay adequate attention to several operational aspects so as to ensure the feasibility and viability of the plan. If the built-in evaluation strategy is to be made success-

- ful, particularly in developing countries, it should not become either too costly or too time-consuming. Its role as only one among several important instruments to be used for bringing about educational change and development should be clearly visualized in a proper perspective. Similarly, according to specific evaluation needs and other related factors, a proper balance should be maintained between quick and qualitative reviews on the one hand, and more quantitative and sophisticated procedures on the other, when selecting different techniques and tools of evaluation for different stages and dimensions of evaluation. There are several other important aspects to be kept in view such as the increased participation, where feasible, of the personnel concerned, in conducting the evaluation, the maximum possible use of evaluation results for feedback and change, and so forth. Thus, planners, administrators and evaluators have to take both technical and practical aspects into account in planning and practising a more comprehensive and improvement-oriented evaluation system for optimising the efficiency and effectiveness of educational activities.

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 2



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOMES AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

Reform Evaluation
by
E Ayotunde Yoloje

Reform evaluation

When we speak of the evaluation of educational reforms, we are in the realm of what is often referred to as 'programme evaluation'. This has been defined as 'a type of applied research in which program process and outcome characteristics are related explicitly to a set of values, such as program goals, objectives and costs' (National Institute of Mental Health, 1976). Conducting programme evaluation is a technical job in many respects. This article does not however intend to present a technical treatise on how to carry out programme evaluation. The focus is essentially on drawing attention to essential issues in programme evaluation. Hopefully the contents of this article will be helpful to the technical expert as well as to the administrator in considering matters of evaluation in educational reform.

E. Ayotunde Yoloye (Nigeria). Professor at the Institute of Education, University of Ibadan; Director of the International Centre for Educational Evaluation at the University of Ibadan; Chairman of the African Curriculum Organization; served in 1980/81 as a consultant to the International Institute for Educational Planning (Unesco) in the training programme. Author of several books and papers on evaluation.

Definitions

For the purpose of this article the term 'educational reform' has a specific meaning, namely 'those policies and programmes that aim at a major and nation-wide change and new development in one or more aspects of an educational system'. The emphases in this conception of educational reform are (a) its pervasiveness and (b) its direction towards change in the status quo. Clearly, in implementing such educational reform, one must be concerned with a somewhat complex network of administrative and communication infrastructure, from a central co-ordinating body to the grass roots.

The term 'evaluation' has a variety of definitions. In ordinary usage it means the assigning of some value to an entity in relation to certain criteria. Defined this way the focus of evaluation would appear to be the passing of judgements. However most people who are concerned with educational reform are concerned not merely with being able to pass judgement on how much reform has or has not taken place. They are concerned with seeing that reform does take place. Therefore there is need for a more guidance-oriented definition of evaluation. For this purpose the definition by Alkin (1970) seems most appropriate. He says, 'Evaluation is the process of ascertaining the decisions to be made, selecting related information, and

collecting and analysing information in order to report summary data useful to decision makers in selecting among alternatives.'

It is of course true that after a period of time during which a reform programme has been in operation, one may wish to pass some judgement on the degree of success that has been achieved in its implementation. We may therefore further distinguish between two kinds of evaluation. First, there is the type of evaluation which guides and aids development. Secondly, there is the kind of evaluation which gives judgement as to the value or worth of the resulting programme and especially the outcomes of the programme. Scriven (1967) applied the terms 'formative evaluation' and 'summative evaluation' respectively to the two kinds.

Whose idea to reform?

Implicit in the term reform is the assumption that the status quo is unsatisfactory. The big question is: 'Who feels aggrieved by the status quo?' Ideally reform should be stimulated by a pervasive feeling of dissatisfaction from the grass roots to the top. Experience shows however that this is not always the case. Sometimes policy-makers at the top apparently perceive certain needs in society that may not be perceived by the masses. When this happens the policy-makers have a selling job to do in addition to implementing a reform. Examples of this situation may be found in the series of campaigns mounted in the United Republic of Tanzania in the last decade to sensitize the populace for large-scale social reforms. At the other extreme reforms sometimes find their origins at the grass roots and the masses perceive certain needs which the policy-makers at the top have not perceived. It is then the turn of the masses to sell their ideas to the policy-makers. Often this is done through petitions, agitations, demonstrations and sometimes violence. As a rule policy-makers and adminis-

trators would prefer a situation where there is communication of feeling and needs between all strata of society so that reform policies are responsive to the needs of the populace who will be directly affected by such reforms.

What should be evaluated?

The tasks of evaluation would vary with the answer to the question, 'Whose idea to reform?' Let us for the moment consider the happy mean when there is adequate intercommunication between the populace and the policy-makers. What events, situations, etc., should be evaluated? Stake (1967) identifies three categories of events: (a) antecedents, i.e. relevant conditions prior to the introduction of a reform programme; (b) transactions, i.e. the various kinds of interactions and activities that take place during the development and implementation of the programme; and (c) outcomes, i.e. the effects of introducing the programme on various categories of people, things, structures, processes, etc.

Evaluation of antecedents would in this situation precede the introduction of a reform. It would deal with the important task of needs assessment at various levels; it would assess resources as well as other aspects of the feasibility of introducing a particular reform.

Evaluation of transactions would be a combination of simple monitoring as well as specific studies aimed at providing decision-makers with rational bases of choosing between alternatives in the implementation process. This is the core of formative evaluation.

Evaluation of outcomes is a natural outgrowth of the evaluation of transactions and is the central ingredient in summative evaluation.

In many developing countries, this happy medium is hardly ever struck, for a number of reasons. For one thing the data base from which a substantial part of the evaluation of antecedents could be done is absent. A number of

countries do not even have valid or up-to-date census figures, not to mention other important demographic data. For this reason many developing countries have adopted one or the other of the two extreme approaches to reform described above. Some writers have even referred to the approaches to reform in many developing countries as 'management by crisis'. In other words, policy-makers in these countries often embark on extensive social reform without adequate evaluation of antecedents. Consequently, questions of feasibility and acceptability are not answered. Rather, crisis situations arise in the process of implementation, and frantic efforts are then made to resolve them. Somehow many developing countries have managed to move forward in their development by this 'crisis' approach, and learn more or less in trial-and-error fashion. The 'crisis' model of management has far-reaching implications for evaluation, for the nature of evaluation questions has to change radically. The evaluators no longer pose the question, 'Is this reform desired or feasible?'; they have to answer the question 'Given this reform policy, how best can it be implemented?'

Management of reform

Given our definition of educational reform, it is obvious that implementation of educational reform requires a rather complex management network. The exact nature of the network will vary with the nature of the reform and the administrative structure of the country concerned. For illustration however, let us take a look at the structure for the Nigerian Universal Primary Education (UPE) programme, which was launched in 1976. Administratively, Nigeria is a federation of nineteen states. Each state is divided into several local-government areas. The UPE policy is a federal-government policy, but the financing and administration is a tripartite responsibility involving the federal govern-

ment, the state governments and the local governments. Already even parents are assuming considerable financial and material responsibility for the programme. Thus there is a management network linking the federal Ministry of Education with nineteen state ministries and their state school boards, over 300 local governments and several thousand schools and teacher-training colleges.

For effective operation of the UPE programme, it is clear that varying levels of decision-making must operate, and an effective intercommunication system between the various levels must be set in motion.

Management of evaluation

Since the management structure for the implementation of a reform is so complex, it is reasonable to expect that the processes of collecting and using evaluation data on the programme would also be complex. In other words, the evaluation programme needs a management structure of its own. The question that arises is, 'Should this be a separate structure from that of the implementation structure?' The answer to that question, in part, depends on whether the programme evaluation is intended to be an internal self-evaluation programme or is contracted out to a specialized evaluation body not necessarily part of the governmental management structure. In the first alternative, the evaluators in effect are part of the implementation management. In the second alternative the evaluators are not part of the implementation management, but in carrying out their duties they must interact intimately with the former. In both cases it is clear that evaluation requires certain specialized skills and must be carried out by people possessing such skills. It is of course quite possible that skills in implementation management and evaluation may be possessed by the same person. Indeed it is desirable in many situations that whether a

person is primarily designated a manager or an evaluator he should possess skills in both areas of functioning. However the successful operation of an evaluation programme for any particular reform depends on several other factors besides the mere possession of the technical skills of evaluation. We now turn our attention to a consideration of these factors.

Evaluation as a threat

Broadly speaking, programme evaluation should have as its major goal the provision of essential information for use in decision-making to shape the services offered and to make the services more responsive to the people who are in effect the clients. However besides improvement of services as a goal of evaluation there is often a second powerful goal: that of *accountability*. The two goals may evoke vastly different attitudes in those connected with the reform. At this stage, it is pertinent to examine a connotative meaning of the term 'evaluation' of which perhaps many people are not consciously aware. I have written elsewhere:

Easily the most common connotative meaning of evaluation is 'threat'. The child, the teacher, and the curriculum innovator feel threatened by the word 'evaluation' . . . when a human being is the object of evaluation the very process of measuring any aspect of his make-up, especially one of a psychological nature, causes that aspect to alter its value. Thus it becomes difficult to describe values obtained as due to the treatment whose effect is being measured or to the process of measurement itself.

If the purpose of evaluation is 'accountability' the 'threat' factor comes very strongly into play and the evaluators must have certain characteristics if they hope to carry out a successful programme of evaluation.

For effectiveness within a reform-management system, the evaluator should: (a) have competence in the skills of evaluation; (b) become a responsive and interactive member of the

management team; (c) support active rather than passive management processes; (d) function as an intermediary and facilitator; and (e) have administrative competence. The point being emphasized in this list of characteristics is that the evaluator needs more than technical skills. He needs, perhaps more importantly, skills in social interaction and administration. As the NIMH *Manual* states (National Institute of Mental Health, 1976, p. 3):

Charters of official status, guaranteed access to management, specific grants of authority, and formal relationships spelled out between evaluator and [project] leadership are insufficient substitutes for real membership in the management team, and are not always productive for attaining real membership. . . . Many of the conditions essential to useful program evaluation. . . . cannot be legislated.

Attitudes to evaluation

Even when a smooth relationship exists between evaluator, management and client, there can be multiple perspectives on what represents appropriate impact of a reform programme. Sometimes, evaluation criteria with respect to different audiences are incompatible. For example in evaluating the University of Ife Six-year Primary Programme—an experiment in teaching through the medium of the mother tongue throughout the primary schools—the evaluators used criteria of cognitive and affective gains within the school curriculum, but the ministry and the public at large used criteria of improved performance in the common entrance-selection examination for admission into secondary schools (Yoloye, 1976).

Of course the perspective of the evaluator also affects very strongly what aspect of evaluation is emphasized. Again, quoting from the NIMH *Manual*.

To some evaluators, what they do is research, discovering or verifying program-specific knowledge, and reducing uncertainty about what is the truth. To others, evaluation is justification, occasionally even

defense or a put-off—something with which to fend off unfriendly outsiders. By those under evaluation, the process may be perceived as punishment, and evaluators as hatchetmen. More positively, evaluators are sometimes seen as supportive management specialists providing useful feedback to administrators and practitioners and useful verifications to those who need them. Trouble ensues when the evaluator, the user of the evaluation, and those whose work is being evaluated have different views on the evaluative function being performed. . . . Trouble is worst when each party to the evaluation imputes discreditable and possibly hidden motives to the others (such as empire-building, self-protection, or excessive professional competitiveness). People tend to avoid bringing such subjects to the surface and do not know how to test for their absence or presence skillfully and without taking foolish risks.

Factors in drawing up an evaluation plan

With all the preceding factors in mind, one may draw up a more comprehensive list of factors that must be taken into consideration in drawing up an evaluation plan. Such factors include: (a) willingness of those concerned to have their activities evaluated; (b) the needs or demands for evaluation, e.g. 'improvement of services' or 'accountability'; (c) the probable 'success' of the evaluation, i.e. feasibility in terms of time, resources, design and data source, and its likely impact and acceptability; (d) the cost of the evaluation.

At a more detailed level, one must consider: (a) priorities in management's needs for answers; (b) the potential value of evaluation to those whose work is being evaluated; (c) the kinds of evaluation questions that need to be answered; (d) whether key issues are resolvable by answering these questions; (e) whether courses for further action after resolution of issues are clear.

Management information systems

A vital factor in the evaluation of any programme is the management information system (MIS).

A management information system is any system of people, equipment, documents, procedures and communications which gathers, stores, processes and presents information for the purpose of assisting managers with a variety of managerial functions.

When evaluation is integrated into the management structure, programme evaluation information becomes part of management information. If evaluation is external to management, however, the management information system becomes a source of data for the evaluation programme. The role which the MIS then plays is to collect, process and present much of the information and data which is used in performing the evaluation.

Information requirements

Three main categories of information are required for programme evaluation.

1. *Internal*. This is the kind of information generated entirely within the system and it covers things like (a) programme objectives, (b) statistical information, e.g. number of schools, pupils, teachers, classrooms, equipment, etc., and (c) accounting information dealing with costs, etc.
2. *External*. This is the kind of information generated from outside the system, e.g. comparative data from other systems, market variables, data from existing research, etc.
3. *Transactional*. This is the kind of data arising from interaction of the system with other systems, e.g. between Ministry of Education and Ministry of Information or Works, etc.

Although many management systems are not consciously aware of it, each management system has some form of management information system. These can range from simple

manual systems to highly automated and computer-based systems. It is vital that the evaluators should be aware of the kind of management information system that exists in the reform they seek to evaluate; for the kind of system has implications for cost, time allocation, techniques and efficiency.

One may identify for purposes of convenience four kinds of MIS.

1. *Manual systems.* These are based mainly on filing and indexing and therefore require skills in office management and clerical work.
2. *Machine-assisted manual systems.* Filing and indexing are supplemented by machines like key-punch equipment, card sorters, desk calculators, etc. These machines greatly ease the steps that make data useful for evaluation although they are not necessarily more efficient for management.
3. *Systems using simple computers.* Many desk-top computers can now be used for much data storage, processing and retrieval. Their use requires a higher level of technical expertise than for the simple machines listed in (2) above.
4. *Systems using computer service bureaux.* Many systems have computer service bureaux as part of the system; for example the West African Examinations Council has its own computing centre. In more developed countries, computer service bureaux as part of the establishment are almost commonplace. Of course the bureau may also exist outside the system. The computers may of course be bought or leased, with varying cost and utility implications. Most systems prefer to lease, as they can then obtain the latest, most up-to-date computers in the market.

With the computer bureau, data may be handled by batch processing, in which all required data must be available for one-time analysis or online processing in which processing can be done in stages usually from the systems' own computer terminal. Online

processing can be more expensive but it is also more flexible.

The important consideration here is that evaluation designs and administration must be compatible with the available management information system. Thus the evaluation designs in many developing countries must be somewhat different from that in developed countries. A number of cost implications and time implications will of course result from these considerations.

It may be also pertinent that the functioning of the more sophisticated MIS depends very often on the functioning of other systems in the society in ways that the manual systems do not. For example the efficiency of electrical power supply, water supply and telecommunications drastically influences the performance of the computer bureaux as well as even the desk-top computers. There is also the important factor of whether repair or servicing facilities are available or not.

Evaluation techniques

Programme-evaluation methodology is not one single thing; it is not a unidimensional entity. It is perhaps better to regard it as a spectrum of activities ranging from rather informal, intuitive administrative judgements to sophisticated and rigorous research studies. For convenience one may classify evaluation activities into six categories ranging from the intuitive to the rigorous.

1. *Administrative judgement.* This is usually based on the administrator's immersion in the living data of the programme being evaluated, and is reinforced by documented answers to specific questions.
2. *Routine monitoring.* Conclusions can often be drawn from routine monitoring of explicit programme data arranged in preselected categories, processed, analysed and displayed to enable implicit or explicit com-

parison with norms or plans of operations or with other experience. One evaluation technique that is particularly relevant in this category is programme evaluation and review technique (PERT) which identifies and sequences appropriately all the essential events, processes or activities involved in a programme. Excellent treatments of PERT exist in various documents, e. g. Gurugé (1975).

3. *Recurring studies* of routinely and periodically reported data. These often lead after analysis to conclusions and recommendations for management.
4. *One-time studies* requiring special care in data gathering procedures. The methods used and questions asked are usually related to managerial utility of results. An example would be cost-effectiveness studies, or surveys.
5. *Quasi-experiments* with clear and specific management focus.
6. *Research studies* involving rigorous designs and careful controls addressed to broad management problems and often testing specific hunches or hypotheses.

Programme evaluation may involve some or all of these techniques, depending on the expected comprehensiveness of the evaluation. Formative evaluation will have greatest reliance on techniques (1) to (4) while summative evaluation will put greater emphasis on (5) and (6) but it should be noted that this is a case of emphasis and not exclusiveness. Specially commissioned evaluation studies often fall into the category of summative evaluation but often rely heavily on techniques (2) and (3).

Financing evaluation

Most educationists would advocate that formative evaluation should be an integral part of every reform programme. As I wrote earlier (Yoloye, 1978):

The primary purpose of formative evaluation is to help as much as possible to ensure that summative evaluation comes out positive. If he wants to do formative evaluation really well, the evaluator cannot be an objective outsider.

The formative evaluator has to become deeply involved in the reform process itself.

In practice however there is often a marked resistance to building more than the routine monitoring and intuitive levels of evaluation into reform programmes. One of the reasons for this reluctance is the cost, especially where resources are scarce. For if one were to cost separately what would be involved in comprehensive formative evaluation, it would come to a substantial amount. Most policy-makers tend to see this as a diversion of scarce resources from the main reform programme to something which is usually regarded as subsidiary. Yet one feels almost intuitively that the cost-effectiveness ratio of a reform programme would be highly improved if that extra cost for formative evaluation were built into the reform-programme implementation.

It is only fair however to point out that resistance to financing formative evaluation has other roots besides the cost implications. Some of the other factors are: (a) reluctance by implementation staff to undertake the additional burden of data collection; (b) desire to minimize disruption of implementation processes by evaluation processes; (c) the fear that evaluation studies may lead to the distortion of programme focus (it is well known in the educational field that the examination often determines the curriculum rather than the other way round; some sceptics refer to this phenomenon as 'the tail wagging the dog'); (d) concern that evaluation results may be used destructively; (e) recognition that programme evaluation results are often not used at all and thus expenditure on it constitutes wastage of scarce resources; (f) scepticism about the validity of evaluation studies anyway.

On the whole, specially commissioned

one-time evaluation studies tend to be less expensive and less disruptive of the reform process, and therefore tend to be generally favoured by financing bodies and implementation staff.

So, in practice, mostly lip service has been paid to formative evaluation, and I believe that the faith of the specialist in formative evaluation has hardly been effectively tested or validated in practice. One knows that in the field of student achievement, extensive research has gone into the validity of formative evaluation, especially as it relates to mastery learning (Bloom, 1976). Results from such studies have been positive. This gives a hope that in programme evaluation formative evaluation would yield similar dividends. One realizes of course that dealing with the individual child is vastly different from dealing with a management team, but it would help if funding bodies would take greater chances by financing formative evaluation so that cost-effectiveness studies in the area may be the validity of the educationists' hopes truly tested.

Indicators of success

One of the most difficult aspects of evaluation of educational reform is the choice of the reform's indicators of success. Education is a social service, and so, besides purely management criteria which may be used in connection with PERT charts, indicators of the reform's impact on the society have to be looked for in the society.

The term 'social indicators' came into vogue in the 1960s to describe various social conditions and trends. It has become the focus of attention in trying to identify the impact of any social reform on the society.

Social indicators have been demanded for various purposes by various groups. Among the purposes are: (a) the establishment of goals and priorities; (b) the evaluation of social programmes; and (c) the development of a system

of social account that could provide guidance among many alternative interventions. Sheldon (1975) gives a concise accounts of the social-indicators movement.

With specific reference to educational reforms, the term 'educational indicators' has come into vogue. Educational indicators have been defined (Gooler, 1975) as 'statistics that enable interested publics to know the status of education at a particular moment in time with respect to some selected variables, to make comparisons in that status over time, and to project future status'.

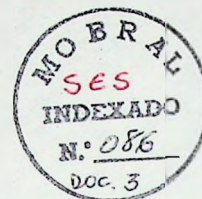
Thus educational indicators are seen as time-series statistics that permit a study of trends and changes in education. Gooler (1975) also provides a neat summary of categories of educational indicators, of which he identifies five as follows:

1. *Access*
How many and what kinds of people participate in educational activities
Retention rates in educational activities
Catalog of existing/available educational activities or services.
2. *Aspirations*
Description of needs and desires of various kinds of people
Individual self-assessments of personal capabilities
Description of institutional goals
3. *Achievement*
What people, know, do, and feel
What people have earned (degrees, diplomas, certificates)
What is taught
4. *Impact*
Consequences of having schooling
Impact of education on social/economic/cultural systems
Consequences of not having schooling
5. *Resources*
Capital, personnel, and materials expenditures
Quality of human resources
Cost-benefit/effectiveness ratios
Quality of educational climate
Time

These various indicators are measurable and a variety of techniques have been developed to measure them. The particular evaluation

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 3



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOMES AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

A Self Evaluation of the National
Literacy Programme
in Botswana
Second Outline

by

Ulla Kann

Botswana : Workshop document

HOW CAN WE SUCCEED?

A Self-Evaluation of the National Literacy Programme
in Botswana.

Second outline.

Ulla Kann
Planning Unit
Ministry of Education
April 1982

Paper developed during Project Development workshop, Moshi
Tanzania 29th March to 6th April 1982. IIEP Research and
Training Project on Evaluation and Monitoring of Educational
Reform Programs in Africa.

