evaluating development projects

[A.2684] $2.50; 15/- (£0.75) (stg.); 10 F

evaluating
development projects

technology and society

Technology and society

evaluating
development projects

a manual for the use of field workers
prepared for Unesco by
Samuel P. Hayes, Jr.

unesco

Foreword

One of the constructive aspects of the decade following the Second World War has been the launching of many projects designed to improve man's condition through planned programmes of social development. But although there are and have been hundreds of social development projects, relatively few have been analysed by techniques adequate for measuring their actual effectiveness or for guiding them towards increased effectiveness. The social sciences have produced simple measurement techniques that can be used in many such development projects, including rural situations. What is needed now is experience in applying such techniques and in adapting them to a variety of field situations.

In this volume, Dr. Hayes endeavours to demonstrate how certain social science measurements can be adapted to help field workers assess initial conditions before a project is begun; to measure the extent to which various attempts at producing social change have been successful; to determine the over-all result of social development schemes; and to identify the factors that are important in influencing the success of programmes of social change.

The Unesco and United Nations Secretariats have called on many social scientists, programme administrators and field technicians to assemble material on this subject. An issue of Unesco's *International*

Social Science Bulletin in 1955 was devoted to 'Evaluation techniques'. Dr. Hayes has taken account of this work, has solicited the criticisms of his colleagues in preparing this volume, and acknowledges their contribution. Dr. Hayes, himself, has outstanding qualifications for the task of preparing a manual designed specifically for the use of field project personnel, being both an academic social scientist, concerned largely with measurement methods, and an administrator and planner of social development projects.

From 1934 to 1943 and from 1945 to 1948, Dr. Hayes conducted social science research and taught the social sciences. From 1943 to 1945 and from 1948 to 1953, he was engaged in planning and directing programmes of economic and social rehabilitation and development in North Africa, Europe and South-East Asia. From 1953 to 1960 he was director of the Foundation for Research on Human Behavior, in Ann Arbor, Michigan, United States of America. From 1959 to 1962 he was professor of economics and director of the Center for Research on Economic Development at the University of Michigan. Since 1962 he has been president of the Foreign Policy Association, a private organization engaged throughout the United States in programmes of adult education in foreign policy.[1] Dr. Hayes is co-editor, with Dr. Rensis Likert, of *Some Applications of Behavioural Research*, published by Unesco in 1957.

1. Additions made in 1965.

Foreword

(Second edition, revised)

Since the first impression in 1959 of the present manual, interest in the evaluation of field development projects has steadily increased in all parts of the world. In accordance with 11 C/Resolution 3.52, the Director General is authorized, in co-operation with the United Nations, the Specialized Agencies and competent non-governmental organizations, '... (c) to continue to promote the development and application of techniques of survey and evaluation'. To this end two seminars were convened by Unesco at the Arab States Fundamental Education Centre (ASFEC) in Sirs-el-Layyan (United Arab Republic) (16-28 December 1961 and 16-20 December 1962) in order to examine the practical problems involved in the application of the methods of the manual to the evaluation of field projects by non-social scientists.

The seminars were organized with the valuable assistance of Drs. Hamed Ammar and Louis Kamel Meleiha, of ASFEC, and of representatives of ILO, FAO and WHO, as well as the National Commission of the U.A.R. for Unesco. Dr. J. D. N. Versluys, Chief, Division of Applied Social Sciences, represented the Secretariat.

The scientific direction of the two seminars was entrusted to Dr. A. J. Wichers, Agricultural University, Wageningen (Netherlands).

In preparation for the first seminar, some sixteen projects currently in operation in the U.A.R. (Egypt)

and Sudan in the fields of health, education and rural development had been selected in consultation with the National Commissions of these two countries. Officers in charge of these projects, belonging to the public administration of these two countries, and who in general were not social scientists, were invited to participate at both seminars.

The first seminar was devoted to the planning of the evaluation, and provided the participants with a certain amount of training and briefing: during the intervening twelve months, the project officers applied the methods and criteria of the manual to the evaluation of their respective projects, and prepared reports which were discussed at the second seminar and appeared to be very satisfactory in general.

In the light of the conclusions emerging from this exercise Dr. Wichers made detailed suggestions for the revision of some parts of the manual, which were transmitted to Professor Hayes for comment and approval.

Unesco wishes to thank the aforementioned persons for their most valuable co-operation and expresses the hope that the manual which was so well received on its first appearance will be an even better working instrument in its revised form.