1. Background and Description of the National Literacy Programme

1.1 Historical Account.

The National Commission on Education (Republic of Botswana, 1977) recommended a major change in the educational policy, namely, a shift from the previous emphasis on secondary and higher education to basic education for everyone through primary education for school age children and non-formal education for young people and adults who had not received primary education.

been/ Though there had previously/several literacy programmes started by voluntary organisations there had not been any national initiative to eradicate illiteracy. Following two pilot projects carried out by the then Botswana Extension College (BEC, 1978) the Department of Non-Formal Education within the Ministry of Education in the beginning of 1979 initiated the present National Literacy Programme. The programme is planned to last for 6 years including the experimental year 1980. The national programme, covering all nine districts of the country, was launched in June, 1981.

1.2 Objectives

The National Development Plan 1979 - 1985 states that the aim of the programme is to eradicate illiteracy by 1985 and provide basic education to those lacking formal education. This aim should be seen in relation to the overall development planning objectives i.e. rapid economic growth, social justice, economic independence and sustained development. As a more precise definition of the above objectives the present National Development Plan contains two main themes i.e. employment creation and rural development. The National Literacy Programme should be seen as one of the means towards achieving these objectives.

The consultation document that prepared the ground for the programme specified the objectives by defining 'literate' as implying that a person can comprehend those written communications and simple computations which are part of daily life.

1.3 Major Programme Activities

The Department of Non-Formal Education is the main executing agency for the Programme. When preparing for the implementation of the programme various steps were taken. Literacy Assistants (with three years secondary education, Junior Certificate) were employed and trained for their tasks as supervisors and trainers of Literacy Group Leaders. The Literacy Assistant (LA's) together with District Adult Education Officers (DAEO's) also did the recruitment of the Literacy Group Leaders (LGL's). Other countries have used well defined groups as implementors of their literacy programmes e.g. school teachers, secondary or university students etc.

This has not been the case in Botswana. The Literacy Group Leaders are persons who are interested in teaching others to read, write and count and who have recruited a sufficient number of people to form a group. They often do not have more than primary education. The teaching method has been designed to suit this group. In addition to recruitment and training of personnel a lot of effort has been put into the preparation of material. A total of five primers have been developed. In addition to primers, handbooks for LGL's, flashcards and literacyboxes (consisting of a box for the LGL's material and at the same time being both a flannellograph and a blackboard) have been developed and produced in large quantities. A monthly broadsheet is also issued to the participants. All material is supplied free of charge to the learners.

The programme has been planned to have an annual cycle adjusted to the annual migration of the participants. Thus in April Literacy Assistants are appointed and trained. During May-June LA's and District Adult Education Officers recruit LGL's and prepare for their training. In July the LGL's are trained so that teaching can start in August when participants have returned from the lands.

The groups continue until people disperse to the lands for ploughing, when rains have started. The time of learning thus can vary from year to year. When possible the learners are regrouped at the lands in order to continue to work on their own.

In addition to the programme administered by the Department of Non-Formal Education other government as well as non-government agencies use the National Literacy Programme material for literacy teaching. The most important of these is the Tlrelo Setshaba with its increasing number of participants. Other groups are to be found within the Prison Department, brigades, Women's Organisations etc. Special "One-to-one" packages have also been prepared and distributed. These packages are intended for people who for some reason cannot attend a literacy group, but who has somebody willing to help in the teaching of literacy and numeracy.

1.4 Size and Scope of Programme

The National Development Plan and the first project documents state the planned scope of the project in the following way:

"The programme will commence in 1980 as an experimental year with an intake of 15,000. Thereafter it is anticipated that 50,000 participants will be enrolled annually, each cohort receiving 12 months of supervised instruction followed by the supply of self-tuition materials for the ensuing twelve months".

Experiences so far have indicated that it is not possible to achieve these targets. Appendix A contains the Composite Statistics for the programme in December, 1981. The numerical targets have been adjusted towards a total of 30,000 a year including various voluntary organisations.

Several donors have contributed generously towards the programme. A revised budget is included in Appendix B.

1.5 The Need and Purposes of evaluation.

The proposed evaluation is made for three different reasons. The documents on which the programme is based stated that there should be a mid-term evaluation in 1982/83. As mentioned above several donor agencies have contributed to the programme. It is natural that they want information on the development and success of the programme. They will most likely base their future support partly on the outcome of the evaluation.

The Government of Botswana and more specifically the Ministry of Education has undertaken to carry out this programme in order to eradicate illiteracy by 1985. In order to succeed in such an undertaking it is important to take stock so that necessary modifications and improvements can be done. In effect this ought to be done with periodic intervals and it is intended that in addition to give information on how the programme has proceeded the evaluation should also lead to a more efficient monitoring system.

The proposed evaluation has a third purpose somewhat different from normal evaluations. It has a training component. There are few people in Botswana who are trained to carry out an evaluation. The proposed evaluation will provide on-the-job training in evaluation for District Adult Education Officers, other Department of Non-Formal Education staff and Literacy Assistants.

2.1 Technical Elements of the Evaluation Design

Appendix C gives a detailed account of programme objectives and activities, assumed users of evaluation information, evaluation questions or aspects to be evaluated, indicators to be used, data collection instruments sources of data and persons to be responsible for the collection of data. An overview of the Appendix is given below.

The design follows the EIPOL model (Dave', 1980). Thus the following aspects of the programme are dealt with:

- (E) Environmental setting (historical background of the programme, relationship to other development efforts, national support etc.)

- (I) Inputs These can be human (staff, participants), material (primers, vehicles, venue) or financial.
- (P) Process This section deals with educational processes (teaching method, training of staff, cultural aspects of the programme, motivation of learners) and with administrative processes (supervision/monitoring, link to other education activities, distribution, communication, financial administration and control).
- (O) Outcomes Relates the immediate outcomes of the programme with the objectives of enrolling 30,000 persons per year in literacy work and to make them literate and numerate.
- (L) Long range impact deals with the impact of the programme on a national basis in relation to national and district development goals.

2.2 Identification of data sources and data - gathering tools Some sources of data do already exist and it is more a question of compilation and analysis. These are, for example:

The monthly attendance register (LI)
The Group Report Form (LG2)
The LA's monthly visit report form (LG3)

The LA's weekly report form (LG4)
The Location of Literacy Groups form (LG6)
The Register of Literacy Participants (LG 9)

As indicated in Section E of Appendix C these forms will be used in the evaluation. However, some new instruments are needed, namely:

- Some form of test or self assessment instrument to make it possible to decide whether the objective of making the learners literate and numerate is achieved.
- interview questionnaires for DAEO's
LA's
LGL's
Participants
for collection of some basic background information and possibly some attitudinal information (motivation etc.)

2.3 Sampling

Some of the information collected will cover the whole programme e.g. number of learners, their sex, number of group leaders, their sex, number of LA's their sex and educational background, number of DAEU's their sex and educational background. Location of Literacy groups on a national basis as compared to latest population statistics, will also be dealt with for the country as a whole in order to facilitate the decision where to locate new groups.

However, for most other issues it is suggested that a sample be drawn. The sampling unit should for practical purposes be a literacy group, not a learner. If the sampling is made on a district basis the sample will be fairly representative. This approach can be described as a stratified, cluster sampling. There are, as far as we know, no special groups, e.g. farmers' groups, women's groups etc. within the national programme. The 1981 programme consisted of 23,600 learners in 1780 groups. The target for 1982 is 30,000 learners in 2,000 groups. A suggested sampling frame is shown in Table 1. In 1981 the resources i.g. Literacy Assistant were not strictly distributed on a population basis. The 1981 census gives a more appropriate base for distributing LA's over the country. It can be argued that consideration should be given to variations in literacy rate between the districts, when distributing resources. The Table gives an indication of the size of sample considered for the evaluation i.e. 5% of the groups gives 100 groups and approximately 1,500 learners.

The evaluation also should include other literacy groups using the material from Department of Non-Formal Education. No information is presently available on the number of groups of this kind. However, by contacting persons or organizations that have participated in training courses for voluntary organizations it should be possible to get the necessary information. The Tirelo Setshaba Scheme and the Prison's Department Literacy groups will also have to be included in the evaluation. For the purpose of evaluating the long range impact of the scheme it will be necessary to include a sample of 1981 participants. The 1981 groups will have been dissolved by October - November, 1982, and the problem of how to reach the 1981 participants will have to be discussed with LA's and LGL's.

Data regarding DAEU's and LA's will be collected for the total population as numbers are small. For LGL's the data collection will be restricted to those in charge of sampled groups.

2.4 Data analysis and interpretation

Though a more indepth data analysis might be carried out at a later stage, using computer facilities, the initial analysis will be made at the primary level by LGL's and LA's. The intention is to prepare worksheets for the compilation of the data collected for each group so as to facilitate the interpretation of the data by the implementators of the programme. Special worksheets will be prepared for district level analysis to be done by DAEO's. This is admittedly a labour intensive form of data analysis, but it is expected to have certain benefits regarding feed-back and use of evaluation results. The worksheets mentioned above will have to be very carefully worked out and tested before use.

DAEO's will be encouraged to choose one area or problem of special interest to their district and make a special study on this problem. There will thus be some special case studies that do not cover the whole country.

2.5 Reporting and use of evaluation results

A major report will be prepared as documentation of the evaluation and for the sake of those interested in detailed aspects of the programme. For wider use a shortened summarized report will be printed. A tentative outline is enclosed in Appendix B.

The main purpose of the evaluation is, however, not to produce a report. Rather the aim is self-evaluation, that is the results of the evaluation should be fed back into the programme in order to take corrective measures, where necessary, and thus improve the performance of the programme. It is also the intention to create a permanent monitoring system on the basis of the experience of the evaluation. For this purpose, the various people involved in data collection will also be called to meetings to assist in the interpretation of the data analysis. Thus DAEO's and LA's should call meetings at cluster level to discuss with LGL's, and DAEO's should call meetings at district level to discuss with LA's etc. These discussions can take place in connection with the meetings held regularly with different categories of Staff. It is suggested that members of the evaluation team participate in some of these meetings.

TABLE 1
PROPOSED SAMPLING FRAME

Districts including urban areas	Census 1981	%	1981		1981			1982 Target				No. of groups 5% Sample
			ULT Target %	ULT Enrolment %	LGL %	LA %	Group %	%	"	LAs	Groups	
Southern + Jwaneng	125,000	13	8	11	12	11	13	13	3,900	18	260	13
South East + Lobatse+Gaborone	109,600	12	15	12	13	15	12	12	3,600	16	240	12
Tloveneng	115,600	12	19	17	17	16	17	12	3,600	16	240	12
Botatleng	42,300	5	10	8	8	10	9	5	1,500	7	100	5
Central+Orapa Selibe-Phikwe	357,300	38	23	27	25	19	24	38	11,400	51	760	38
North East+F/town	67,000	7	5	6	5	5	6	7	2,100	9	140	7
North West	76,300	8	8	9	9	8	8	8	2,400	11	160	8
Ghanzi	19,700	2	6	2	3	9	2	2	600	3	40	2
Galagadi	24,000	3	6	8	8	7	3	3	900	4	60	3
TOTAL	936,600	100	100	100	100	99	99	100	30,000	135	2,000	100

3. Organisational/Management Aspects

3.1 Personnel, institutions and agencies to be involved

Administrative Location

The Department of Non-Formal Education will be the administrative headquarters for the evaluation.

Persons responsible for design and execution

The Head of Informal programmes, the Literacy Coordinator and the Evaluation officer of the Department of Non-formal Education in co-operation with a Planning Officer at the Ministry of Education will be the main responsible persons for design and monitoring of the evaluation. It is planned that the major part of the execution will be done by District Adult Education Officers and Literacy Assistants.

Composition of evaluation team

In addition to the persons mentioned above as donor/ responsible / agencies will be invited to appoint representatives.

Composition of advisory committee

The advisory committee consists of two representatives from the Department of Non-formal Education, the Dean of Faculty of Education, one representative from the Institute of Adult Education and a Planning Officer from the Ministry of Education. Additional members might be considered e.g. NIER.

Preparation of Staff

In-service training of Department of Non-formal Education staff, especially District Adult Education Officers and Literacy Assistants regarding evaluation methods and procedures to be used.

The training sessions will consist of short half-day meetings in connection with regular staff meetings, staff training courses and National Literacy Committee meetings. Some longer training sessions are also planned i.e. one one-week course in the beginning of September for training of DAEO's in carrying out the evaluation and in basic monitoring skills. If funds can be secured some of the resource persons from the Project Development Workshop in Tanzania (29) March - 6 April, 1982) will be requested to assist in this training course.

	DAEO's will carry out training at district level of LA's in charge of clusters in which sampled groups are located. This training will take place in the beginning of October. As a follow-up and further development of monitoring skills the outcome of the evaluation will be dealt with during the March, 1983 refresher course for LA's.
--	---

3.2 Financial Implications

No detailed budget for the evaluation is available as yet. The major part of staff costs is included in the budget for the literacy programme. The literacy group leaders are employed on a honoraria basis and will have to be paid for extra hours worked in connection with the evaluation. Printing of instruments will be done within the Department of Non-Formal Education. Costs for material (paper, printing plates etc.) will have to be born by the evaluation study. Though most of the execution of the evaluation will be done within the normal work of the department staff some extra travelling costs will no doubt be incurred. It is assumed that donor agencies carry their own costs for participation. Donor agencies will also be requested to contribute towards the extra costs budgeted below. See Appendix E.

3.3 Timetable

Evaluation activities to be undertaken at different points in time are indicated on the enclosed chronogram. Appendix F.

4. Intentions, considerations and constraints

The evaluation outlined above is quite comprehensive and ambitious. It might appear "descriptive" and "summative". This is correct to a certain extent. However, it is also "explanatory" and "formative". It is important at this stage (mid-term) to make a description of the progress of the programme so far, not least for policy makers and donor agencies. It should be stressed, however, that the main purpose of the proposed evaluation is the improvement of the monitoring system. This is also the reason for the heavy reliance on programme personnel especially the LA's in the collection as well as analysis of the data. The presentation of the evaluation plan and its objectives will be a crucial factor. It is essential that everyone involved realises the importance of the evaluation.

The long term impact of the programme is the issue receiving least attention in the evaluation. The reason for this is not only that the issue is difficult to tackle but also that it is early in the programme to make such an attempt. Consideration will be given on methods to be used for estimating impact at a later stage.

The environmental or contextual factors will also be covered but only superficially. These might be very important per se but by definition these factors can not be monitored or changed by the programme. It is also felt that this aspect is difficult to approach within the framework of the kind of self-evaluative study proposed here. Rather it should be dealt within a separate study.

How realistic is the evaluation outline? It is possible to implement it? A basic feature of the outline is that it draws on already available resources. This will ensure implementation to a certain extent. However, if no extra resources will be available the outline will have to be adjusted somewhat.

PROGRAMME OBJECTIVES/ACTIVITIES(L) LONG RANGE

- eradicate illiteracy by 1985.
- provide basic education to those lacking formal education
- support attainment of national development goals
- support attainment of district development goals

(O) OUTCOMES

- To enrol 30,000 per year in literacy work (including voluntary organisations)
- To enable a person to comprehend those written communications and simple computations which are part of daily life.

(P) PROCESSESEducation

Teaching method
 Training of staff
 Cultural aspects taken into consideration
 Motivation of learners

Administrative

Supervision/monitoring of programme include National Literacy Committee, Non-Formal Education activities. Life long education/link to other education activities.

Distribution of material, communication.

Financial administration and control.

(I) INPUTSHuman

Staff
 Participants

Material

Primers
 Teaching support material
 Broad sheet
 Vehicles
 Meeting place of literacy groups
 Resource material at District Office
 Follow-on material

Financial

Financial support to the programme

(E) ENVIRONMENTAL SETTING

USES OF EVALUATION INFORMATION

LONG RANGE

- Policy makers
- Donors

OUTCOME

- Policy makers
- Donors
- Ministry of Finance & Dev. Planning
- Ministry of Education include Non-Formal Education

PROCESS

Educational

- Non Formal Education Staff

Administrative

- Non-Formal Education staff
- Ministry of Education
- Ministry of Finance & Development Planning
- Donors

INPUTS

Human

- Policy makers
- Donors
- Ministry of Finance & Development Planning
- Ministry of Education
- Non-Formal Education - staff

Material

- Non-Formal Education - staff

Finance

- Ministry of Education
- Ministry of Finance & Development Planning
- Donors

ASPECTS/EVALUATION QUESTIONS

LONG RANGE:

- use of literacy skills
- other impacts: scholarization of children
participation in kgotla meetings
participation in VDC activities
increased production

OUTCOME

- enrolment in literacy groups, MFE and non-MFE
- achievement in literacy and numeracy skills.

PROCESS

- How many hours of training is the programme?
- Is the teaching method effective, adapted to the situation?

Education

- What is the training programme for DAFO's, LA's LGL's?
- Are there any deficiencies in this training programme?
- How is the language "problem" dealt with? "Non-Setswana speakers".
- What, and how strong is the motivation of the learners?

Administrative

- National/regional coverage by programme?
- What supervisory/monitoring procedures exist, deficiencies?
- How are learners recruited?
- What is currently planned and done for the continued education of participation?
- What is the relationship to other extension activities including libraries?
- Does the necessary material reach the group?
- Does the financial administration and control function work satisfactorily?

INPUTS

Human

- | | |
|--------------------------------|--------------------------------|
| - Non-Formal Education staff) | |
| - DAFO's) | |
| - LA's) | What resources do these groups |
| - LGL's) | represent? |
| - Participants) | |

Material

- What kinds and amounts of material has been produced?
- Any locally produced material?
- Is it well suited i.e. design, illustrations, size of script?
- Various modes of transportation is used. Advantages-disadvantages.
- Which officers do have satisfactory transport? How is it used?
- What venues are used for literacy groups and how are they equipped
- How are district offices equipped and how is the equipment used?
- What is the content of the teaching material? Is it adjusted to the needs of the learners?

Financial

- What is the amount and source of financial support to the programme so far
- Cost per learner
- Estimated and actual cost
- The historical background of the programme. Relationship

INDICATORS/LEARNERS TO BE USED

LEARNER RANGE

1. - Number of letters written last month.
- Number of letters received and read last month.
- Number of other material read last month: specify
specify uses of numeric skills.
2. - Number of contacts with school/headmaster, teacher last month/
less-same-more than last year.
- Number of kgotla meetings, attended last month/Less-same-more than
last year.
- Number of VDC activities in which learner participants/less-same-
more than last year.
- Number of children of school going age/number of children attending
school.
- Paid employment during the year yes-no/Paid employment last year
yes - no.

OUTCOME

1. - Number of learners enrolled ILE, Tirelo Sechaba, Prisons, other
groups.
2. - Literacy test (D1) result 1981 and 1982 participants
- Numeracy test (D2) result 1981 and 1982 participants.

PROCESS

1. - Total number of hours of instruction received by 1981 participants
by District.
2. - Qualitative judgement of teaching method.
3. - Description of various aspects included in training programmes
of staff.
4. - Length and timing of training programmes of staff.
5. - Relevance of training of staff.
6. - Drop-out rate of participants.
7. - Mean number of hours attended per month by District.

Administrative

1. - Number of LA's and Groups by District and 1981 census (if possible
by educational achievement in District).
2. - Description of organisational structure.
3. - Organised feedback of information to headquarters? What kind?
4. - Description of promotion of programmes at national, district & local level.
5. - Number of articles in Daily News.
6. - Number of radio programmes
7. - Number of kgotla meetings
8. - Description of formal-non formal links, if any, and links other.

9. - Number of groups which had all necessary material when they started.
10. - Number of audit queries.

INPUTS

Human

- 1 - Number and function of non-formal education headquarters staff,
- 2 - Number, sex, age, education background DAE0's
- 3 - Number, sex, age, education background LA's
- 4 - Number, sex, age, education background of LEL's (previous and present work).
- 5 - Number, sex, age of participants, physical handicaps.
- 6 - Number of participants indicating various types of motivation for joining the programme.
- 7 - Number of participants with other mother tongue than Setswana.

Material

1. - Number of Primer 1 printed
2. - Number of Primer 2 printed
3. - Number of Primer 3 printed
4. - Number of Primer 4 printed
5. - Number of Primer 5 printed
6. - Number of Literacy kits produced
7. - Number of Broadsheets (issues and numbers)
8. - Relevance of material
9. - Relevance of design
- 11 - Number and kind of transport facilities per District (+their use) }
- 12 - Number of venues of different kinds: primary school classroom)
district/council facility) OR LG
LEL's home)
other)
- 13 - Equipment of venues: Chairs for all - some - none)
Desks for all - some - none)
- 14.- Equipment of district officers: typewriter
duplicating machine
stire
- 15 - Kinds of locally produced material:

FINANCE

- 1 - Sources of finance with amounts,
- 2 - Financial breakdown of cost, salaries, material, training
- 3 - Per learner cost
- 4 - Budget vs. actual cost

ENVIRONMENTAL SETTING

- 1 - Description of national setting: background of programme, coordination with other extension efforts, structural hindrances.

DATA COLLECTION INSTRUMENTS/SOURCE OF DATA

V = available

X = to be collected

			V X	Data Collection	Person or Office responsible for compilation
L	1.	Interview of participants	X	LA or LGL	Evaluation Team
	2.	Interview of participants	X	LA or LGL	"
O	1.	NFE: LGL	V	DAFO	"
		Tirelo Sechaba : Quest.to field officers	X	T.S. Field Off.	"
		Prisons: Quest.to Home Affairs Voluntary Org: quest to NFE Course participants	X	Prisons	"
			X	NGO Lit leaders	"
	2.	D 1 Literacy test or self assessment	X	LGL?	"
		D 2 Numeracy test or self assessment	X	LGL?	"
P	1.	L 1 forms for sample 1981	V	DAFO + NFE Eval.	NFE evaluation
	2.	Observation of 2 groups IG2	V	LA	Evaluation Team
	3.	Narrative of training progr.	(X)	Head informal Pro	Head info. pro.
	4.	Narrative, chronograph	(X)	Head informal Pro	"
	5.	Summary of evaluation + quest.LA	(X)	Head informal Pro	"
		" " "	X	LA or LGL	Evaluation Team
	6.	L 1 forms for sample 1982 September, October	V	NFE Evaluation	NFE Evaluation
	7.	L 1 forms for sample 1982 September, October	V	NFE Evaluation	NFE Evaluation
Ad	1.	LGL + 1981 census results	V	NFE evaluation	"
	2.	Org. structure chart	V	P.O.	P.O.
	3.	Interview with NFE Staff, NIC activities, IG2, IG3, IG4.	(X)	P.O.	P.O.
	4.)				
	5.)	Narrative	(X)	Head Informal Pro	Head infor.m.P
	6.)				
	7.	Quest. LA.	X	DAFO	Evaluation Team
	8.	Narrative + DEB + VEB + Kettle meetings	(X)	CEO (NFE)Eva.Team	CEO (NFE) Evaluation Team
	9.	Inf. of LGL	(X)	LA	"
	10	Public reports + P.O. report	V	MTDP (PO)	MTDP (P.O.)

I	1.	Establishment register	V	CEO (MFE)	CEO (MFE)
	2.	Quest. DABO	X	Evaluation Team	Evaluation Team
	3.	Quest. LA	X	DABO	"
	4.	Quest. LIL	X	LA	"
	5.)				
	6.)	Quest. participants or	XV	LA or LIL	"
	7.)	LG 9			
Ma	1.)				
	2.)				
	3.)	Votes lodger or printer's	V	MFE Eval. +	MFE Evaluation
	4.)	lodger		Head Informal	
	5.)			programmes	
	6.)				
	7.)				
	8.)	Content analysis	X	Evaluation Team	Evaluation Team
	9.)	judgement analysis	X	Evaluation Team	"
	10.)				
	11	Quest. DABO	X	"	"
	12)	LG 2	V	LA	"
	13)				
	14	Quest. DABO	X	Evaluation Team	"
	15.	Quest DABO	X	"	"
F1	1.	MFE P.O. Reports to	V	MFE P.O.	MFE P.O.
	2.	Monthly computer print outs	V	MFE P.O.	MFE P.O.
	3.	Monthly computer print + LG 6	V	"	"
	4.	Estimates + Final accounts	V	"	"
E.	1.	Background documents P.O.s etc.	V	Min. of Ed. P.O. MFE	" MFE

REPORT OUTLINE

- Front Cover** : Title of programme and Location
Name of evaluator(s)
Period covered
Date of report
- I Summary** : What was evaluated?
Why was the evaluation conducted?
What are the major findings and recommendations?
- II Background Information:** Origin of the programme
Goals of the programme
Participants
Characteristics of the programme materials, activities, and administrative arrangements. Staff involved
- III The Evaluation Study** : Purposes of the evaluation
Evaluation design
Outcome measures: Instruments used
Data collection procedures
Implementation measures (process)
: Instruments used
Data collection procedures
- IV Results** : Results of outcome measurements
Results of implementation measurements
Informal results
- V Discussion of Results** : How certain is it that the programme caused the results?
How good were the results of the programme? (Modifications suggested by staff).
- VI Costs and Benefits** : Method for calculating Costs and Benefits
Costs associated with the programme
Benefits associated with the programme.
- VII Case Studies**
- VIII Conclusions and Recommendations** : Conclusions
Recommendations regarding the programme
Recommendations concerning monitoring of the programme.

Proposed timetable for planning and implementation of
evaluation of National Literacy Programme

	1982											1983				Staff Staff Involved
		A	M	J	J	A	S	O		J	J	F	M	A		
<u>Planning</u> <u>Development of Ev. Design</u>		---	---	x												P.O. Ev. team
<u>Training of staff</u>		x	x		x		x	x					x			P.O. NFE DAEO LA
<u>Development of instruments</u>				---												P.O. NFE Ev. team
<u>Printing of instruments</u>						--										Printer
<u>Inf. collection non-NFE GROUPS</u> <u>Sampling</u>						x	-									NFE ev P.O. NFE
<u>Implementation</u> <u>Literacy test</u>								---								LGL
O. DAEO							x									Lit. Coord.
LA								x								DAEO
LGL								x								LA

	1992										1993				Staff involved
	J	A	J	J	A	S	D	D	J	F	J	A			
Participant							---							LA?	
Hor-NFE groups							---							NFE ev.	
Other data collection		-----												See detailed data coll. par	
<u>Data Analysis</u> Cluster level analysis							--							LGL LA	
District Level analysis								--						LA DAEO	
National level analysis								---						DAEO ev. team	
<u>Interpretation and reporting</u> Cluster level							--							LA	
District level								--						DAEO	
National level									--					Ev. team	
<u>Dissemination</u> Shortened version of final report										--				Ev. team	
Meeting with donors		x			x	---				x					

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 4



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOMES AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

Appraisal Studies of the Adult
Education Programme in
Bihar, Gujarat, Maharashtra,
Rajasthan and Tamil Nadu

by

RS Mathur
Prem Chand

ADULT EDUCATION IN GUJARAT: An appraisal by
Atul Sarma, Vimal Shah, Bhanumati K. Parikh
Ahmedabad, Sardar Patel Institute of Economic & Social Research. 1979 *

Introduction

The study was undertaken by the Sardar Patel Institute of Economic & Social Research, Ahmedabad, at the request of the Ministry of Education and Social Welfare, Govt. of India, and with the concurrence of the Govt. of Gujarat (Education Department) and the State Resource Centre for Adult Education. The study was initiated in the third week of August, 1978, field work lasted about a month (end of September to third week of October) and the report was ready by middle of January, 1979.

Objectives

The main objectives of the appraisal study were:

- to add to the data on the adult education centres;
- to examine the working of the AECS in relation to the objectives of NAEP;
- to identify the strengths and weaknesses of the AECS and factors responsible for these; and
- to indicate areas for action.

* Summary by R.S. Mathur, Deputy Director, Directorate of Adult Education, Govt. of India, New Delhi.

Methodology

(a) Sampling Design

A sample survey was designed. The universe for the sample consisted of all adult education centres which started on or before April 15, 1978 and functioned for 4½ to 7 months, until the start of the study (August 31, 1978). Thus, 1753 AECs formed the universe. A ten percent sample of these centres was chosen randomly, agency-wise. The sources of data collection included:

- the voluntary agencies (10 per cent adult education centres)
- Instructors of sample adult education centres.
- Learners (3 per centre randomly)
- Drop-outs (1 per centre randomly)

(b) Instruments for data collection

The following instruments were designed and used:

- Questionnaire for the voluntary agencies: Information about the organization and administrative set-up of the agency, previous background in organising educational and welfare programme, past experience in conducting adult education work, procedures used in selection of areas and functionaries, preparation and supply of materials to the centres, problems faced etc. was to be obtained through this questionnaire.
Method: Mailed questionnaire.
- Questionnaire for the Instructors: Information about Instructors' socio-economic and demographic background, educational level, motivation, cooperation from other development functionaries, difficulties encountered etc. was elicited through this instrument.
Method: personal interviews.

- Questionnaire for the Learners: Information on learners' achievement in literacy, social awareness and functionality, availability of materials for learning etc. got covered under this questionnaire.

Method:personal interviews.

- Questionnaire for drop-outs: It was used to ascertain the magnitude of drop-out and identify the characteristics of the drop-outs.

Method:personal interviews.

- Schedule for Adult Education Centres (AECs): The records available at the AECs were to be consulted for purposes of collecting details of the learners enrolled, their date of joining the centre, attendance, etc.

Method:verification by the investigators.

- Investigators' Diary: This was used for noting the observations regarding physical facilities and environment at the AECs at the time of visit. Difficulties faced by learners and instructors in conducting the programmes were to be noted.

Method:unstructured observation.

(c) The Survey

The field teams were organised in a unique manner. 95 persons (64 college teachers and 31 government officials) made up the survey team. About 30 teams consisting of one college teacher and one government officer, and 34 individual teams comprising of college teachers were formed. Each team covered from one to three adult education centres. The teams were briefed in a two-day orientation programme organised. In all 169 instructors, 530 learners and 65 drop-outs were contacted for purposes of investigation.

Major findings

Data Audit (functioning of the AEC)

The appraisal revealed that almost all the voluntary agencies, except three, had been engaged in activities related to programmes of social reconstruction and welfare for a considerable period of time and 94 percent of the AECs started functioning within a short period of five months' time. Of the total sample, 11 AECs had already closed and half of these centres belonged to two voluntary agencies. Only 6.2 percent of the centres were not found functioning at the time of visit by the investigators. The reasons for such a situation were not studied. Further, it was noted that one third of the centres functioned for five months or less, which showed that the information with the State Resource Centre about the period of operation was not correct.

Location

It was found that 37 percent of the AECs were organised at the residence of the instructors, 28 percent in public buildings and another 28 percent in open places. It can be inferred that it facilitates the centre to be located at the Instructor's house and it is found convenient by the learners also. However, the acceptance of an Instructor by the learner and his other characteristics would be the important determinants in such an effort. Organizational and environmental inputs such as lighting and sitting arrangements at these places left much to be desired. These are bound to have implications for motivation of the learners to take advantage of the programme.

Distribution and Performance

About 30 percent of AECs were for women, about 40 percent for men and the remaining 30 percent were for both the sexes. About 70 percent of the centres were organised for weaker sections - 48.5 percent for Scheduled Tribes and 8.9 percent for Scheduled Castes and 10.1 percent for other Backward Classes. The

distribution of centres was very high (89 percent) in rural areas. The performance of the AECs could be judged by the fact that one fourth of the AECs conducted on an average 25-30 sessions per month, half of them between 20-25 sessions and the remaining showed a very unsatisfactory level of performance.

Participation of Learners - Enrolment, Attendance, Dropouts

An average of 32.4 learners were enrolled at the sample AEC. However, the proportion of learners present on the date of the visit was found to be 67.6 percent of the total shown in the attendance register and the mean size of an AEC worked out at 22.7. According to the records, about 67 percent of the centres showed an average attendance in the percentage group of 75-100. But the attendance on the day of the visit expressed in terms of percentage to the total enrolment, exclusive of drop-outs, shows that nearly 43 percent of the centres fell in the percentage group of 75-100. The obvious inference is that the data recorded in the attendance register has to be taken with a grain of salt. About 74 percent of the sample AECs had upto 10 percent of the total enrolment as drop-outs. The average drop-out level was estimated at 1.94 percent of the total enrolment. The problem of drop-out was also found to be connected with that of substitute enrollees. This indicates the need for enunciation of a policy towards drop-outs and substitutes.

Learners' Characteristics

The distribution of learners by age, sex and occupation indicated that, by and large, the target age group got the benefits from the programme and only a negligible percentage of learners was below 15 years. About 43 percent of the learners were women, 78 percent learners interviewed belonged to Scheduled Castes, Scheduled Tribes and other Backward Classes categories and 64 percent of them had agriculture as their main occupation. It was further noted that 16 percent of the learners had some background of schooling earlier to their enrolment in the AEC.

About three-fifths of male and female learners had acquired the skill to write. Additional 13 percent men and 18 percent women had acquired the ability the ability to write their names. The arithmetic ability of 38 percent learners was considered 'good' by the Investigators. Against the back-ground of prior schooling of about 16 percent of the learners and the fact that two-third of the centres had functioned for six months or more, such a level of achievement may be considered as very modest.

Achievements of the learners in social awareness and functionality were found to be on the lower side. Favourable responses to the questions on these two aspects varied from 44 to 68 percent. Ill equipped AECs was the main reason attributed to such a situation which was caused by non-availability of suitable teaching materials with the instructors. This needs greater attention than given at present by VAs and the SRC.

Instructors' background

A study of the distribution of the Instructors revealed that 34 percent of them were women. About 64 percent of the Instructors belonged to the age group of 20-30. More than one half of the Instructors were below SSC level and one fifth had passed the SSC examination. It is obvious that the educational background of the Instructors influences the instructional process, but this can be compensated to some extent by locating such instructors who have dedication and other personal traits which can be developed through a process of training. It is because of their inadequate education, perhaps, that the instructors "did not seem to have made much effort to make the classes more profitable to the adult learners by preparing charts and other teaching aids". Nearly 80 per cent of the instructors had undergone some training before taking up the work of AECs. Nearly 19 per cent of the instructors had problems regarding delayed payment of remuneration and 71 per cent of

SSC - means Senior School Certificate and can be equated with High School/Matriculation level.

the instructors felt that it was very inadequate. This appears very reasonable, particularly when it comes from those who joined the AECs for eking out their livelihood. While the percentage of such people does not seem to be very large (14 per cent or less), the question deserves some consideration of the policy makers.

Performance of Voluntary Agencies

The appraisal also helped in knowing the performance of the voluntary agencies against the selected criteria. This information was to be used in taking rational decisions about the capability or otherwise, of the agency in organizing the work. Decisions on whether to continue, expand or discontinue the work of any voluntary agency could thus be taken on the basis of evidence.

The voluntary agencies were ranked in terms of important characteristics under three types.

In the first type of ranking, the performance of the learners in terms of the three R's taken into consideration. The measures of the three R's performance were: (1) the percentage of the learners who could, on their own, fill up the required portion of the learners' questionnaire and (2) the percentage of learners (to the total interviewed under the VA) who were judged 'good' and 'fair' by the investigators after giving small tests in elementary operations of arithmetic.

In the second type, the performance in terms of the three R's, social awareness and functionality was considered. It included seven additional characteristics such as information relating to Primary Health Centres, agriculture and veterinary centres, cooperative societies, family planning, dowry system, minimum age for marriage and occupational information.

In the third type, 25 important characteristics, including those in the second type, were included. Thus, additional 16 characteristics were: information

about availability of learning materials to learners, regularity of classes, possibility of spread effect, target group bias i.e. the proportion of the AECs run for SC, ST, and OBC, Instructors' educational level (SSC and SSC+), instructors' training, Instructors' opinion on the availability of teaching materials on time, inadequate quantity, pertaining to occupation, pertaining to rights and duties, pertaining to common diseases, instructor's initiative, his satisfaction from work, regularity in receipt of remuneration, his satisfaction with the amount of remuneration and his willingness to continue in subsequent years.

The broad methodology followed in working out the rank was as follows:

- the percentage of favourable responses to the total number of responses for each of the VAs was worked out;
- scores were given to each of the characteristics for each VA in a descending order corresponding to the percentages;
- the scores were averaged out across the characteristics;
- the VAs were ranked in a descending order on the basis of the average score;
- all the characteristics that figures in computing the ranks were treated equally.

The agencies were ranked on the basis of the above methodology in seven class intervals. The inference drawn was that the organisation did not appear to be neutral in the performance of AECs.

On the whole, while the N/AEP in Gujarat was generally found to be addressed to the target groups kept in view under the N/AEP and it was found to have some other commendable aspects, all things considered, its achievement in terms of spread of literacy (3 R's) were rather modest and more so, in terms of social awareness and functionality.

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 5



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOMES AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

Feedback Action - The Essence of Monitoring

by

RS Mathur

in Reading Material. National Workshop on Monitoring and Evaluation (June, 10-15, 1982) organised by Directorate of Adult Education, Ministry of Education and Culture, Govt. of INDIA, New Delhi

FEEDBACK ACTION - The essence of monitoring*

Hundred per cent coverage of the illiterate adults in the age group 15-35 by 1990 through non-formal education is the goal set forth in the Sixth Five-Year Plan document under the Minimum Needs Programme. Besides, removal of adult illiteracy, together with the spread of universal elementary education, forms part of the New 20-Point Programme. The new 20-Point Programme is a selection of Plan Schemes. Its essence is full commitment to achieve the targets and objectives which have been identified and selected.

The achievement of the goals and targets under the Adult Education Programme necessitates careful attention to be given among other things to the following:

- a) preparation of detailed operational plans, indicating year-wise, district-wise, programme-wise targets proposed to be achieved upto 1990,
- b) resource mobilisation - in terms of money (financial provisions), men (personnel to handle) and materials (both instructional and physical), and
- c) proper and effective monitoring of the implementation process and continuous assessment of the outcome/results.

It is a known fact that proper implementation of a programme is not automatically guaranteed just because it has been conceptualised with utmost care and that it has been

* Prepared by RS Mathur, Deputy Director, Directorate of Adult Education, New Delhi, 1982.

planned and financed adequately. Some very imaginatively drawn up schemes, and innovative programmes have had a sad experience in the past, mainly because of flaws in implementation and absence of effective monitoring system. Therefore, provision of a built-in system of monitoring and evaluation is advocated to keep a watch on the implementation process, ascertain regularly how close or behind are we in relation to the targets, why are we in that state and what is the net result of the investments being made. In short, monitoring helps in.

- i) checking the wastage (by identifying the reasons for poor performance and taking remedial action),
- ii) maintaining a control over the quality of the programme, and
- iii) optimising the results

Successful implementation of a mass literacy programme depends to a large extent, on an efficient monitoring system.¹ It would be considered effective if -

- the flow of information is fast;
- the data furnished is accurate and reliable;
- the data generates requisite feedback;

Let us see how feedback is central to monitoring; how it constitutes the essence of monitoring. Consider, for example, the first two of the above features of an efficient monitoring system. Monitoring, as distinct from evaluation, requires early

 1. Unesco - International Institute for Educational Planning, Paris. Planning and Administration of National Literacy Programmes - Final Report on IIEP Workshop: Arusha, Tanzania, 1981.

information about how the programme is operating. Therefore, it has to be fairly quick and fast and if the flow of information is delayed, one should look into the reasons for it and try to remedy the causes for delay - by constantly reviewing the flow, identifying bottlenecks and removing them. This itself results in feedback action to improve the flow of data. Again, when the data furnished is scrutinised and analysed, it has to be seen if what is given is correct, truthful and reliable. Data audit may be necessary to verify facts stated earlier. Thus, feedback on its reliability has to be provided; in the absence of which tendency to misreport may develop. Therefore, as distinct from pure reporting, here the purpose is to ascertain factual position, and give suitable feedback advice to the concerned person/agency.

Thirdly, the data, as such, is not used for compilation purposes. The returns from the monitoring system have to result in action, where called for. It is not merely for purposes of generating statistics but for ensuring action that monitoring is necessary. Without adequate feedback, monitoring is purposeless. Moreover, it has to be participatory - the reporting authority must know about his/her own efficiency in reporting and has also a right to know what is being done at the reported level about the issues/problems/facts stated in the monitoring returns. If the two-way flow is absent, it is not monitoring. In the absence of this, the status of the programme will remain unaffected, the quality will not improve, the results will suffer. The diagnosis of the ailment, in the absence of proper treatment, will mean decay and loss.

Where are we, in relation to the above expectations from the monitoring system which has fortunately been operational

since the programme assumed large dimensions in 1979. This monitoring system is reviewed periodically, at least once a year, where the "monitoring process itself is monitored" and further reorientation of the personnel involved in it is done.

It is assumed here that details of the monitoring system as applicable to the AEP are already known. The subsequent analysis is, therefore, restricted to the three factors as mentioned on page 2 and focus on feedback arrangements as conceptualised and as operational.

a) Flow of Information

The monitoring system has been able to maintain a steady flow of information from the Adult Education Centre to the Project, from the Project to the State/UT and from the State/UT to the Central (National) level. During the early phases of the programme, monitoring returns were not regular and were often delayed. The time lag is continuously reducing. Regularity in reporting is improving and there are now fewer defaulting projects/States than before. However, the Monitoring System has yet to improve substantially with respect to coverage under the Colleges/Universities and the Nehru Yuvak Kendras, practically in all States/UTs. Compared to the programme of school education, the monitoring of AEP can be described as satisfactory.

It may, however, be noted that there is still

sufficient scope for reducing the time gap in information flow at every level and this is being attempted with faith and conviction. With the support of field functionaries and their proper orientation this should be possible.

b) Data Quality

It is difficult to make any reasonable generalisation about the quality of data yielded by monitoring system. The quality of data available from States/UTs at the national level shows that there are various shades of 'good' and 'bad' examples. This difference is noticeable by States/UTs, by projects, by type of information sought and so on which suggest that there is a need for

- (i) special orientation programmes in Monitoring & Evaluation for the field personnel (P.Os, Supervisors and Instructors), (ii) training of Statistical Assistants in processing tabulation and compilation procedures,
- (iii) sample checks of the data reported to establish reliability and curb the tendency to present "estimates" rather than "actuals",

in howsoever small proportion this tendency may be found. The project agencies have a special responsibility in this regard, because if data

reliability is questioned, the whole foundation of monitoring collapses. Involvement of local community in the planning and implementation of the programme would act as a deterrent to mis-reporting. Similarly, formation of district parties, Project Committees (whose members may pay surprise checks to centres) would assure provision of more reliable information.

The appraisal studies of the programme in States like Bihar, Gujarat, Maharashtra, Rajasthan and Tamil Nadu suggest that there is a definite need to keep the attendance and other project records more 'straight' as appreciable differences were found in the attendance figures recorded and actually present in the AECs on the date of visit by the evaluators. The authorities at every level - State, Project, Supervisory and Centre - have to take note of this and should work out mechanisms to ensure a reasonable degree of reliability in flow of information.

a) Feedback arrangements

Apart from regular flow of information provided in the monitoring system, what is interesting is that there exist intermediary levels where

information flow is specifically meant for action. The Supervisors, the District Adult Education Officers and the State Governments consider specific action points. Besides, between every two levels - the reporting and reported - there is an arrangement for two-way flow of information. There are four distinct lines of flow²:

- 1) The Principal flow - Information going from the adult education centres to project agency, then to state adult education office and finally to the Directorate of Adult Education. The latter providing feed-back down the line by the same channel.
- ii) The half-way feed-back - This is intended to be provided by each stage without waiting for the feed-back from the Directorate of Adult Education. Thus, the project agency and the state adult education officer are expected to provide immediate feed-back on the basis of information received and they are also supposed to advise the intermediate stage people, (district adult education officer and supervisors) to follow up the feed-back sent to the operational levels (project agency and adult education centres).
- iii) The intermediate level feedback - ~~The operational level people have~~ to send a copy of the information to the intermediate level - thus the adult education centre is supposed to send a copy to supervisors, the project agency to district adult education officers and the state adult education officer to the State Government. These intermediate level agencies are expected to send a feed-back immediately.

2. Anil Bordia, Planning and Administration of National Literacy Programmes - The Indian Experience - Research Report - 47.
International Institute for Educational Planning, Paris. 1982 (Unesco)

- iv) Policy level feed-back - The Directorate of Adult Education, which is supposed to receive information in respect of the entire programme, is expected to brief the Ministry of Education, the Council of Ministers and the Parliament's Consultative Committee on Education. On Policy Issues the Ministry is expected to advise the Directorate of Adult Education and also directly the State Governments.

However, in operation, the feedback aspect has been found to be inadequate. Usually, the reports received do not evoke immediate feedback. The concept of feedback and its significance has yet to receive full attention?

It has also been observed that the items on which feedback action is often desired relate to

- 1) Financial matters - Usually the problems mentioned concern delays in release of funds, rigidity of financial pattern, non-existence of budget provision for certain items, problems relating to operation of budget, particularly because of savings or inability to spend on certain items because of rise in costs, poor honorarium to instructors, etc.

Since any matter relating to financial pattern has larger policy implications these problems are referred to Ministry of Education, Government of India/UGC/as the case be. The exact difficulties are taken up for consideration at the appropriate level and decision is taken after

examining its feasibility for accepting it. A concrete instance of how problems emerged from monitoring got resolved is cited at Appendix 1.

11) Administrative matters

Such problems generally relate to delays in creation of posts, non-appointment of personnel as per norms suggested in the project structure, inadequate staffing, etc. However, for most of these problems it is the State Government's decision which counts. But if the difficulties noted are found to hamper the programme seriously, these are referred to the State Government with suitable advice. Moreover, such problems are also discussed at higher levels and in forums like the Regional Conferences of Education Secretaries and appropriate decisions are taken during such meetings.

To cite an example, one of the States of the Eastern Region experienced the following difficulties:

"One of the major problems faced in implementation of the programme in the State is that the norm of appointing one supervisor for every 30 centres has not been accepted by the State Government. The supervision work, therefore, suffers to a great extent. Proper supervisory arrangements are needed. Full-time District Adult Education Officers have not been provided and the posts of supervisors have not

been created. The responsibility of supervision has been given to Social Education Organisers working under Community Development Blocks. This arrangement has not worked satisfactory and the functioning of the centres was found to be deteriorating on this account." This point would have to be clarified at the earliest in consultation with the Government concerned.

iii) Technical matters

There are very infrequent references to problems under this category. It is just possible that the problems encountered are followed up at lower levels, but such reports of feedback action are conspicuous by their absence. Sometimes problems of lack of motivation are raised, but those too are in insignificant proportions. Inadequate cooperation of development departments at field level is another difficulty which sometimes gets a passing reference. Similarly, very little is said about organisation of training programmes, teaching-learning materials and other technical inputs. Lack of adequate cooperation in providing monitoring returns by certain agencies like the Nehru Yuva Kendras, Colleges/Universities, etc. to the State level officer incharge of monitoring is, however, invariably reported.

It is here that the feedback is most crucial as it has a direct bearing on the output, performance and results.

The feedback available through the monitoring returns is pursued for action as is necessary, at the national level but it would be interesting to know how and what type of feedback action is taken at the Supervisory, Project, District, Directorate and State levels. Examples of reports on such feedback action would strengthen the monitoring system as it would then get the due recognition from the reporting agencies. The other sources of feedback are personal visits, discussions, meetings, conferences and findings of evaluation/appraisal studies. The findings of appraisal studies do indicate areas for action so that the implementation process gets improved. These are also followed up for action and most of them have already been jointly discussed by the concerned parties.

- ii) The savings, if any, out of the admissible sanctioned grant for each project could be utilised on the following essential items which are presently not included in the financial pattern of the scheme:
- a) purchase of tat-patties/durries for seating the participants in the adult education centres;
 - b) purchase of boxes/trunks for storing of teaching-learning materials - one for each adult education centre; and
 - c) in case the adult education centres are electrified and kerosene oil/lanterns are not necessary, the funds earmarked for lighting in the financial pattern could be used on replacement of bulbs/tube-lights, etc.

Yours faithfully,

Sd/-
(KUMUD BANSAL)
DEPUTY SECRETARY

Copy to Directors of Adult Education (all States/UTs)

Sd/-
(KUMUD BANSAL)
DEPUTY SECRETARY

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 6



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOMES AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

Learner Evaluation Report of National Workshop on
Monitoring and Evaluation (June, 10-15, 1982)
organised by Directorate of Adult Education,
Ministry of Education and Culture, Govt.
of India, New Delhi

II

LEARNER EVALUATION

One of the criteria to determine the success of an educational programme is to gather the evidence about the impact it has on the learners. In determining the impact an important factor relates to the changes that the programme brings about in the learners and find out the gains that they derive from the programme. Such changes or gains in the learners can be known through a process of evaluation and therefore evaluation of learners becomes one of the central considerations in finding out the usefulness of a programme.

In the adult education programme the concept of learner evaluation got clarified in the form of guidelines that were finalised by the Directorate of Adult Education in 1980. The guidelines explained the importance of learner evaluation and also answered some of the crucial questions such as these - when can learner evaluation be done? what steps are involved in learner evaluation? who will do learner evaluation? how will the learner evaluation be done? and the stages at which learner evaluation can be done? A number of States/UTs started operationalising these guidelines. Some 4 or 5 SRCs also organised workshops on learner evaluation and tried to give a concrete shape to the guidelines on learner evaluation, prepared and circulated by the Central Directorate

of Adult Education earlier. They took steps to prepare tools for testing and measuring the achievements of learners. The progress has not been uniform - in certain SRCs substantial work was done and tools were prepared after organising seminars of project officers, supervisors, instructors and even learners. These tools were then put to a trial in the field and revisions introduced for larger application to determine the progress and achievements in literacy, functionality and awareness. However, there were quite a number of States/UTs which still persisted with their own methods of evaluating the learners and did not follow common norms or standards against which learner evaluation was to be done. It was, therefore, considered desirable to take stock of things and consider the following issues:

- review of current practices in learner evaluation?
- how reliable and valid are these practices in learner evaluation?
- how uniform are the evaluation practices and whether such uniformity is desirable?
- whether in the light of existing practices, systematic and dependable comparisons between projects, between States, etc. are possible?
- whether the learner evaluation is done against an accepted norm or each project follows its own standards/norms?

- how are the evaluation results interpreted? Whether the need for a common approach in interpretation of results exists?
- whether there is a need to work out equivalents of the educational levels attained by the learners in AECs with those at the primary level?

These issues¹ demanded serious consideration especially because on the one hand we wish to be sure about the number of persons who become literate and understand, in precise terms, the capabilities that they develop after their participation in the programme and on the other hand, we know the limitations of the person who organises the centre (instructor) and his background. The instructor is not always a committed person (the selection procedures, remaining as they are do not necessarily succeed in finding the right persons). Moreover, the training offered is, more often than not, inadequate. The kinds of factors which constantly act as demotivating elements (poor motivation of learners, lack of attractive honorarium, absence or lack of proper coordination between him/her and other development functionaries) also sometime render him helpless to perform his/her functions well.

-
1. Some Issues in the Evaluation of Learner Evaluation by RS Mathur in the Reading Material prepared for the National Workshop on Monitoring and Evaluation of Adult Education Programme.

It is not to suggest that we do away with measurement of results. The problems are genuine but not insurmountable. The argument is not that we eliminate the testing procedures, but that these should be not too ambitious and should be simple, practicable and worthy of use. At the same time it may be noted that very often the testing procedures themselves influence the quality of instruction/education and therefore there is a need not to sacrifice the testing methods to seek adjustments with the instructors quality but to ensure that the test material is reasonably objective and maintains a standard matching the programme objectives/goals. How this could be done was the task before the Workshop.

A review of the current practices revealed that different patterns of evaluating the learners were in vogue. Information gathered on a questionnaire circulated to the participants in the Workshop revealed that:

- out of 20 States/UTs, 7 States had not yet developed any clearly defined models of evaluating the learners.
- in other States/UTs which mentioned that such models were prepared, on further probing revealed that such models were not in the knowledge of project functionaries and in a number of cases the guidelines in this respect were not properly conveyed.

- in a few other States the guidelines on learner evaluation were fairly systematically formulated and were also being followed but such examples were not many. States like, Bihar, Haryana, Kerala, Maharashtra, Tamil Nadu and Uttar Pradesh had formulated the guidelines and also tools for learner evaluation which were being administered to determine the achievements of learners.
- in a few other States/projects it was noted that periodic evaluation was claimed but was not done in a very systematic manner. In some projects the tests were prepared and conducted but the validity of these tests could not be properly stated.
- only in a few States the models of learner evaluation prepared were helpful in determining the real outcomes of the learners.
- regarding the periodicity of conducting evaluation, the responses received from the participants indicated that in a number of States no systematic procedures for testing existed and in a few others the tests were conducted only at the end of 10-month period. In Andhra Pradesh, Bihar, Haryana, Kerala, Maharashtra, Rajasthan, Uttar Pradesh and West Bengal there was a practice of holding one or two tests during the course and such tests formed the basis of measuring the achievement level of learners. In one State no learner evaluation was systematically done and attendance was considered as the sole criteria to declare a person 'successful'.

- the manner of interpreting the results also revealed substantial variations. Practice of giving marks/scores was generally followed but the test results were interpreted either in the form of grades (A,B,C) or Good, Average, Poor or even by classifying them into successful and unsuccessful categories.

The above review necessitated that some commonly accepted procedures for learner evaluation be agreed and till such time as learner evaluation became an integral part of the teaching-learning process, minimum norms for testing learners achievements at least at different points of time be formulated. There was a consensus that such tests could be conducted at the end of 3 or 4 months (by which time sufficient progress in teaching-learning would have taken place) at the end of 7th month (by which time most of the readers, primers would have been completed and in some projects the second reader might have been introduced) and finally at the conclusion of the course to find out final output.

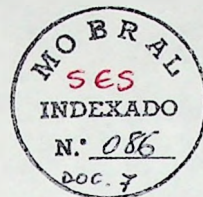
In view of the above considerations it appeared necessary to do the phasing of norms for learner evaluation according to the duration/content covered through the programme. The task was therefore undertaken of determining the expected level (norms) of achievement at the end of 4 months: suggesting the procedures that could be adopted for testing the achievements in relation to these norms

and to suggest the manner in which interpretation of results could be done. Similar exercise was undertaken for evaluation to be done at the end of the 7th month and on conclusion of the 10-month period. This task was undertaken in groups and a suggestive scheme for learner evaluation at the end of various phases is presented separately for literacy and functionality and awareness components at Appendix - IV and V. In addition, one of the groups also operationalised the steps involved in learner evaluation in relation to the period covered by the host project (Nilgiris Project in Tamil Nadu). The concretisation of this effort has been presented in Appendix - VI.

It was felt that in view of the experiences available, there was a need for those States/UTs to frame guidelines on learner evaluation where these had not yet been formulated. Further, concrete models would have to be developed for being practised so that by the time the next review takes place, the lagging States/UTs move forward and attain matching progress with those where already evaluation plans are fairly well developed.

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 7



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOMES AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

Conceptualizing Evaluation for
Lifelong Education

by
R Skager

in Lifelong Education and Evaluation Practice. Advances in Lifelong Education
Volume 4, Pergamon Press/UIE, 1978

CHAPTER 2

CONCEPTUALIZING EVALUATION FOR LIFELONG EDUCATION

The purpose of this chapter is to develop a conceptualization that will assist in drawing implications about the practice of evaluation under an educational system incorporating the principles of lifelong education. In order to do this it will be necessary both to specify what are the common features of activities that are referred to as "evaluation" as well as to isolate a set of characteristics that lead to useful distinctions between various types of evaluation that differ in significant ways. The strategy here involves making distinctions from which implications about practice can be drawn. Although it has already been pointed out that all aspects of evaluation cannot be treated in detail in this report, the distinctions to be made will at least lead to a comprehensive typology or framework for all of evaluation.

Defining Evaluation

Educational evaluation incorporates a great many diverse activities and functions. Assigning marks to learners, selecting individuals for special opportunities or training, appraising the performance of teachers or the quality of instructional materials, assessing the effectiveness of various approaches to instruction, and many other kinds of activities are all commonly referred to as "evaluation". As a result, attempts to formulate neat definitions of the field with clear-cut inclusion and exclusion criteria tend to produce unsatisfactory results. It is wiser to follow the advice of Payne (1974) and propose qualitative characteristics or dimensions, all of which must be present to a substantial degree if an activity is to be described as evaluation. The four characteristics given below have been formulated so as to apply to all of the specific

types of evaluation to be distinguished later (1).

Evaluation in education always involves an appraisal of the desirability of events or conditions associated with learning and teaching. The special nature of evaluation is closely tied to the appraisal of desirability. But in order to make such appraisals, information and evidence must be interpreted in terms of some value system. It has been common in much evaluation practice to ignore this underlying value base as if the grounds for establishing desirability were self-evident and universal. An evaluation theory and technology developed to serve traditional schools in a relatively homogeneous social context can of course pretend for a time that there is nothing controversial about whatever criteria are being applied. Still, even from within evaluation serious questions about this very issue have been raised of late, Stake (1973) providing a salient example.

All comprehensive views of what education should be like are built around values. Specific evaluation criteria are in turn derived from those values. Since prevailing philosophies of educational practice differ in many respects, it follows that associated evaluation criteria will differ as well. It is readily conceivable that an educational practice might be judged as desirable under one set of criteria and undesirable under another. Tractable, diligent, but teacher-dependent children might be a credit to a school in the eyes of some, while others, favouring independence and liveliness in children, would find such an outcome highly undesirable.

Evaluation is thus inevitably based on value considerations which are not empirically given and not necessarily universal. The need to interpret events in terms of desirability in part distinguishes evaluation from basic inquiry in the social sciences. The latter can often operate under the more detached perspective of description, prediction, and, where possible, explanation.

Evaluation in lifelong education cannot avoid the explicit examination of its own value base. The broad principles of lifelong education are likely to be interpreted differently in different societies, or even by groups within the same society. This is especially the case at the level of concrete evaluation criteria. The choice of evaluative criteria is in a real sense the concretization of values and as such the most critical aspect of any evaluation.

Evaluation is an experientially grounded activity carried out in a systematic and orderly manner. It was suggested at the beginning of this book that evaluation of any kind, whether casual or formal, always involves a process of "finding out". Some kind of evidence or experience is necessary as a starting point if an evaluative conclusion is to be reached. Concern here will be primarily with evaluation in its more formal aspects. That is, the types of evaluation considered will all depend on systematic examination and investigation rather than on haphazard observation or casual impression. Systematic evaluation more often than not involves the collection and interpretation of a formal, numerical type of data. However, other types of evaluative activity such as the content analysis of a written curriculum may be just as systematic and equally pertinent.

Evaluation is typically a field activity in that the information it utilizes is collected mainly in situations where learning activities occur in their natural settings. Most of the information collected in all types of evaluation evolves from what are often described as "field" or "naturalistic" settings. In contrast, basic educational research often makes extensive use of information collected in deliberately artificial, controlled situations such as the laboratory experiment. When learner performance is the phenomenon being evaluated, achievement is often measured in formal and somewhat artificial situations such as that typified by the standardized achievement test. However, even in the latter case the information generated will, it is hoped, reflect skills and capacities developed by learners in schools or other educational settings.

Almost any educational idea can be made to look successful under ideal conditions. Much can be learned in experimental demonstration schools, and evaluation has its place there as well, but generalization to the world of typical schools requires testing in that world. Evaluation in Scriven's (1967) terms may be "formative" in the sense of being concerned primarily with improving various aspects of educational programs or curricula, or it may be "summative" in the sense of rendering an overall judgment of quality. But both approaches become more and more useful as the context in which they operate approximates the context in which the program or curriculum is ultimately to be used.

Evaluation also tends to be most useful when those involved have direct contact with whatever is being evaluated. Teachers

evaluating learners ordinarily have such contact. However, the evaluation of educational practice is often conducted by specialists rather far removed from the actual source of the data under appraisal. This situation is admittedly hard to avoid when the evaluation is large in scope, involving many institutions. Still, steps can be taken to mitigate the separation.

When evaluators do not have direct contact, significant events may go unnoticed. Much has been written about the importance of identifying "unintended consequences" associated with instruction. It is particularly easy to miss such consequences when goals are spelled out in advance in terms of relatively specific learning objectives and where the information collected in the evaluation reflects only those pre-determined objectives. Often relatively little is learned when this approach is taken. Scriven (1972) has even advanced the concept of "goal-free" evaluation. He suggests that real goals and objectives should be deduced directly from educational activities and materials rather than from formal statements about goals and objectives which, while designed for public consumption, may have little correspondence with reality. In this model the evaluator would remain naive as to the goals of the educational activity while collecting the data. An attempt would then be made to "guess" those goals based on what was observed. The "real" goals and the "guessed" goals could then be compared for congruency.

Evaluation is always undertaken in order to facilitate decision-making or policy formulation. This principle may appear to be something of a truism. It nevertheless deserves constant reiteration because it is all too easy to fall into the practice of collecting information merely for its own sake. Educational research (as distinguished from educational evaluation) is often conducted in order to contribute to general knowledge rather than to any particular decision needs. In contrast, evaluation should always be guided by concern for how the information is ultimately to be used and for what purpose it is to be used. Evaluation involves deliberate expenditure of time and resources that might otherwise have been assigned directly to the teaching and learning process. It must have a strongly utilitarian orientation. This does not mean that evaluations need be so rigidly planned and structured that the unintended outcomes referred to above cannot be detected.

A distinction is made here between decision-making and

policy formulation. These terms refer to levels on a hierarchy of application. Decisions are often specific manifestations of broader policies, and evaluation is relevant to both. In the understandable effort to be highly precise about the nature of evaluation a number of influential writers on the subject have developed conceptualizations which stress immediate, situation-specific information needs, and which tend to ignore evaluation for the purpose of guiding policy formulation. Stufflebeam (1968), Provus (1971) and to some extent Alkin (1969), all writers particularly concerned with the evaluation of curricula or programs, have seen the uniqueness of evaluation as lying in its close tie to educational decision-making of a very immediate kind. There is a corresponding tendency to develop evaluation models for situations in which identifiable "decision-makers" face specific "decision-alternatives".

This point has been carried further in the notion that evaluation should not attempt to generate information that is useful beyond a particular situation. Thus, Worthen and Sanders (1973) suggest that the function of evaluation be narrowed to the point where, "...the object of the search becomes non-generalizable information on performance characteristics of a specific program or process..." (p.23). This statement implies that an activity is truly evaluation if and only if its findings cannot be applied in any other situation. But this would mean that evaluation is irrelevant to the formulation of policy, since by definition the latter does generalize to a variety of situations. Certainly this limitation is unnecessarily restrictive.

Other writers, especially in the field of curriculum, have taken a broader perspective by refraining from any attempt to draw a firm line between evaluation and general educational research. Heath (1969) suggests that one of the three basic functions of curriculum evaluation is generalization, e.g., contributing to a body of knowledge about the design of curricula. Payne (1974) sees the distinction between curriculum evaluation and educational research as merely a matter of emphasis. This is the approach taken here. The more delimited conceptions of evaluation referred to above are nevertheless useful, especially where applied to innovative programmes being installed in schools or other highly structured situations.

Taken together, the four principles just discussed define a broad domain of functioning for educational evaluation. Each is open to great variation in degree of application. It is not

surprising that individuals may disagree as to whether or not an activity is primarily evaluative or not. This does not detract from the usefulness of the principles. It merely reflects the fact that evaluation and general educational research occupy different positions on common continua.

Distinguishing between Different Types of Evaluation

There appear to be five critical characteristics which differentiate between different types of evaluation. Three of these appear to be especially important within the perspective of lifelong education. For present purposes these characteristics will be presented as defining discrete categories, in all but one case dichotomies. Later on it will be shown that all combinations of these characteristics are plausible, yielding the comprehensive classification system for evaluation referred to earlier.

In overview, the basic nature of any evaluation can be described in terms of the following:

1. *Referent*, or whether or not the evaluative judgment is to refer to the learners themselves or to the conditions that accompany learning.
2. *Level*, or aggregate level at which a decision is to be made or policy formulated, beginning with (a) individual learners, and extending through (b) organized groups of learners, (c) educational institutions, and (d) educational systems incorporating a variety of component organizations.
3. *Function*, or whether the strategy of the evaluation is (a) formative or (b) summative.
4. *Agent*, or whether the person or persons responsible for conducting the evaluation occupy (a) internal or participant roles or (b) external or non-participant roles with respect to the educational activities being assessed.
5. *Goal*, or whether the purpose of the evaluation is to assess educational outcomes which can be (a) fully specified in terms of desirable terminal states or (b) are open-ended in the sense of never being completely specifiable in terms of terminal state, possible instances, etc.

Each of these characteristics can now be defined more fully and their implications for the practice of evaluation made explicit.

The Evaluation Referent

Evaluation incorporates two very large classes of activity ordinarily kept separate in theory and practice. The first of these has as its purpose the evaluation of learners themselves, and will be referred to in the discussion that follows as *L-Evaluation*. In this form of evaluation some sort of appraisal is to be made about the desirability of one or more characteristics of the learner, ordinarily for the purpose of facilitating decisions that have to be made about that learner's future activity. Sometimes this is a decision that learners make themselves, and at other times it is made by others, as will be evident when the additional distinction is made between the agents of the evaluation. There are obviously a variety of types of *L-Evaluation* including traditional academic evaluations for the purpose of helping learners who are experiencing difficulties, placement evaluations for the purpose of deciding at what point, in a sequence of content, instruction should begin, and personal evaluations that precede individual counseling.

The second basic class of evaluation referents is made up of the conditions under which learning takes place. *C-Evaluation* assesses the effectiveness or desirability of all conditions, planned or accidental, that are potentially significant aspects of intentional learning activity. These include the learning/teaching process, the instructional materials, the institutional (or non-institutional) context in which the activity is located, and the inter-action of learner characteristics with other conditions. This evaluative referent is often referred to as either "program" or "curriculum" evaluation, though writers using the respective terms typically emphasize somewhat different evaluative criteria. The evaluation of teachers is implicitly included under *C-Evaluation*, although it could be assigned an independent status as a third category if it were useful to the discussion.

In many ways *L-* and *C-Evaluation* are parallel activities, although *L-Evaluation* is most often conducted by participants involved in the learning process, especially teachers and learners themselves. Systematic *C-Evaluation* is usually the domain of people who are external to the learning process, frequently with more of a researcher's, as contrasted to a practitioner's,

orientation. Still, the two types of evaluation often utilize the same kinds of information at a different level of aggregation. But they contribute to different types of decisions, and it is to the relationship between evaluation and decision-making that we should turn next.

Level of Decision-Making and Policy Formulation

Goodlad (1966) has suggested that all educational decisions can be classified within one of three different levels, each corresponding to one of the basic types of educational resources that a society may make available to its citizens. In a more recent treatment of this classification, Goodlad (1975) describes the levels as "*instructional* (tutors, teachers, books, responsive machines, for example), *institutional* (institutions whose primary function is designated as education), and *societal* (aspects of living having potential for serving an educational function)", p.13.

Perhaps the most important implication of Goodlad's analysis is its recognition that direct means of delivering education are by no means confined to interactions among teachers and learners in the classroom. An institution such as a school can have an impact that is more than the sum of what is occurring in its separate classrooms. Institutions as a whole have "climates" that presumably can affect learning. Teaching and learning can even be organized on an institution wide, rather than a classroom basis. Likewise, large systems that transcend or incorporate individual institutions such as schools, or that are quite separate from formal schooling, may have societal outcomes that are traceable to the system as a whole.

This idea of resources and decisions at different levels has been elaborated by Suchodolski (1970) in his concept of a "third pedagogy". On the one hand this third pedagogy is identified with the "...contents and structure of the whole school system in which the different schools function", a system which, for example, may "allow an easy flow of youth to the highest levels or which may set obstacles in the way" (2). Likewise, "...the education of youth is not the exclusive business of schools but is the result of activities on the part of many other formal and non-formal entities". Both Goodlad and Suchodolski thus suggest that direct educational impacts can be engineered at the societal level through school systems as well as through other systems that are independent of schooling. This same emphasis is implied in a variety of ways in

writings on lifelong education. Any comprehensive treatment of evaluation must take into account the level at which decisions and policies are to apply.

In a paper concentrating on the topic of evaluation it seems useful to make two modifications in the conceptions provided by Goodlad and Suchodolski. First, an additional distinction can be made at the instructional level where evaluation can be concerned either with the individual learner or with learners assembled in interactive groups. For example, a teacher could select a set of instructional materials for a single learner or for a group of learners. Learners could be evaluated as individuals or as co-operative learning groups where all individuals receive a mark for the group, and so on. This additional distinction has been made at the instructional level because the lifelong education literature emphasises both individual, independent learning as well as co-operative or inter-learning.

In addition it appears to be better for present purposes to use the term "system level" instead of the term "societal level" or "third curriculum". This is solely by way of being more specific about the referent of the evaluation. At this third level evaluation addresses the information needs of large systems incorporating more than one educational institution in addition to systems which, while perhaps originating in a single institutional base (television network, publisher, government agency, etc.), are widely diffused throughout a society. Evaluative conclusions are then more readily seen as pertinent to a specific system operating within a society rather than as referring to a society-wide "third" curriculum. The systems referred to here, by the way, may or may not extend through the entire society. Thus, local school systems under a decentralized educational authority would be found at the system level.

The distinction between L- and C-Evaluation and the four levels of decision-making taken together define a preliminary classification system incorporating the basic classes of decisions facilitated by evaluation. Table 1 provides examples for each of the resulting categories to help clarify what has been presented up to now.

The examples listed in the two columns of the table corresponding to L- and C-Evaluation should clarify the fact that the former always involves a conclusion about learners and the latter a conclusion about the conditions that mediate learning.

TABLE 1

A PRELIMINARY CLASSIFICATION SYSTEM COMBINING
EVALUATION REFERENT AND LEVEL OF DECISION OR POLICY ADDRESSED

Level of Decision or Policy	Referent	
	L-Evaluation	C-Evaluation
Individual	Independent learner conducts self-evaluation of own progress in achieving a goal	Search by an individual learner for most satisfactory way of accomplishing some learning goal
	Assessing present reading level in determining where to start a learner in e.g. a sequenced reading program	Comparing effectiveness of various individualized approaches to teaching e.g. reading for learners with certain characteristics
Group	Assigning learners to learning groups so as to facilitate certain patterns of co-operation among members	Comparing alternative patterns of grouping the members of a given class to promote learner satisfaction and progress
	Monitoring actual progress of a learning group against rate of progress estimated to be appropriate	Comparing different approaches to group study for learners with certain defined characteristics
Institution	Assessing competencies and skills of learner pool in order to establish priorities for a school curriculum	Comparing different curriculum materials for appropriateness at schools enrolling certain types of learners
	Identifying and recruiting learners with defined characteristics for a local work-study program	Comparing different patterns of combining practical with academic learning in work-study programs
System	Identifying sub-populations needing re-training because of changes in certain work-roles	Comparing various formal and informal modes of education in terms of utilization by a given sub-population
	National survey of independent study skills of learners at the completion of secondary school	Assessing a written national curriculum against criteria derived from the principles of lifelong education

There is a deliberate parallelism between juxtaposed examples in the two columns, and as a result it is probably most useful to read the two examples in each horizontal pair successively before going on to the next pair. Thus, the first pair of example has, under L-Evaluation, an individual learner assessing his or her own progress, presumably toward some personally defined goal. The corresponding example under C-Evaluation has an individual learner choosing between alternative methods of achieving a goal, implying that he or she is evaluating the conditions under which learning is to take place.

In the C-Evaluation example listed in the Table it should be clear that the decisions or policies at which the evaluation is directed are to be implemented at the level in question, but they do not necessarily apply only to the given individual, group, institution or even system in which the evaluation was conducted. In other words, the information resulting from the evaluation could be situation-specific or it could be generalizable across a variety of similar situations. The first example under C-Evaluation, group level, is specific to a particular classroom. It implies a kind of internal evaluation that a dedicated teacher might conduct. The second C-Evaluation example in the same level is intended to produce generalizations applicable to any classroom enrolling learners from the population studied in the evaluation. This merely reflects the earlier conclusion that evaluation can and does produce generalizable, in addition to situation-specific, information.

It might also be pointed out that portions of the national evaluations conducted as a part of the multi-national curriculum evaluation study reported by Skager and Dave (1977) and described in the previous chapter, fit under the last example in the Table (system level C-Evaluation). For countries involved in the study, all with centralized educational authorities, the written curriculum is of course an example of an influence that operates at the system level.

Table 1 presents what amounts to a slice or section of the full treatment of evaluation being developed here. However, the classification system in Table 1 does stand on its own in defining the basic categories of decisions addressed by evaluation. The remaining characteristics distinguish between types of educational goals, evaluation processes, and evaluation agents in ways that appear to be pertinent to lifelong education.

Formative and Summative Evaluation Functions

The distinction between formative and summative evaluation made originally in the paper by Scriven (1967), has been elaborated extensively by other writers. Probably the best argument for importance of formative evaluation is still the earlier paper by Cronbach (1963), to which Scriven's later paper was in large part a response. Cronbach, writing on the role of evaluation in "course improvement" (in present terms C-Evaluation at the instructional level), maintained that, "The greatest service evaluation can perform is to identify aspects of the course where revision is desirable" (p.676). The idea of generating information to be used for revising or improving educational practice is at the core of the concept of formative evaluation. Cronbach saw little utility in summative evaluations comparing one course against another, such as a new course or curriculum against prevailing practice. This nevertheless remains a typical approach to summative evaluation, implying some sort of final, over-all comparison of one alternative against another. Cronbach's position was that differences between courses or curricula observed in summative evaluation might easily be eliminated or even reversed if formative evaluation were used to improve one or other of the alternatives.

The concept of formative evaluation has broadened since Cronbach's paper was written. Bloom, Hastings and Madaus (1971) extended it to L-Evaluation at the individual level in their analysis of evaluation in the process of teaching and learning. Here formative evaluation concentrates on providing feedback to learners and teachers for the enhancement of learning.

Formative evaluation appears to be a much more positive and constructive function than does summative evaluation. It seeks to improve or facilitate educational practice. Summative evaluation is much more concerned with judging and selecting in both L- and C-Evaluation. Nevertheless, summative evaluation is a function that cannot be dispensed with. It is true that in the past the summative evaluation of learners by traditional academic marking systems and for the differential distribution of educational opportunity has had many negative aspects. The literature on lifelong education contains frequent references to the socially restrictive aspects of traditional grading and promotion systems (e.g. Parkyn, 1973). Some of the central proposals of lifelong education were formulated as alternatives to traditional practices in L-Evaluation. Still, many of the examples of types of evaluation in Table 1 are sum-

mative in nature. Elimination of the more negative aspects of summative evaluation has to do in any case with the restructuring of education itself.

It appears, then, that both evaluation functions are relevant to evaluation under lifelong education. Presumably there would be a greater emphasis on the constructive aspects of formative evaluation than there is at present, especially in the treatment of learners.

The next distinction, that between internal and external agents of evaluation, is important because of the probably greater emphasis under lifelong education on the role of internal evaluation agents. Internal evaluation by participants in an educational process would tend to be formative in nature, and this also argues for greater utilization of the formative function.

External and Internal Evaluation Agents

Much of the earlier writing on formative and summative evaluation has recognized at least implicitly that those who conduct evaluations can have one of two kinds of relationship with whatever is being evaluated. On the one hand the evaluator can at the same time participate in the educational process as a self-directed learner, as a teacher attempting to apply or use some method or procedure, or as a developer seeking to improve learning materials or the context in which learning occurs. Part of the notion of participation is contained in the idea that the individual in question has a role in the situation that includes responsibilities in addition to evaluation. Participation also implies some sort of communality of concerns and activities between evaluators and other participants. It may even imply that all participants in an educational enterprise share some of the responsibility for evaluation. One example of this will be presented in a discussion of the advocacy model in the final Chapter. Rather than assessing the efforts of others, *internal evaluation agents* assess the desirability of events or conditions for which they are themselves at least partly responsible.

On the other hand, evaluators are *external evaluation agents* whenever their responsibilities in a given situation incorporate only evaluation. Even the teacher can be an external agent when evaluating learners in L-Evaluation, particularly when the learner is the passive object of the evaluation.

In C-Evaluation, typical evaluation practice has usually taken it for granted that an evaluator is properly an external-agent with a different and in some respects higher status than the participants in the activity under evaluation. Administrators evaluate teachers. Outside experts, from government or universities, evaluate curricula, programs, and other conditions supposed to facilitate learning. External evaluation agents are often preferred because their lack of personal responsibility for the events or conditions under evaluation presumably enhances their independence and objectivity.

It is understandable that formative evaluation tends to be associated with internal agents and summative evaluation with external agents. Formative evaluation for the purpose of improving an educational activity requires close contact with a situation and participation is one way of achieving such contact. Summative evaluation is more judgmental and often terminal. Here the objectivity of the external agent has appeal. However, there is no one-to-one relationship between the nature of the evaluation agent and the evaluation function. Evaluation agents who are external and who are responsible only for evaluative functions can and do conduct formative evaluations. Internal agents who have responsibilities in an educational process also may be responsible for summative evaluation.

It is clear that lifelong education would place much greater stress on the utilization of internal evaluation agents than has been the case in the past. The principle of *self-learning* directly implies both willingness and competency on the part of learners in regulating their own activities through self-evaluation. This is recognized by Dave (1973) who also stresses the enhanced role of the formative evaluation function, especially in L-Evaluation:

"The chief purpose of evaluation should be to improve achievement rather than just measure it for the purposes of classifying students or issuing certificates to them. Evaluation of educational achievement should be improvement oriented". (p.41)

The principle of *democratization* likewise suggests greater emphasis on internal evaluation agents under lifelong education. This principle emphasizes participation and the reduction or elimination of many types of status barriers among participants in the educational process. The concept of the internal evaluation agent is thus not restricted to self-evaluation by learn-

ers or groups of learners working co-operatively. All participants can act as internal agents, and this includes instances of C-Evaluation in which participants in an educational process evaluate their own attempts to create and maintain optimal conditions for learning.

There is no intent to suggest here that internal evaluation is necessarily more valid than external evaluation, or that the latter does not have an important role in lifelong education. Both types of evaluation could be invalid by being based on inaccurate information. Both could also be referenced to poorly thought out or irrelevant standards.

Internal evaluation on the part of the learner does not exclude comparison of one's work with that of others. What others are able to do is a valuable source of information about one's own progress, assuming that the information is not used in a self-destructive way. But the learner in this case is still the evaluation agent. Also, internalization of the evaluative information (as evidenced in its application or use by the learner) can occur whether the agent is internal or external. Many would doubtless argue that internalization is much more likely if the learner or participants are active as evaluation agents. But seeking out the evaluations of others, getting them to serve as external evaluators, does apply an initiating and, hence, partly autonomous role.

Specified and Open Educational Goals

The discussion of evaluation criteria in the previous chapter made it clear that lifelong education incorporates structural and organizational goals as well as a variety of goals relating to the personal development of learners. It was suggested that under the lifelong education framework the former are not merely valued as "means" toward certain types of outcomes in learners, but are in fact assigned independent criterion status. Just as it is desirable to foster certain types of personal growth in human beings, it is also desirable to facilitate certain types of development in institutions and structures associated with education. This relates very closely to the notion just discussed of possible independent (e.g., over and beyond events in the classroom) impacts of institutions and larger systems on learning and learners.

OPEN-ENDED GOALS: The illustrative criteria and specifications presented in the first chapter are virtually all "open-

ended". That is, they indicate desirable directions of continuous development rather than closed, terminal states about which it can be said with finality that a given goal has been fully achieved. Even something so straightforward, for example, as integration between the educational functions of home and school is not really definable in a meaningful sense in terms of an ideal state of home-school integration. Hundreds of instances of home-school integration could doubtless be proposed. But even such a long list would probably not be exhaustive of all the possibilities and might well include instances that would be unacceptable in some societies and desirable in other societies. In other words, a state or condition of home-school integration is an ideal that can be approached but never fully achieved, just as the limit of a mathematical function can be approached but never reached by the curve representing that function. It is possible to propose and recognize instances (and non-instances) of home-school integration. It is even possible to demonstrate that one institution or system shows more instances of such integration than does another. But it is not possible to state with confidence in any concrete way what a complete or ideal state of integration between the home and the school would be like. Anyone who was foolish enough to attempt to do so, would doubtless threaten cultural, political, or economic values held by large segments of the earth's population.

Integration between the home and the school is a structural principle subsumed under the *Horizontal Integration* cluster, but similar examples could be given for criteria from the other clusters. Select any criterion under *Democratization*, for example, and see if it is possible to define concretely an ideal state in which all possible instances are included and all non-instances clearly excluded. There are many ways in which any of these principles can be implemented, but full implementation remains an elusive abstraction.

The same is true for goals relating to learners. The various criteria of *Self-Growth* and *Self-Directed* learning mainly suggest ideal states that provide direction to efforts rather than inform that a goal has been reached. The importance of open-ended educational goals was stressed by Cronbach (1971) in his distinction between three stages of education: (a) the establishment of basic (though static) knowledge and skills through training, (b) the development of intelligent analysis and problem-solving and (c) the fostering of creative and self-expressive production. The explicit goals of lifelong education

tend to fall into the latter two categories, although the first or basic skills category is certainly implied as well, since it is essential to the development of the other two. The latter two stages were also characterized by Cronbach's as "open-ended". Analysis and problem-solving usually allow for multiple processes of solution. Creative and self-expressive acts are not even predictable. They can be recognized when they occur (though not necessarily immediately), but they cannot be encapsulated within a definition that anticipates all of the possibilities.

There appear to be at least four possible characteristics of open-ended educational goals. Each taken alone would define an open-ended goal, but they are probably more likely to appear in various combinations, with the first characteristic probably always present. Open-ended educational goals:

1. are *continuous* in the sense of not having a concretely definable or predictable terminal point of attainment;
2. may incorporate an infinite set of manifestations by being *transferable* or applicable to a wide variety of situations and contents;
3. may represent *meta-* or *second-order* abilities that interactively combine other, less complex skills and competencies; and
4. may be represented in characteristics which combine in complex ways cognitive and affective facets of the personality leading to *self-transcendent* behaviour manifested in commitments to ideals.

The four characteristics are stated in terms of the development of the personal characteristics in learners. It should be recalled that organizational and structural goals can also be open-ended, especially in the sense of (1) above.

Open-ended goals can be measured. Chapter 4 will in part deal with the construction of definitions that lead to the assessment of instances of such goals. Appropriate definitions make it possible to recognize instances of open-ended goals when they occur and hence to count or otherwise assess their degree of attainment. But such definitions cannot specify a terminal state in which the goal has been entirely achieved nor do they predict all of the possible instances.

SPECIFIED GOALS: The first stage of Cronbach's (1971) continuum incorporates "closed" goals for which a desired state of accomplishment can be clearly stated. When applied to education the notion of something being "closed" is of course pejorative, and it seems more accurate to describe such goals as *specified*, simply because susceptibility to precise delineation is their most important characteristic. The process of specification involves making a relatively abstract goal statement more concrete by developing one or more precise rules as to how manifestations of the goal are to be recognized. Such rule-statements may refer to the actual performance of learners or to characteristics of an educational process or institution.

Specified goals by no means need be trivial. The kind of specified goals we are interested in form the basic building blocks from which higher order learning must evolve. They comprise much of the early curriculum and often encompass knowledge domains that are very large in terms of the number of elements contained. But at the same time it is true that school curricula can emphasize the achievements of specified goals that literally lead nowhere (e.g., memorizing material that does not contribute to further learning) or persevere at this level by failing to challenge learners to develop higher order skills. Old-fashioned, rigid curricula emphasizing memorization and rote learning were of this kind. The danger is not in specified goals themselves, but rather that the curriculum might be aimed solely at this level.

The next chapter will compare several models for defining specified educational goals and also describe the various functions of L-Evaluation in the achievement of such goals. However, at least two examples of specified goals can be given here in the form of the familiar behavioral objective.

After reading an editorial taken from a nationally distributed newspaper in his or her first language, the learner will be able to select those points of view presented in the editorial from a longer list of points of view about the same topic.

This first example involves a relatively low level reading comprehension skill. It asks only for recognition rather than for production (e.g., where learners themselves generate the list without the stimulus of a prepared list). The task as defined does not require inference (assuming the editorial was clearly written) or analysis, let alone synthesis and evalua-

tion. Still, being able to recognize the basic points that a writer is making in connected discourse is an essential step in the development of the ability to read with understanding. The performance is specified, but non-trivial.

Specified goals may also refer to performance at higher levels of cognitive or affective functioning.

After reading a discussion in which the author makes a prediction about some future event or state, the learner will identify (by listing) the bases for the conclusion, distinguishing between factual evidence, argument and assumption.






The second objective requires the analysis of written material as well as the ability to distinguish between evidence, argument and assumption. It appears to incorporate at least two types of higher level functioning as well as the ability to produce, rather than merely recognize, a correct response. The objective could be elaborated with specifications as to difficulty of the reading passage, subject matter, and so on, but it does serve to illustrate that specifiable goals can refer to relatively high level performance.

A potential source of confusion between educational objectives derived from open-ended goals and those reflecting specified goals lies in the fact that it is usually possible to derive a specified objective from an open-ended goal. For example, the *Unusual Uses Test* which is purported to measure certain aspects of creativity requires that respondents list as many different uses for various common objects (e.g., a brick) as they can think of. The test is commonly scored for the number of meaningful responses over and beyond the common use of bricks for construction. This domain of responses is genuinely open-ended in that unique, unpredicted responses may always be given by individuals. There are no limits defining the maximum possible score. However, a specified objective can be derived for this open-ended task simply by specifying an arbitrary standard of performance, such as the following: "The learner will give at least five acceptable responses to the *Unusual Uses Test* administered under standard conditions". The imposition of an arbitrary limit when the test is scored for total number of correct responses establishes a specified domain of performance with absolute limits within a much larger open-ended domain.

The distinction between open- and specifiable goals is especially important for evaluation under lifelong education. It will be shown in the two chapters that follow that the two types of goals lend themselves to quite different strategies of evaluation, especially as related to the crediting or certifying of learner achievement.

A Classification System for Evaluation

Considered together, the five characteristics discussed in the previous section comprise a system for classifying different types of evaluation. This system appears to be applicable to evaluation operating under virtually any form of education. It is useful here as a means of identifying particular types of evaluation that would be especially important or common under lifelong education.

<u>Referent</u>	<u>Level</u>	<u>Function</u>	<u>Agent</u>	<u>Goal</u>
				
L-Learner	1-Individual	1-Formative	1-Internal	1-Open
C-Conditions	2-Group	2-Summative	2-External	2-Specified
	3-Institution			
	4-System			

An Evaluation Index for Differentiating Between
Types of Evaluation Based on Five Characteristics

FIGURE 1

The five boxes in the figure can be thought of as registers which might display an index combining any of the sub-categories of each of the five characteristics. In all there are 64 possible combinations, far too many to illustrate here. All, however, appear to be plausible, and a few illustrations are certainly in order.

L-1111: This combination refers to a very common (and highly relevant) kind of evaluation. It has already been referred to as self-evaluation, in this case as relating to the learner's own development along some open-ended goal continuum. An individual using self-selected criteria to assess the quality of his or her own developing capacities in design, esthetic production, etc. would be engaging in this type of evaluation.

L-1121: This category of evaluation is perhaps equally common. It is illustrated whenever a teacher evaluates an individual learner's performance in some open goal area for the purpose of helping the learner improve.

L-1122: This particular type of evaluation would occur in classrooms in which teachers systematically collected information on the progress of individual learners toward the mastery of the kind of specified goals that comprise much of basic school mathematics, vocabulary, and reading skills. If it is a group that is being so monitored the closely related category, *L-2122* is seen. These important categories incorporate the kind of formative evaluation for the purpose of managing the instructional process that is the central element of several contemporary approaches to the organization of teaching and learning. Both categories will be elaborated extensively in the chapter which follows. They are by no means relevant only to the classroom. Evaluation by supervisors during on-the-job training would often fit this pattern. So do informal evaluations made by parents who are guiding the home-learning activities of their children.

L-3111: This interesting category would find an institution examining its own learners, perhaps for the purpose of assigning priorities in the curriculum. Since the evaluation agent is internal the learners would be participating in the process in some active way by helping to define goals, reporting on their own perceived strengths and weaknesses, and the like. A seemingly slight modification in this category can produce a very different evaluation situation. Thus, *L-3221* would incorporate evaluation of learners by an institution for

various selection purposes. The common use of aptitude and general achievement tests for admission, grading, and the provision of special opportunities are typical examples.

L-4122: This systems level category would be illustrated by a national (or other large scale) survey conducted to identify individuals who might benefit by various kinds of educational opportunity. It is formative evaluation in that the object presumably would be to identify potential learners and guide them toward the development of personally and socially useful skills and competencies.

All of the examples so far have been confined to L-Evaluation. Parallel examples could easily be given for C-Evaluation, the difference being merely that here it is the conditions of learning rather than learners themselves that are being evaluated. For instance, the first example given (L-1111) when transformed into C-1111 would be illustrated by an individual learner deciding what is the best approach to the further development of some open-ended domain of competency. This would presumably involve systematically collecting and weighing evidence as to the desirability of various alternatives. If a teacher rather than the learner were making the decision the category would change to C-1121. Rather than make the remaining transformations of the L-Evaluation examples already given, it would be more interesting to identify two or three of the most common varieties of C-Evaluation.

C-2112: Perhaps the majority of C-Evaluations involve comparing various educational methods, materials, etc. in terms of how they differentially affect the performance and attitudes of groups of learners. Much of the recent conceptual as well as technical writing on evaluation referred to earlier pertains to this single category and to its typical counterpart companion for summative evaluation, C-2222. In the first case we have an internal formative evaluation, and in the second an external summative evaluation of a program, curriculum, mode of organization and the like. This second category is probably viewed as the most relevant evaluation mode by virtually all national and international agencies funding developmental projects in education. It is also favored whenever special funding for developmental purposes occurs within educational institutions of systems. This type of evaluation is thus used particularly in the context of decisions about funding and the formulation of policy. It is dependent on external evaluation agents and is considered advantageous because of the widely perceived need for

independence and objectivity in these kinds of situations.

C-4221: This category is illustrated by any assessment of educational outcomes that is designed to reflect on the quality of a system as a whole. It could be concerned with either open or specified goals, so category C-4222 is equally relevant here. It might involve the comparison of alternative systems, as in the case of evaluations that had been conducted in a number of countries for the purpose of comparing comprehensive versus traditional forms of education, or it might be confined to a single, total educational system. The National Assessment of Educational Progress conducted annually in the United States in recent years has focused primarily on goals of the specified variety in what Wilson (1974) describes as an attempt to provide "... direct measure of educational outcomes which could be utilized by school systems to improve the educational process" (p.27). These two categories of evaluation obviously have their major role at the level of policy formulation.

At first thought it might seem unlikely that internal evaluation agents could be employed at anything so broad as the systems level and, as a result, that systems level categories incorporating internal agents would be empty. This turns out not to be the case, however. For example, a consortium of training institutions at the professional level which co-operatively assess the competencies of their own graduates in the interest of redesigning curricula would be engaging in an internal, C-Evaluation at the systems level.

The classification system just explored does appear to highlight critical features that distinguish between the various types of evaluation. This has already helped make it easier to point out some of the considerations that would shape the practice of evaluation under lifelong education. The final section of this chapter deals directly with this question.

Implications for Evaluation under Lifelong Education

The principles of lifelong education would have a significant impact on the practice of evaluation in ways that correspond to each of the distinctions made in the classification system. In the first case this impact would be confined to a single category of evaluation. In the other instances several categories are involved.

with lifelong education. Greater provision for flexible crediting and certifying systems is clearly associated with implementation of the principle of vertical articulation, as is apparent in Lebouteux's (1973) report on the deliberations of the Pont-à-Mousson Conference:

"If...this new form of assessment is accompanied by the award of "units" or "credits" that are valid beyond the school stage then the gap between school education and occupational life will have been truly bridged...The "credits" system is giving birth to a new idea, that of recurrent education...the idea of a programme of education that begins at school but is carried on throughout one's life". (p.12)

The important notion in the above is that credits obtained through schooling would not simply be accumulated toward a final degree or certificate, but would be valid after the termination of schooling. Assessment for crediting and certifying would as a result occur more frequently in the lives of individuals than is the case at present. The possibility of having various kinds of learning recognized in a formal way would presumably act as a motivating factor for lifelong learning.

The next and following chapters will argue that genuinely open-ended educational goals are not compatible with the concept of crediting and certifying adopted here. Rather, these functions will be identified with the accomplishment of specified goals and objectives. It will be suggested that a great deal of what is important in education from the point of view of experience and personal development would not be susceptible to the fairly rigorous approaches to assessment for crediting and certifying that will be presented in the next chapter. However, this by no means implies that crediting and certifying apply only to unimportant or low-level accomplishments. It has already been argued that specified goals are far from trivial in education and work. There are two very significant reasons - one pertaining to the society as a whole and the other to the individual learner - why accurate, fair, and flexible crediting and certification systems would be essential under a fully implemented system of lifelong education.

From the point of view of the society accurate crediting and certification offer a vital protective shield against incompetency and charlatanism. Lifelong education, by providing a variety of non-formal means for educational attainment, would

at the same time tend to reduce the role traditional schools and universities have played in the maintenance of educational standards. Under these circumstances only a genuinely utopian society could afford to ignore the need to develop an alternative means for verifying educational attainment.

Certification and crediting also serve a highly important function from the point of view of the learner. These evaluation processes can guarantee that learners are judged on what they can do rather than on the time, place, or manner in which their skills were learned. Effective certification directs attention to specific competencies rather than to general school credentials. LaBelle and Verhine's (1975) summary of the large number of studies on non-formal education and occupational stratification in Latin America attests to the importance of this issue. Their findings clearly reveal that non-formal education does not enable individuals to move out of low occupational strata as long as formal school credentials constitute the primary basis for advancement. The alternative appears to be a universal, competency-based system of crediting and certification.

2. *Lifelong education would place greater stress on evaluation by internal agents for both L- and C-Evaluation.* The emphasis on democratic participation, autonomy and independent learning embedded in the principles of *Self-Learning* and *Democratization* suggests that for both L- and C-Evaluation there would be a much larger role for all categories in the classification system which incorporate internal evaluation agents. For L-Evaluation the development of motivation and skill in individual and collective self-evaluation becomes a central goal of schooling, an indispensable part of the broader educability concept. The views of Schwartz (1970) are quite representative in this regard.

"Evaluation of the results would no longer be the prerogative of the teacher but would have to be undertaken jointly with the pupil. Everything must be done to replace evaluation by self-evaluation in the sense of 'enabling the pupil to assess himself continuously and to analyse his own mistakes'". (p.64)

Evaluation, always an important element in the so-called "hidden curriculum" of schools, would join the explicit curriculum. Learners would also be encouraged to use multiple means of evaluation.

When evaluation is internal, desirability or worth is reflected ultimately in the level of satisfaction individuals feel about the activities in which they engage. This would be true for internal L-Evaluation where learners are evaluating their own activities as well as for internal C-Evaluation where participants evaluate the conditions they have created and maintained for promoting learning. Satisfaction should not be equated with mere happiness or comfort. It has to do with accomplishment, with what Fenstermacher (1975) has described as "...enduring gratification yielding a sense of worth and pride in accomplishment..." (p.238), with a sense of self-fulfilment.

3. *Under lifelong education the formative functions of evaluation would receive significantly greater emphasis.* This point has already been made in the earlier suggestion that categories involving both internal agents and formative functions appear to be highly consistent with many of the comments about evaluation appearing in the literature on lifelong education. As far as possible evaluators are to play facilitating rather than judgmental roles. Traditional approaches to evaluation, especially L-Evaluation, are repeatedly characterized in terms such as those used by Parkyn (1973) as, "...based largely on the assumption that the school should be a selection and rejection mechanism for channeling people to different levels of a predetermined vocational and social hierarchy" (p.31).

Traditional grading and examination systems are frequently associated with the less egalitarian aspects of formal schooling. This view is closely connected with opposition to grading or marking systems based on comparisons between the performance of an individual and that of some reference group. Thus, a 1971 study of the Council of Europe on permanent education concluded that, "...the value placed on a pupil's performance should not be based solely on the average performance of a group of students competing against each other" (p.22). It should be noted that the emphasis on evaluating for crediting and certifying in (1) above is a constructive alternative to the kinds of comparative or relative grading practices that have been so heavily criticized in the literature on lifelong education.

4. *The usefulness of evaluation under lifelong education would be significantly dependent on the development of a greatly enhanced capability for the measurement of goals of the open-ended variety, especially those associated with the general concept of "educability" in the case of L-Evaluation and with in-*

tegrative and interactive factors in the case of C-Evaluation
The overall notion of educability sums up a variety of goals of lifelong education with respect to the development of the individual. These goals are derived mainly from the *Growth* and *Self-Learning* principles and represent a genuinely open-ended complex of personal characteristics.

The nature of educability needs considerable elaboration at a conceptual level before the practical measurement of its components can be undertaken. Dave (1973) has made a start in this direction by describing five categories of relevant learner characteristics. These will serve as a starting point for the discussion of the educability concept in the fourth chapter.

It should be evident that the measurement of many criteria that might be associated with educability would require careful conceptualization followed by adequate research and development. There are real dangers inherent in the careless application of educability criteria for L-Evaluation. Certainly when effective strategies do not exist for fostering aspects of growth relating to educability, learners can hardly be held accountable for failing to develop in desired directions.

Enhancement of educability is the fundamental goal of schooling. Educability really refers to any skill or capacity which contributes to the potential for later learning. It is easy enough to criticize the concept on the grounds that it would be difficult to prove that almost any item of knowledge, even seemingly trivial factual knowledge, could not in some way contribute to later learning. This misses the point. There are a number of areas in which commitment to the educability concept on the part of schools would be recognizable in the light of the relative emphasis given to various kinds of activities on the part of learners.

It is admittedly true that much of what would be included in an "educability curriculum" goes on already in most schools to at least some degree. Educators in "traditional" schools virtually anywhere in the world would acknowledge the importance of developing such characteristics in learners. However, it would be difficult to find examples of evaluation instruments and systems which are directed at more than a few educability criteria in addition to the most concrete (even unimportant) areas of subject-matter content. Evaluation always reflects what is held to be important in a given place and time.

For C-Evaluation a variety of open-ended goals are implied in the Horizontal and Vertical Integration principles as well as in Democratization. These principles are structural, organizational, and sometimes interpersonal. Anything that has been said in the case of educability about the need for careful conceptualization and development of modes of assessment applies here as well.

NOTES

1. The four dimensions described here have evolved from a slightly earlier set of three presented in Chapter 3 of the report by Skager and Dave (1977) on a multinational curriculum evaluation study conducted within the life-long education perspective.
2. Translations of relevant excerpts from Suchodolski were kindly provided by Mr. P  ter Inkei (see acknowledgments section).

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 8



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOME AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

Outline of The global evaluation of Literacy
within the OACV (Ground-nut and Crops
operation), Kita, Mali 1978

Propositions of the evaluation scheme of the results
of the functional literacy within the OACV

I. General Points

1. A first series of research undertaken by the group in charge of the study about basic education (documentary research, visits of researchers, students' work) has made it very evident that the best results on literacy were obtained within the ground-nut operation and crops. A profound evaluation of the obtained results within the framework of this operation will thus have an impact, not only as far as the pursuit and improvement of the work within the OACV is concerned, but also for the whole of the literacy projects in existence and for those which are to be created within other development operations.
2. The two literacy zones within OACV have already begun their own internal evaluation which is conveyed in the collecting of an important number of indispensable data at every attempt at external evaluation. One of the objectives of the work confided to the group in charge of the study of basic education will be to strengthen the devices of self evaluation which exist already in order to increase their efficiency. Thus, there should not be any opposition between self evaluation and external evaluation, but rather complementarity.
3. The present evaluation scheme bears largely in mind the propositions made by the Bank and transmitted to the Government by letter dated the 14 April, 1975.

II. The Proposed Evaluation Scheme

A. SUPPORT TO THE LITERACY SERVICE

In conformity to the announced principle under point 2 above, the group charged with the study of basic education will begin by giving its support to the people responsible for the two zones of KITA and KOULIKORO to bring them close to the evaluation work. Their understanding of the objectives which fixes external evaluation, of the methods which will be used by the external evaluators, their participation on the level of collecting the basic data is in fact a 'sine qua non' condition of all serious evaluation.

It has also been proposed to begin the work by organizing the two training sessions collecting the heads of the zone and their assistants, the heads of ZAF as well as the heads of the sector and the sub-sector of OACV. During that session the objectives and the methods of the evaluation for the service and for the operation will be shown. Their participation will be equally specified. This is very important on the level of a permanent collecting of data which the evaluators will strive to interpret but which is impossible for them to record from day to day.

It is also indispensable

- (a) to measure the starting point of the listeners and to organize them in groups of homogenic levels;
- (b) to obtain a regular holding of registers of presence to allow evaluation of
 - the degree of regularity of courses,
 - the recorded losses, and

- the total hours of presence of each listener allowing to obtain precise information on the different attainment rhythms.

(c) to obtain what is organized at a regular periodicity of the tests of a intermediary level allowing to appreciate the rhythms of progress according to the different apprenticeships (reading, writing, calculations).

B PUTTING INTO WORK THE "EXTERNAL" EVALUATION

1. It has been proposed that the responsible group is composed of:

- two researchers of the group of the study of basic education: one sociologist and one linguist;
- of the evaluation cell of the literacy service;
- of a specialist consultant of the functional literacy programmes acting as a mediator for 3 months at the beginning of the phase and for 3 months at the time of exploitation;
- the survey coordinator about basic education will assure the general coordination of the evaluation study.

Moreover, a call will be made for the support of the cell of the technical and socio-economic evaluation of the OACV to define the most adequate criterions in view of measuring the impact of literacy in these two domains.

2. Objectives of the evaluation

2.1 Evaluation of the appropriate results of literacy (internal evaluation)

2.1.1 Quantitative data

It will be a question of measuring the efficiency of the programme in terms of acquisition of new knowledge. Three parameters will be taken into consideration:

- (1) The different types of apprenticeship: reading, writing, calculation. (N.B. The evaluation which was done at Koulikoro brought to the fore a very distinct weakness in calculation whereas the first surveys on motivation had shown on the contrary a large interest of the farmers for calculations. There will be reason to verify whether the same phenomenon appears uniformly at all centres and whether one finds it by the same degree at Kita. There will also be cause to elucidate the reasons for: the insufficiency of instructors, the inadequacy of the method, etc.)
- (2) The hours of presence: the measuring of the hours of presence is indispensable in order to compare the acquisition rhythms of the different listeners amongst themselves and the real rhythms of acquisition in comparison to theoretic rhythms anticipated in the adopted progression.

Only the regular session of registers of presence will allow for the establishing of significant correlations.

- (3) The age of the listeners: several converging indications permit one to think that there could be a very distinct correlation between the age of the participants and the obtained results, the best results were obtained with the 15-25 age group. The evaluation should allow to encompass this point with precision. If this has been verified it is in fact evident that such a situation will not be without important practical consequences. (N.B. This quantitative evaluation endeavours to be exhausting, that is to say, to be directed at all centres and all listeners.)

2.1.2 Qualitative data

Further to the quantitative data mentioned above, the systematic observations of the literacy sessions will allow to evaluate the level of the instructors' competence and the quality of the practised education. It has also been planned to edit a certain number of monographs on the centres which have obtained the best results. Often a monograph on a successful experiment can tell more in effect than a series of statistic tables.

2.2 Evaluation of costs

The evaluation seeks to establish a scale of real costs in terms of:

- achievements at the different centres; and
- achievements at the different literacy zones

in order to encompass the optimum conditions of internal profitability of the literacy investment.

2.3 Evaluation of the literacy impact (external evaluation)

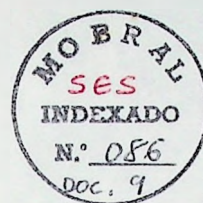
This evaluation is done by a systematic survey in comparison with neo-literates. This survey should permit to specify:

- the "sociology" of neo-literates: their age, their economic and social status, the characteristics of their exploitations, their greater or lesser openness to innovations. The literate farmers can thus be placed according to their environment;
- their motivations for leaving;
- what literacy has changed for them;
- their expectations and aspirations as new literates (efforts will be made to collect information concerning their appreciation of the content of literacy: what appears useful to them, what they would have liked to learn and what did not figure in the "programme", etc.)

This survey will be conducted:

Unesco Institute for Education
Feldbrunnenstrasse 58
2000 Hamburg 13
Federal Republic of Germany

PRG 4.32
October 1982
Document 9



EXPLORATORY STUDY ON MONITORING AND EVALUATION OF
LEARNING OUTCOMES AND LARGER IMPACT OF LITERACY,
POST-LITERACY AND CONTINUING EDUCATION
PROGRAMMES IN DEVELOPING COUNTRIES

The Experimental World Literacy Programme:
a critical assessment
The Unesco Press. UNDP, Paris, 1976

Introduction

A fully literate world: such was not the immediate aim—still less the main result—of the Experimental World Literacy Programme. What the programme did attempt was to begin to halt the worsening situation with regard to illiteracy in the world. During the decade 1950-60, although the percentage of adult illiteracy (over age 15) dropped from 44.3 to 39.3, the number of illiterates increased from 700 million to 735 million. Projections made in 1960 predicted, moreover, that the number if not the percentage of adult illiterates would continue to rise.

Awareness of this fact led the General Conference of Unesco to decide at its thirteenth session (1964) to

initiate in 1966 a five-year experimental world literacy programme designed to pave the way for the eventual execution of a world campaign in this field.

This decision was the culmination of a long series of studies and activities devoted by Unesco, since its inception, to the problem of literacy, adult education and what was earlier called fundamental education. It further testified to a growing awareness in the international community of the political, intellectual and economic consequences stemming from the illiteracy of a large proportion of mankind.

The main objective of the experimental programme was to test and demonstrate the economic and social returns of literacy and, more generally, to study the mutual relations and influences which exist or may be established or strengthened between literacy training—particularly among the working population—and development.

A few months later, the Secretary General of the United Nations, when presenting the proposals adopted by the Unesco General Conference to the General Assembly, expressed the hope that the experimental programme would make it possible

to provide valuable information on the relationship of literacy with social and economic development; to ensure that the programme will make a considerable

impact on economic development during the Development Decade in the countries where projects will be conducted; and to prepare the way for an eventual World Campaign for the Eradication of Mass Illiteracy [34].¹

The World Conference of Ministers of Education on the Eradication of Illiteracy met in Tehran in 1965 to consider, in particular, the manner in which national plans for the eradication of illiteracy could more effectively contribute to social and economic progress and to the objectives of the United Nations Development Decade. The conference gave international expression to this change in outlook when it stated in its conclusions that:

Rather than an end in itself, [functional] 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 civil life and a better understanding of the surrounding world, and should ultimately open the way to basic human culture [45].

Guided by these decisions and recommendations, yet without ceasing to give support, within the limits of its resources and in response to requests from Member States, to activities undertaken as part of conventional literacy campaigns, Unesco, the United Nations Development Programme (UNDP) and various national governments have been engaged for the last ten years in conducting a series of pilot projects and micro-experiments.

These took place in a score of countries within the framework of the Experimental World Literacy Programme (EWLP), making use of a network of institutions and regional and subregional centres responsible for assisting national literacy programmes.

For national and international staffs alike, the launching and implementation of EWLP required intense activity including, for example, initial identification missions to some fifty countries, and the organization of training programmes and regional conferences. Perhaps more than in many multilateral aid efforts, EWLP was a dynamic attempt to achieve in a relatively short period measurable progress against a particularly intractable aspect of Third World poverty.

EWLP comprised five major types of activity. The first three types included independent and diversified projects covering one or several experiments or subprojects of four to five years' duration. Among these, nine were carried out with financial assistance from UNDP. Four projects were launched in 1967 (Algeria, Ecuador, Iran, Mali), four in 1968 (Ethiopia, Guinea, Madagascar and the United Republic of Tanzania) and one in 1969 (Sudan).

1. Figures in square brackets refer to the bibliography at the end of the book.

Secondly, one project launched in 1971 (Zambia) was undertaken with international assistance financed from funds-in-trusts. A single project, launched in 1968 (Venezuela), was carried out with national financing.

The final two types of activity included pilot projects where functional literacy was a component organically and administratively integrated with a development project which was the responsibility of an institution other than the literacy project *per se*. Thus, two projects carried out with financial assistance from UNDP were launched in 1970 in co-operation with the Food and Agriculture Organization (FAO) (India and Syrian Arab Republic). The fifth type of project was undertaken with bilateral assistance. This type included two projects launched in 1971 with aid from Sweden (Afghanistan and Kenya) and another project undertaken in 1971 with aid from a Swiss foundation (Niger).

The question uppermost in the minds of concerned individuals and groups at the end of such an international programme as EWLP is: What have we learned? A first important lesson is the extreme complexity of the undertaking, in the case of EWLP at least. The degree of complexity encountered was quite unexpected at the programme's outset and daunted a good many participants, instructors and administrators, not to mention analysts and evaluators. The complexity was all the greater since the present generation of specialists in industrialized countries—who were expected to provide much of the expertise for designing, implementing and assessing EWLP—simply does not have personal experience of literacy as a chronic national problem.

EWLP's complexity, however sobering, is also challenging. The programme had certain tangible and immediate results. Over 120,000¹ men and women were made functionally literate according to the notion of 'functionality' as interpreted by EWLP. In the process numerous obstacles were encountered and it was the various ways in which the programme came to terms (or otherwise) with these obstacles that posed the greatest challenges and offered future literacy action the richest lessons.

It is only natural that these lessons be drawn and converted into guidelines for future action, first and foremost, in and by the countries that participated in EWLP. Nevertheless, an important reason for the provision of assistance to national EWLP projects by the United Nations system was the hope that broader lessons would emerge and be of use for the international community as a whole.

The present report attempts to set out just such lessons. It is based essentially on the experience of countries that received assistance from UNDP and Unesco under EWLP: Algeria, Ecuador, Ethiopia, Guinea, India,² Iran, Madagascar, Mali, Sudan, Syrian Arab Republic,² United Republic of Tanzania. Some of these countries have already reoriented their literacy efforts in the light of their participation in the programme. Other

1. This figure is the estimated result of projects in five EWLP countries on which data are available.

2. Projects also aided by FAO.

countries in turn may find in this report valuable guidelines for future action. Since the programme was a major international effort and therefore more than the sum of its individual parts, certain of the lessons suggested here may also have implications for multilateral and even bilateral co-operation in the education field and beyond.

The data have been drawn primarily from the following national and international sources:

Project files at UNDP and Unesco headquarters (these included interim and final project reports by chief technical advisers and evaluation specialists; periodic reports of UNDP resident representatives; reports of UNDP-Unesco joint evaluation missions; financial statements on projects).

Preliminary drafts of the technical reports on various aspects of EWLP prepared by the Evaluation and Research Unit attached to Unesco's Literacy Division.

Reports and research documents from universities and governmental and non-governmental institutes and organizations, and studies carried out by individuals.

Interviews with a variety of specialists at different levels who were involved with aspects of the Experimental World Literacy Programme.

The report synthesises information from the various sources and uses material considered most significant in the context of all data available to the authors. Much of the background documentation used is not available in published form and thus does not appear in the bibliography. Those published items directly referred to in the report do appear in the bibliography. A more extensive general bibliography of the Experimental World Literacy Programme is available from Unesco's Literacy Division.

This report approaches EWLP in two ways. First, it analyses EWLP vertically, with a series of project profiles. That is, each project is described and analysed so as to elicit particularly salient information in the following areas, as appropriate: time perspective; policy and objectives; participants; administration and organization; teachers and other personnel; curriculum and methods; teaching materials; costs; and evaluation and research. Each profile ends with a summary which attempts to single out some of the major lessons learned from that project.

Secondly, all projects are analysed—and lessons drawn from them—horizontally. That is, global lessons are sought under the following four headings: Why EWLP?; The wherewithal; Doing the job; The results. Particularly in these chapters a certain paradox appears. On the one hand, problems encountered by EWLP resulted from its innovative aims and character. But on the other, at least some of its limits were imposed by the fact that its implementation depended on traditional forms of technical assistance. From this contradiction emerged, happily, not only problems but also numerous positive guidelines for the future.

This report attempts to make a genuinely honest assessment. Indeed, because it seeks to derive lessons for the future, it may tend on occasion to

focus on what went wrong with EWLP rather than what went right. In this way it heeds a warning sounded in 1972 by Unesco's International Commission on the Development of Education in its report *Learning to Be*. Speaking in the name of those concerned with literacy, the commission stressed the need for frankness, so as to avoid the reproach that 'we are succumbing to demagogy in this particularly delicate field of social action' [10].

Thus no claim is made for quantitatively airtight argumentation. The report does not hold up a single 'model' of literacy action for replication or even adaptation. Nor does it defend any single approach to literacy. Rather, it strives to offer information and ideas that could be of use nationally and internationally, in projects ranging from micro-experiments to mass campaigns. In this pluralistic and self-critical spirit, the report asks more questions than it even begins to answer.

The point, then, is to encourage rather than avoid debate, to stoke up rather than damp down even unorthodox thinking about the various subjects covered here. It has been felt that this would be the most useful—and perhaps most readable—manner in which to make known the issues raised by EWLP. If these issues have a rather uneven quality about them, this is because the programme's implementors were understandably more concerned with doing than with recording. As a result, *post facto* information available was of varying reliability. This made an essentially qualitative approach the only—if not necessarily the best—way to interpret the results.

The results—effects on participants; what did it cost?

In trying to determine what the Experimental World Literacy Programme actually achieved and what changes resulted from it, it is well to be prudent. Perhaps because of its newness among international non-formal educational efforts, EWLP's evaluation encountered severe difficulties.¹ On the whole, it did not yield internationally comparable results, and even regarding national projects extreme care must be exercised in interpreting the data. Considerable precautions have, of course, been taken to collect reliable data and to base analyses on such data. Nevertheless, nearly every figure quoted in this chapter must be considered as at best an estimate of a given national reality, made under generally difficult circumstances. Moreover, it must be remembered that comparisons between doubtful representations of national realities makes it even more difficult to assess the viability of the functional literacy approach as applied in EWLP.

By and large, it is also impossible to state with much confidence what would have happened if EWLP had not taken place. World attention might be less aware of the growth of illiteracy today were it not for the programme. Such affirmations, however, can only be speculative, the more so since EWLP did not take place in an international void. It was part of a sharpening awareness of, and growing effort in, literacy. Certain of its apparent results may, then, be due to other factors or forces. Similarly, the programme may have had far-reaching impacts that cannot yet be positively attributed to it.²

With these provisos clearly in mind, what can be said about EWLP's results? The basic questions are: What was learned? What was done with what was learned? What were the costs and economic advantages of literacy as achieved in EWLP? And what were the programme's impacts on further educational action?

1. Problems affecting the EWLP evaluation were discussed above on page 151 et seq.

2. Perhaps summative (terminal) evaluation of such international aid efforts could usefully wait five or ten years, since this larger time span may be necessary to ascertain who learned what lessons with what result.

What was learned ?

Pass rates

What did the million-odd illiterates reached by the programme actually acquire in the way of skills and knowledge? There was wide country-to-country variation among pass standards and rates of learners who took examinations on completing different parts of their course. The following table gives success percentages for projects on which data are available. The figures refer to samples, of limited scope in some cases, of participants who finished a full literacy programme (usually two cycles). Thus, success percentages among initial enrollees were comparably smaller. The percentages given here range from the classes with lowest success rates to those with highest rates in each project. The wide variation is due, in some cases at least, to the fact that teaching materials and methods were employed during the project with the result that lower tests gave higher success percentages.

	Ecuador	Ethiopia	India (Jaipur)	India (Lucknow)	Iran	Sudan	United Republic of Tanzania
Writing	38-97	50-93	74	71-80	40-100	62-92	
Reading	73-100	58-97	96	89-94	10-89	70-97	
Arithmetic	2-89	55-91	76	75-80	3-91	40-93	7-17
Vocational training	6-95	50-92	85	88-93	—	41-96	

In rounded absolute figures¹ the number of people progressing from (a) initial enrolment through (b) sitting for final examinations to (c) success in those examinations were as follows (in projects for which data are available):

Country	Initial enrolment	Examinees	Pass
United Republic of Tanzania	466,000	293,600	96,900
Iran	97,400	46,900	13,900
Ethiopia	36,800	21,700	9,300
Ecuador	17,500	10,000	4,100
Sudan	7,400	2,400	600

The actual significance of these pass figures in terms of learning is hard to determine. Test levels and criteria varied widely from subject to subject, context to context, project to project and perhaps also within projects. Arithmetic considered 'basic' in an automobile repair course might be 'advanced' in a weaving programme. Automobile repairs (or weaving) also require very different amounts of learner effort depending on previous familiarity with cars (or looms). Similarly, 'written expression' means one thing for examinees writing in their mother tongue and something quite different and much more arduous for those who have had to master both writing and a new national or foreign language.

Another important factor affecting examination performance was the level of pre-course knowledge. It was generally assumed in EWLP that the

1. In all cases these figures are estimates of varying reliability.

beginning learners were fully illiterate. This was far from always the case. In Ecuador, at least a few participants were primary-school leavers anxious to acquire vocational skills. In Iran, a baseline survey found that many learners were already able to do arithmetic on enrolment. Indeed, ignorance of reading and writing is probably never automatically linked with inability to do mathematics since all societies have their own kinds of numeracy that do not require reading or writing.

Conditions and methods of measuring knowledge and skills also influenced results, both internally and with regard to national and international comparability. 'Yes/No' test questions were used in at least one case. They are, however, patently an irrelevant imposition in cultures where it is impolite to say 'no'. Measurement of individual achievement is alien to communities that prize group solidarity. In such settings it is perhaps also unreasonable to expect a farmer-instructor to exercise Western-style detachment when marking papers written by his neighbours and friends.

There was widespread frustration among EWLP organizers and evaluators over the obstacles to reliable testing. The examples just given suggest that these obstacles may sometimes have stemmed from the irrelevance (or threat) of 'modern' assessment methods to the 'traditional' mores of illiterate communities. In future, a sensible and sensitive approach might be to attempt to devise tests more compatible with local customs, even taking into account the extreme difficulty and expense of such an effort. Attempts in this direction could even enhance (rather than undermine) certain nationally favoured values. Why stress and test individual achievement if group solidarity is fundamental to a culture? Could not joint examinations requiring co-operation rather than competition be devised?² The question is perhaps worth at least some experimentation.

*Sensible
and
sensitive
testing*

A more flexible international approach to assessment of learner progress would have implications reaching far beyond the redesign of testing instruments. Also necessarily at stake would be curricula, teaching methods and personnel, and the very function of literacy in a given zone or country. More pluralism would doubtless make international comparability of result measurement even more difficult than it was in EWLP, but that is possibly not too great a price to pay for achieving the relevance that, in principle, lay at the heart of the idea of functionality. Giving priority to the needs of national—rather than international—decision-makers would also better reflect the idea (now generally accepted though not always practised) that aid should be partnership, rather than an unequal relationship between 'donors' and 'beneficiaries', with accountability solely owed by the latter to the former.

Seeking to place the levels of learning attained by new literates in a broader educational perspective, EWLP organizers and evaluators inevi-

1. Perhaps sometimes equated with 'backward'.
2. Some two decades ago, *animation rurale* educational programmes in Morocco relied heavily on group testing whose instruments were based on Moroccan folklore and mores.

tably drew comparisons with primary schooling. Naturally, schooling fulfils certain specific roles not played in functional literacy, and vice versa. Excluding these non-comparable functions, however, it does seem possible to make broad comparisons about the area of overlap, i.e. the skills and knowledge comprising the fundamental training that both literacy and primary schooling seek to provide.

Analysed in these terms, it appears that the level of basic training acquired by new literates can be roughly likened to the amount of basic training acquired in a given length of primary schooling in the same country. The basic training acquired by a primary-school pupil over a period of two years was more or less achieved by new literates during courses lasting a total of 700 hours (Ethiopia), 600-700 hours (Mali), 560 hours (Syrian Arab Republic) and 270 hours (United Republic of Tanzania). In Algeria and Ecuador, it was four years, for respective totals of 426 hours (three cycles in Algeria) and 400 to 746 hours (two cycles in Ecuador). In Iran, a five-year equivalent was attained in two-cycle programmes varying from 208 to 554 hours. Conventional wisdom¹ has it that a minimum of roughly four years is necessary for primary schooling to become self-maintaining. In the light of the above figures, it would seem indispensable for EWLP countries to offer further education to new literates if these are not to relapse into illiteracy.

What was done with what was learned? EWLP's overriding goal was that functional literacy should change the new literates' relationship to their socio-economic milieu, and the milieu itself. The programme evaluation sought to detect these changes in three major respects: 'insertion into the milieu', 'mastery of the milieu' and 'transformation of the milieu'. About 160 observations and measures were made in various projects. In overall terms, the influence of functional literacy was judged to be plausible and favourable in about 42 per cent of the observations.

However, a relatively low percentage of new socio-economic behaviours sought through the various curricula were actually adopted (e.g. estimated at 30 per cent, in Iran); of these changes, a number were never the less considered significant.

Looking more closely at the results, one may begin with 'insertion into the milieu'. Here, indicators were devised to ascertain changes in areas such as interest in further education, management of personal finances, exposure to mass media, the seeking out of technical advice, use of the three R's, and participation in formal organizations. Eighty-six per cent of actual changes observed were judged satisfactory (56 per cent were statistically significant) as against 14 per cent declared unsatisfactory (6 per cent were statistically significant).

These results, and particularly the values attached to them ('satisfactory' or not) raise a fundamental question: into what milieu is the illiterate to be 'inserted'? Point by point examination of certain 'insertions' attempted

*Insertion
into what
milieu?*

1. cf. for example, D. Harman, *Community Fundamental Education* [17].

and observed in India and Iran—probably reflecting similar efforts in other EWLP countries—offer at least the beginnings of an answer.

The Jaipur and Lucknow subprojects in India gave rise, among their farmer-graduates, to significantly increased desire for technical information and to significantly intensified behaviour designed to satisfy that desire. That is, they were stimulated and enabled to seek out in the maze of possible sources of technical information those most likely to help them solve their daily and mainly farming problems, usually—although not always—posed in individual terms.¹ This seems to have been linked to the fact that the project stressed the economic gain that the individual could expect to reap by becoming literate.

On the other hand, the results regarding participation in formal organizations were less conclusive. There was only a tendency (not statistically significant) for new literates in the Jaipur and Lucknow subprojects to increase participation in bodies such as co-operatives, youth clubs, political parties and village councils. Certain observers explained this trend as expressing a lack of propensity for social affiliation in the two districts studied, and/or defective organizational structure (one might also wonder if, in fact, it is known how to teach these behaviours).

Perhaps, too, the organizations selected as indicators were not perceived by learners as useful means of attaining the individualistic and productivistic aims stressed by the subprojects. Additionally (or alternatively) such organizations may be perceived as part of a local power structures not particularly favourable to at least certain strata of new literates. Studying the role of *panchayats* (village councils) in the first decade of Indian independence, S. C. Dube² concluded that, despite the democratization function assigned to them by the central government, they tended to be dominated by land-holders. Consequently, the intentional concentration of power in the councils' hands by higher authorities

Participation in formal organizations

has been construed by some villagers as an effort of the government to maintain the status quo within village communities, and indirectly as a step to support the domination of landowning groups [8].

In Iran, an effort was made to evaluate literacy's impact on the use of writing and arithmetic in the maintenance of savings and bank accounts. A favourable trend (not statistically significant in the first case, but significant in the second) was found, but in the second case it could not be determined whether the trend was attributable to literacy courses, or to the concurrent opening of a bank branch accessible to the new literates under study.

Also in Iran, functional literacy was determined to have had a favourable impact on family planning, i.e. was identified as a plausible cause of the trend toward reduction of the family size considered ideal by new

1. Some of these problems were considered 'collective', but only a minority actually demanded joint action for their solution. The 'collective' problems were 'collective' in that most were faced by many or most individuals in the community.

2. Observations by O. Lewis [23] and M. Opler [31] echo this report in vivid terms.

literate to two or three children. This effect was described as placing new literates before the respective advantages of large and small families, a choice between more children or the hope for a higher material living standard, with stress on the latter. It may be interpreted as having (among other things) enabled the people concerned to understand better the long-range implications of such choices and thus to choose more rationally, although subsequent behaviour has yet to be demonstrated.¹ In this interpretation, literacy is seen to be at the very least a vehicle for increasing the responsibility of new literates in selecting family size.

According to another view, underdevelopment does not result from large populations so that development does not follow automatically from birth control. In this interpretation (which is not of course shared by all observers), the poor family that assumes responsibility for limiting its size, simultaneously and tacitly accepts responsibility for its poverty—without obtaining any assurance of being able to improve its lot. In essence, it is probably unrealistic to assume that functional literacy with a population education component will in itself have significant impact on family size decisions unless the educational effort is in the context of other development activities which offer tangible incentives for small families.

More examples could be given of EWLP socio-economic effects regarding the insertion of new literates into the milieu. These suffice, however, to suggest the outlines of an answer to the question: insertion into what milieu? From the foregoing, one can begin to see the kinds of behaviours deemed important by the planners and evaluators of EWLP projects. The participant who was deemed 'successful' seemed to: (a) actively seek information likely to help solve mainly personal problems, generally posed in vocational terms; (b) prefer such activity to participation in formal community organizations; (c) take advantage of his or her new literacy and numeracy skills to maintain personal bank and savings accounts;² and (d) aspire to reduce the size of his or her family in exchange for the prospect of a higher material living standard.

This admittedly sketchy profile (which—it must be stressed—is by no means necessarily representative of all programme results) diverges slightly from the ideal 'inserted' new literate projected by the evaluation criteria. These criteria reflected 'functionality' as understood (or misunderstood) nationally and internationally in EWLP.

As regards these four behaviours, at least, the programme was most successful (in terms of its own evaluation criteria) when it reproduced the narrowly economic interpretation of functionality. This was notable in respect of knowledge and skills that could be put to work for primarily personal material benefit, whether immediate (wanting and seeking

A profile

1. This in an area where larger-term socio-economic effects of literacy (like other international aid programmes) cannot yet be measured.

2. Given the relatively favoured status of people liable to have bank or savings accounts in most Third World countries, this expected behaviour lends further substance to the 'elite of the impoverished' idea already discussed.

vocational information), intermediate (replacement of children with material goods) or longer-term (savings).¹ It was less successful when it expected new literates to participate in formal organizations, either because the new literates were not gregarious or possibly because such organizations did not enhance their prospect of primarily personal material benefit or did not favour them in the local power structure.

The milieu into which new literates are inserted—i.e. the behavioural profile into which functional literacy helps them to grow—would thus seem to be defined by the values of a society in the process of a certain kind of modernization. This modernization appears to correspond rather closely to a model of development embraced and propagated by industrial nations (perhaps particularly those of Western Europe and North America). An earlier discussion (see page 151 et seq.) concluded that just such a model was projected by at least certain of EWLP's evaluation criteria. It can now be hypothesised that, in the instances just examined, the programme was more successful than not in achieving the insertion of new literates into a milieu corresponding to this model. Before testing this hypothesis against other classes of socio-economic results—mastery of the milieu and transformation of the milieu—a word may be said about the process of insertion.

In the cases just described, EWLP's function was apparently to provide both a stimulus (the prospect—if not always the reality—of primarily personal material benefit) and the means (various information and skills) for giving a 'satisfactory' response to that stimulus. Was, then, the process akin to behaviourist conditioning? The very notion of 'insertion' suggests that something was done to illiterates by external agents, and that the something in question was determined by those external agents, who also decided whether responses to their stimuli were 'satisfactory'.

Without more complete data on the 'insertion' effect of functional literacy, the question can only be asked, not answered. But it must be asked, for it is fundamental. Did the Experimental World Literacy Programme tend to stress knowledge and know-how at the expense of know-why (critical awareness of one's role in society)? If so, it may have been successful in adapting the new literate to existing circumstance. But education should also help learners to know how to shape the environment.

New literates' mastery of the milieu was the second broad class of intended socio-economic effects examined by EWLP evaluators. Here, indicators were prepared to measure changes in such areas as behaviour at work, knowledge of modern technical practices, adoption of such practices, and conservation and reproduction of the labour force. In this class, 93 per cent of changes recorded were judged satisfactory (51 per cent significantly so), while 7 per cent of changes were deemed unsatisfactory (2 per cent significantly so).

In India, the Jaipur and Lucknow subprojects were evaluated for their success in encouraging the effective use of certain new resources (improved

*Knowledge
know-how
and
know-why*

1. One EWLP primer exhorted learners to 'make money—don't waste time'.

seeds, fertilizer, pesticides, related tools) and in the case of Jaipur a certain number of associated practices (row planting, extensive cultivation, irrigation). On the whole, results were found to be positive, more clearly so in Lucknow than Jaipur, where only a positive trend was found—probably because of the disruption in farming caused by a concurrent period of drought.

There were, however, interesting variations within the generally positive results. In Lucknow, for instance, although new literates were found to be much better informed about sorghum than the control group of non-learners, they did not tend to apply their knowledge. This seems to be due to the declining profitability and popularity of sorghum compared to wheat. Regarding wheat, a widespread and marketable crop, the Lucknow subprojects achieved results assessed as extremely favourable. A trend in the same direction was found in the Jaipur subproject. There, the aim was to bring about adoption of new high-yield strains of wheat and millet, and learner response to this innovation was judged positive.

In both cases the socio-economic strata reached seem to have been largely made up of small peasants. Indeed EWLP's apparently greater success in contributing to adoption of at least some technical practices recommended in agricultural courses (a positive programme-wide average score of + 54 per cent) compared to industrial courses (a positive programme-wide average score of + 45 per cent) may perhaps be attributed to the fact that agricultural participants were frequently autonomous farmers or small landholders and thus better placed to apply immediately the new practices taught them than workers who were limited by the constraints of enterprise production.

*Joining
the green
revolution*

In India, and possibly elsewhere, it seems that thanks to the Experimental World Literacy Programme small peasants—members of what was earlier called the 'élite of the impoverished'—were enabled to join the green revolution. Thus the programme helped answer criticism (notably by Ladejinsky [21]) that India's green revolution was primarily benefiting already wealthy landowners, which most Lucknow and Jaipur learners plainly were not. By joining the green revolution in India and elsewhere, however, new literates exposed themselves to its servitudes, even if their personal incomes tended to rise. A brief discussion of the green revolution is therefore vital to understanding the milieu the new literate was to master, and the ways in which this mastery was to make itself felt.

The green revolution requires considerably increased investment, not only in the new strains of seeds but also in the additional resources and tools that, as EWLP taught, are necessary to achieve increased crop yields per land area. According to Rochin's report on West Pakistan, larger farmers may often be in a better position to purchase such accessories than small landholders [36]. Paradoxically, newly literate farmers who join the green revolution may find themselves in a double dilemma since if they are able to purchase new seeds and necessary accessories, how much of their new income has to be re-invested in this way?

Like EWLP, the green revolution tends to be more or less highly selective. In geographical terms, it is generally limited to areas most likely to yield short-term profit. According to Cleaver's study, the green revolution in India seems likely to be incapable of affecting meaningfully zones outside the 20 per cent of cultivated land that is irrigated (only half of it with sure water) [5].

Socially, the green revolution is also selective. It tends to favour the already favoured minority of the rural population, but its benefits do not generally reach the majority. According to Rudra *et al.* [39] and Ladejinsky [21], wealthy Indian farmers already seem to be extending their holdings, thus raising land prices. Cleaver [5] predicts a possible 'considerable rise in rural unemployment in those areas where mechanization proceeds rapidly'. Bardhan [3] points to a trend, in areas of India where the green revolution has been effective, for real wages of agricultural labourers to decline.

In the light of this discussion, what seems to be the provisional profile of a new literate as 'master of the milieu'? There is no doubt that, in the case discussed at least, he or she does exercise increased mastery of his or her milieu. But is it not mastery of a rather narrow technical and economic kind corresponding to the profile suggested earlier of the new literate 'inserted into the milieu'? In acquiring this kind of mastery of the milieu, to what extent has the new literate become dependent on which external socio-economic processes and forces? Has literacy enabled the new literate to know and understand these processes and forces? To come to grips with them? To have a voice in controlling them? What implications has the new literate's accession to mastery of the milieu for the fate of his or her less favoured neighbours and compatriots?

It is probable that these questions have no simple or single answer, either for any specific case examined or for EWLP as a whole. They are none the less probably legitimate and basic enough to require answers in future literacy (and other aid and development) efforts, particularly where the United Nations and its Specialized Agencies—with their special commitment to human rights—are involved. In seeking answers, President Nyerere of Tanzania has some provocative thoughts on education and human beings' relationship with their milieu. He makes

a serious distinction between men and women who are skilful users of tools, and a system of education which turns men and women into tools. I want to make quite sure that our technical and practical education is an education for creators not for creatures . . . that our educational institutions are not going to end up as factories turning out marketable commodities. I want them to enlarge men and women, not convert men and women into efficient instruments for the production of modern gadgets [29].

*Creators
or
creatures?*

The third and final broad class of EWLP's socio-economic effects examined by international evaluators was the participation of new literates in the transformation of the milieu. In this respect, indicators were devised to

to detect changes in areas involving the means of production, the volume of production, cash income, income in kind and the consumption of durable goods. On the whole, 90 per cent of recorded changes (the absolute numbers of which were relatively small) were deemed to be satisfactory (41 per cent significantly so) and 10 per cent were judged unsatisfactory (7 per cent significantly so).

In the Lucknow subproject (India), a list was drawn up of sixteen material goods whose pre- and post-course possession was noted among EWLP learners and non-participants in the literacy programme. The list included items such as bicycles, sewing machines, pressure lamps and chairs. At the end of the course, new literates were found to possess markedly more of these goods than the non-participant control group. Evaluators concluded that increased consumption was a probable consequence of income growth due, in turn, to enhanced agricultural knowledge and skills and resultant innovative professional behaviour.¹

*Stimulus,
response,
reward*

The results help to fill out still further the provisional profile of the new literate emerging from EWLP, in the case analysed at least. The new literate 'inserted into the milieu' has been stimulated by the prospect of personal material gain and equipped with the knowledge and know-how to supply appropriate response to that stimulus. As a 'master of the milieu', the new literate does seem effectively to supply the appropriate response. It is only logical then that, completing the circle, the new literate should be rewarded: he or she (and family) appears to gain access to increased personal consumption of material goods.

In summary, considerable caution must be exercised in interpreting the success criteria discussed above. Relatively few (probably less than a third) of all socio-economic changes advocated in the various programmes were, in fact, ever adopted by participants. And these changes, on the whole, were those which were perceived by participants as having some immediate personal benefit and requiring little expenditure of time and resources to achieve. Thus, notwithstanding the high percentages of such modest changes which were regarded as 'positive' by the Evaluation Unit attached to Unesco's Literacy Division, the fact is that these positive changes were very modest in nature and few in number.

If such modest changes in the three major classes of behaviour evaluated can be said to prove that EWLP was successful, and to the extent that this logic of project success served the development model chosen by participating countries, the programme was worth while. However, the question remains as to whether the narrowly economic and individualistic nature of EWLP's logic did not diverge from—or even contradict—the kind of development to which at least a few participating countries aspired.

Whatever the answer to this question, it is clear that evaluation

1. The validity of this study is open to question. Considering the high dropout rates, the successful participants were a highly self-selected group in the experimental (literacy class) population that cannot be appropriately compared to a random sample of the control (non-literacy class) population.

providing immediate information could only reveal the short-term and most mechanistic socio-economic effects of the world programme. This necessary focus may have biased at least some of the questions asked in this chapter. It certainly makes it difficult even to guess at some of the longer-term and possibly profounder effects wrought by EWLP. Such results doubtless already exist in latent form, and will make themselves felt in the future, if for no other reason than the liberating potential of literacy.

Possibly of necessity, evaluation took into account chiefly the intelligibility and utility of certain messages for learners who were viewed primarily as receivers. Nevertheless, the distinguishing feature of verbal and written symbolism is that it can have a wide variety of functions. It is endowed, in particular, with the ability to broadcast as well as receive, and to broadcast an almost infinite number of entirely new messages. Unlike computers, human beings cannot be programmed to respond only to certain stimuli, or to give only certain responses to those stimuli.

An already palpable effect of EWLP in certain Malian villages appears to have been at least the beginnings of a new social dynamic involving a redistribution of socio-political initiative. In the United Republic of Tanzania, Hall and Mhaiki report as follows on the results achieved by 1972 by adult education, including that country's EWLP project:

*Other
uses of
literacy*

Stories are now being told of farmers learning to read and write and denouncing individuals who had cheated them by mis-weighing their crops when they brought them for sale [16].

Such are the inevitable uses of literacy where domination and exploitation persist, as Shakespeare well understood. In *The Tempest*, the powerful Duke Prospero uses magic (which might be likened to modern technology) to conquer a tropical island and enslave its one inhabitant, Caliban.¹ Having given Caliban enough instruction to understand orders, Prospero can hardly be surprised when his slave turns learning to other purposes:

You taught me language, and my profit on't
Is, I know how to curse. . . .

Costs and project economics

What did the Experimental World Literacy Programme cost? Who paid for what? What does it cost to produce a new literate? How could costs be reduced? Was basic training provided under the programme cheaper than basic training provided by alternative sources? Attempts to answer these

1. 'Caliban' is an anagram for 'cannibal', a word stemming in turn from 'Carib'. Fernandez Retamar views the Prospero/Caliban relationship as typifying that of colonizer to colonized [11]. A symbolic parallel is found in Mozart's *The Magic Flute* when the simple bird-catcher Papageno is punished for mocking the powers-that-be by being obliged to sing a whole aria with his mouth padlocked shut.

questions are perhaps particularly subject to the warnings expressed at the beginning of this chapter. Nevertheless, even attempts to find answers may be of interest. The total cost of EWLP during six years of operation was of the order of \$32 million¹ with expenditures per national project ranging from \$300,000 to \$6 million. The governments concerned paid roughly two-thirds of global costs, although the proportion of national financing varied in function of the size and nature of projects. National financing accounted for 70 per cent of expenditure in projects that tended toward a massive approach, but as little as 25 per cent of spending in the more properly micro-experimental projects. This means that the smaller and more experimental a project, the larger was the financial responsibility of international financial resources. This correlation probably expressed the earlier overriding but now changing preference of the international agencies concerned for what was previously called the technico-scientific approach to experimentation, contrasting with the activist-pragmatic approach. The latter approach seems to have been more popular with national governments, particularly those who viewed their EWLP project as a prelude to broader literacy action.

Regarding international aid to national projects, some project country nationals may have expected that it would take the form of direct cash flows, capable of easing (no matter how little) sometimes disastrous economic situations.² But EWLP followed the traditional pattern of technical assistance under which help was furnished primarily in the form of expert services, fellowships and equipment—but not cash. Disappointed nationals may, then, occasionally have seen the programme as something of a 'Now you see it, now you don't' conjuring trick.

How were budgets used? A study of one EWLP country pointed to the need to separate operational costs from research and development costs. Excluding spending on research and development, the cost per enrollee was \$17, and per participant completing both parts of the literacy programme approximately \$31. These costs rise to \$26 and \$50 respectively if spending on research and development is included (see the cases in Part I for figures on each project).

In more detailed terms, this study found the following breakdown into percentages of overall expenditure: remuneration of instructors, 30; preparatory and research studies, 23.5; administration expenses, 21; transport, 8; training of instructors, 6.5; teaching materials, 5; classroom equipment, 4.5; audio-visual material, 1.5.³ Figures available on other

1. This figure does not include the cost of a certain amount of staff time, or of aspects of the international evaluation.

2. International aid's overall impact in this broader context was pointed to in 1970 by the Principal Secretary of the United Republic of Tanzania's Ministry of Finance when he observed (quoted by Rodney) that 'losses from sisal price declines [sisal is a chief Tanzanian export] have more than equalled total gross aid receipts over 1964-69' [37].

3. These figures do not seem to include capital expenditure, e.g. construction costs, since they refer to the expansion—rather than to the experimental phase—of the project in question. Nor do they appear to cover the cost of international experts. The item 'preparatory and research studies' seems to include baseline studies, curriculum development and evaluation.

projects suggest that this may not be an especially atypical budget, and that instructor remuneration was generally the most expensive single item of global spending, combining national and international outlays for operations and research and development. If salaries of all kinds were grouped, this category would make up over 80 per cent of all project costs, with remuneration of instructors making up less than half of all project salaries. Other salaries, of course, were for those involved in research studies, administration, preparation of materials, etc. Such high non-teaching personnel costs indicate a rather large and perhaps cumbersome infrastructure in most projects.

What did it cost to make a person literate? The following figures should not be compared at face value since they are unweighted. That is, they do not take into account sometimes wide country-to-country variations in testing levels or in the amount and quality of learning required to achieve what was defined as literacy.¹ Keeping these differences in mind, the rounded costs per participant taking the final examination in countries on which data are available were as follows.² These figures include national and international financing, take capital as well as running costs into account, and compare total expenditure (operations plus research and development) with operational costs alone (not including research and development):

	Total estimated cost per enrolled participant	Total estimated cost per final participant	Estimated cost per final participant excluding research and evaluation (operations only)	Estimated total cost per participant passing final examination
	U.S.\$	U.S.\$	U.S.\$	U.S.\$
Sudan	7.00	272.00	110.00	269.00
Madagascar	112.00	126.00	48.00	—
Ecuador	70.00	123.00	70.00	300.00
Iran	49.00	100.00	76.00	332.00
Algeria	71.00	99.00	83.00	—
Ethiopia	54.00	92.00	63.00	212.00
Mali	14.00	35.00	17.00	—
United Republic of Tanzania	7.00	10.00	9.00	32.00

One way of reducing costs might be to carry out more careful analyses of the administrative and organizational patterns in each project than was possible in the profiles of Part I of this report. It would certainly seem (as mentioned above) that many projects spent large percentages of their budgets on administrative and project support infrastructure. Others spent much less by assigning responsibility for various aspects of the programme to officials in charge of other technical or social services.

Secondly, the foregoing table shows that costs per successful

1. In the United Republic of Tanzania, for instance, reading, writing and computational skills roughly comparable to second year of primary school were accepted as criteria of literacy compared to fourth- or fifth-grade levels in several other countries.
2. In all cases, these figures are estimates of varying reliability.

participant were much higher than costs per enrolled participant because of the high dropout rate. Reasons for high dropout rates in many projects are not entirely clear, and relate to factors over which projects may have had some control as well as factors over which they had little influence. It is reasonable to conclude, however, that those projects which had lower dropout rates did something that was more effective in holding participants, and reduced wastage and effective cost per successful participant.

Thirdly, given the experimental nature of EWLP and the rather large amounts already spent on research and development, it may be that economies of scale (the bigger the project, the lower the cost per learner) could be achieved by extending literacy work to larger audiences. Since the methodological research and development have already been carried out, expansion, according to such reasoning, would essentially require the replication of operational spending.

The prospect of achieving economies of scale also seems substantiated by the fact that the lowest *per capita* costs appear to have been achieved in projects reaching the largest audiences. Iran, Algeria, Ethiopia, Mali and the United Republic of Tanzania, the countries with comparatively low total costs, respectively reported the following numbers of candidates for final tests: 46,239, 38,784, 21,722, 50,000 (approximate) and 293,586. Conversely, countries with relatively higher total *per capita* costs—Sudan, Madagascar and Ecuador—had a rather smaller number of final candidates—2,363, 3,826 and 9,988 respectively.

*Economies
of
scale*

Unfortunately, the possible economies to be reaped by extension of the scale of functional literacy do not seem likely to be as great as initial theory suggests. Research and development costs may decrease in cases where widespread application of programmes already devised during EWLP is possible. As earlier chapters of this report have suggested, however, curricula and pedagogy probably require considerable revision. Moreover, pre-existing programmes will have to be modified, or new programmes prepared, if extended literacy is to be functional in terms of the specific problems and aspirations of additional groups with geographical, professional and/or ethnic particularities. Technical assistance should keep these possibilities in mind when helping governments to develop national capacities. Equally compromising for the prospects of economies of scale are trends already detected by evaluators within EWLP regarding operational costs. Although these do seem to fall off with the expansion of literacy, there appears to be little hope for a radical reduction.

Finally, a mass campaign implies a commitment of the government to move towards a fully literate society requiring an infrastructure of reading materials, newspapers, libraries, reading rooms, extension materials, and all the other things necessary to support the intellectual interests of literate people. Thus, a long-term plan for a literate society must take into account expenditures that probably far exceed the cost of the original literacy programme in order to establish and maintain such an infrastructure.

Limiting ourselves for the moment, however, to the costs of literacy

programmes, what can be done to reduce such costs? Basing themselves on available EWLP data, Unesco/UNDP analysts identified four 'sensitive' cost factors and attempted to project the magnitude of savings that might result from slight variations of each.¹

Average class size was the first factor. It was assumed that class size would be increased by 1 per cent. A reduction in overall costs of between 0.4 and 3 per cent—depending on the project—was considered a probable result. This means that for each new group of 100,000 enrollees, savings ranging from \$22,000 to \$300,000 could be realized, depending on the average size of classes. Naturally, pedagogy as well as economics should inform any decision to increase class sizes. Although classes with a 1 : 8 instructor to learner ratio (e.g. Guinea) might not suffer irreparably from adding a few more pupils, the EWLP-wide ratio of 1 : 30 or more should probably not be stretched much farther, assuming that there is any truth in conventional wisdom on class size and assuming the need for participation in discussion and in technical activities involving demonstrations and the like.

The second sensitive area covered the educational know-how required to construct programmes. Such know-how results from research and experimentation which, as already seen, can absorb a percentage of budgets that grows as projects become smaller and more experimental. Here, a 1 per cent reduction results in a saving of between 0.6 and 0.9 per cent of research costs per learner. In certain cases, this economy would not result from standardization of programmes, given the already-mentioned need to remain functional in terms of specific learner groups. It would, rather, arise from the improvement and normalization of programme design techniques. Once again, however, projections should be conservative rather than optimistic, since constant self-assessment would probably require continuing readjustment of even these techniques.²

Most projects, of course, assumed that a materials development and validation infrastructure was needed for the project, quite independent of anything else going on in the country. Many countries have (or are developing) various kinds of curriculum centres designed to prepare more effective teaching materials for the schools. In addition, many countries have agricultural, health and other extension services. With an increasing concern for lifelong education that does not compartmentalize educational opportunity as rigidly as in the past, it would seem reasonable to consider educational materials research and development centres which would not be limited to the preparation of literacy materials or school materials or extension materials. The process of definition of educational objectives, the design of instructional or educational materials to suit a particular organizational 'strategy' involving formal or non-formal education, and the validation and revision of such materials involve similar skills and

1. In each case, all other costs were supposed to remain unchanged.

2. Indeed, 'normalization' may be an illusion in a field like literacy, where innovation is permanent, and the only basic constant may be perpetual change.

resources. The need to have separate expensive centres for literacy materials would appear to be open to question. In fact, the creation of teaching materials and literature in the national languages, is compatible with this idea.

A third flexible cost factor is expenditure on teaching staff. Given the above-noted importance of such expenditure in virtually all EWLP projects, this is perhaps the most influential single item of operational budgets. A reduction of 1 per cent here leads to a saving of more than 0.5 per cent, i.e. an average possible economy of about \$40,000 for each new group of 100,000 enrollees. Because of relatively high salaries, fringe benefits and other career perquisites, school-teachers and civil servants seconded to literacy work generally cost more than non- or para-professional instructors. It is therefore in this second category that imaginative efforts might reap considerable savings. The United Republic of Tanzania's experience suggests that it is cheaper (although perhaps not radically so) to use literate peasants as instructors rather than professionals. Other categories could also be tapped. Students at universities, teacher-training colleges and secondary schools, for instance, might be required to make a given number of people literate as a condition for graduation. In such cases, the teaching work could add an invaluable experience of reality to the education of student-instructors, who tend to come by definition from the more favoured socio-economic strata.

Finally, evaluators attempted to project savings that might accrue from reducing running costs. Unfortunately, this factor was found to be too dependent on other factors to be treated in isolation from them. It was nevertheless thought likely that savings could be achieved if literacy action were more integrated into the institutional framework (farms, factories, development agencies, etc.) of the localities where it takes place. Duplication of staff, plant and so on could be reduced, for example, if a given enterprise assumed prime responsibility for organizing courses.

*Functional
literacy
v.
primary
schooling*

Was functional literacy provided by EWLP cheaper than education furnished by alternative sources? Data allowing a reliable and comprehensive comparison between EWLP projects and other kinds of literacy in programme countries are, regrettably, unavailable. This leaves only the possibility of comparing functional literacy with equivalent amounts and kinds of primary schooling. The difficulty of such comparisons has been referred to in the discussion of EWLP result measurement. Despite this difficulty—and keeping in mind the very vulnerable and provisional nature of the following figures—it does seem possible to make at least an educated guess about comparative costs. This guess is based on the comparison of equivalent amounts (lengths) of equivalent kinds of education (basic training), and explicitly excludes contents and functions of literacy and schooling respectively that do not refer directly to basic training.

It was possible to make the comparison for eight EWLP countries. In only one (Sudan) does basic training furnished by primary schools seem to have been cheaper than that provided by the EWLP project. In the seven

others, literacy was cheaper by percentages that ranged (per candidate at the final test) from 55 per cent in Madagascar to 86 per cent in Algeria and 89 per cent in Mali, and (per successful new literate) from 2 per cent in Ethiopia and 28 per cent in Ecuador to 63 per cent in Iran and 85 per cent in the United Republic of Tanzania. In other words, it appears that for an identical expenditure EWLP was able in each case to provide basic training to a certain number of additional people. Keeping in mind the above criteria (cost per candidate at the final test in Algeria, Madagascar and Mali, and per successful new literate in Ecuador, Ethiopia and the United Republic of Tanzania), the additional numbers of learners trained by the programme compared to what schools could have done at the same cost, seem to have been roundly as follows: Ethiopia, 310; Ecuador, 2,010; Madagascar, 2,360; Iran, 16,600; Mali, 44,430; Algeria, 46,320; United Republic of Tanzania, 130,700.

In contrast, an identical expenditure on schooling and literacy in Sudan may be supposed to have enabled the school to provide basic training to 320 more pupils than the EWLP project.

It can certainly not be concluded from these figures that the Third World should immediately replace primary schooling with EWLP-type functional literacy, if only because of the possible margin of error affecting the collection and analysis of the figures themselves. On the other hand, this comparison does offer two conclusions that may be drawn with fair certainty. First, functional literacy as organized under EWLP does seem to have provided important aspects of fundamental education rather more cheaply than existing forms of primary schooling could have done. Secondly, cost-conscious educators could therefore usefully re-examine existing forms of primary schooling with a view to introducing certain elements of functional literacy. EWLP may not offer irrefutable evidence of the necessity to 'de-school' society, but it does strongly suggest the need (and a way) to 'de-school' the school.

*To de-
school or
not to de-
school*

Impacts on further educational action

EWLP did not, single-handedly, 'de-school' the schools, but there are signs that, here and there (and mostly in programme countries), it at least influenced them. In one case, the school seems to have had greater influence on the EWLP project than vice versa. Elsewhere the reverse may be true. Organizers of one project noted with satisfaction that school-teachers who doubled as literacy instructors began to display less authoritarian attitudes toward children. In two countries, adult education methods have been introduced into the curricula of teacher-training colleges. Iran is studying the possibility of reforming all training of school-teachers to make it more responsive to local reality, i.e. more functional.

EWLP's influence has also been felt outside the framework of formal schooling. One project was requested to assist in training community development instructors. More to the point, several EWLP countries have