5 March 1965

Contents

Preface

This book is an outgrowth of a conference held in Geneva in July 1954, under the joint sponsorship of the Department of Social Sciences of Unesco and the Technical Assistance Board of the United Nations.[1] That conference, in turn, grew out of the lively interest and early initiative of the United Nations in technical assistance for economic development, evidenced in its Expanded Programme of Technical Assistance. This programme was undertaken in pursuance of the General Assembly's resolution 304 (IV), adopted unanimously at its fourth session, in 1949, and constituting approval of the earlier resolution 222 (IX) of the Economic and Social Council. These resolutions not only directed the undertaking of the expanded programme but also the carrying out of 'a critical examination of the activities undertaken and the results achieved under the expanded programme....'

The United Nations and Specialized Agencies have, pursuant to these resolutions, carried out most energetically and fruitfully a wide variety of technical assistance activities, and they have also directed a good deal of attention to evaluating the results achieved. This interest in evaluation led to the Geneva conference referred to above, to several other conferences on this topic, to a series of papers on particular techniques of evaluation (in the *International Social Science Bulletin*, 1955, Vol. VII, No. 3) and to this manual, which was prepared under contract with Unesco's Department of Social Sciences as part of its programme of activities in evaluation.

The first edition of this manual was published by Unesco in 1959. In order to test out the practical applicability of the concepts and procedures presented in the first edition, Unesco convened two seminars at the Arab States Fundamental Education Centre in Sirs-el-Layyan (U.A.R.) in

1. See *Final Report*, Meeting on Criteria and Techniques of Evaluation of Technical Assistance for Economic Development, Geneva, July 1954, prepared by Louis Moss (Unesco document SS/12, Paris, 15 February 1955).

December 1961 and again in December 1962, under the scientific direction of Dr. A. J. Wichers, Agricultural University, Wageningen (Netherlands). Participants in the first seminar returned to their posts in Egypt and the Sudan and carried out twelve evaluation projects as part of their daily routine, using the original manual as a guide. Reports and discussion of their experiences took place at the second seminar.[1] The present revision takes into account the lessons learned during that evaluation of methods of evaluation.

While this revision was put in final form by the original author, extensive additions and changes in the original text were drafted by Dr. Wichers and were incorporated here. The author wishes to express his great appreciation for the important contributions by Dr. Wichers to the text of this revised edition.

It is always important to define the audience for whom a publication has been designed. This manual—especially after being revised in the light of the seminars and field testing described above—is directed to those who have some measure of responsibility in the making of policy and in its execution, but who are not too far from actual field operations. It is assumed that most of them will have had the equivalent of some college training. This is not to deny that those at intermediate levels of responsibility are important in evaluation activities, should have instruction, and might well benefit from this manual. But, as this is not generally the level where evaluation activities are initiated, it is not the primary audience for whom the manual has been written.

Permission to reproduce certain copyright material has been received from the authors and from the Dryden Press, *Economic Development and Cultural Change*, the Russell Sage Foundation, and the Netherlands Universities Foundation for International Co-operation. This permission is gratefully acknowledged.

December 1964   SAMUEL P. HAYES, Jr.
                                                 New York. N.Y.

1. See *Report on the Unesco Regional Technical Assistance Project on the Development of Evaluation Techniques* (Unesco/SS/39; Paris, 20 April 1964).

Introduction

Most nations of the world today give high priority to projects and programmes intended to advance economic and social development. Governments have accepted greater responsibilities and are using greater resources than ever before to aid development. Intergovernmental organizations and their agencies are increasingly concerned with development activities. Business concerns and private, non-profit organizations are more and more adopting economic and social development at home and abroad among their major goals, and are devoting substantial resources to help achieve these goals.

For economic and social development to take place, leading to improved levels and a heightened quality of living, any or many of a wide variety of changes may need to take place. The needed changes may be psychological —in motivation and attitudes. They may be educational—in knowledge, concepts and skills. They may be technological—in new methods of production, organization, administration and distribution. They may be sociological—in relations among persons and groups. They may be economic —in the allocation of resources to various industries, in the distribution of income, in the denial of present consumption to permit increased investment. They may be political—in the control and exercise of power.

The tasks of the development administration planner are threefold: to determine the changes in his country which most need to be speeded up, to design actions or activities (development projects) by government or other organizations which give most promise of bringing about the desired changes with the highest ratio of benefits to costs, and to administer projects (including altering them) as efficiently as possible (which may require altering them).

All three tasks depend for their success upon the collection, analysis and evaluation of pertinent data. The more objective and the more

quantitative the data, the better. As Lord Kelvin pointed out, 'When you cannot measure what you are speaking about, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarely in your thoughts advanced to the stage of science, whatever the matter may be.'

It is especially important to obtain' information about the changes that specific projects have produced and about the costs they have incurred. And the more directly this information can be related to the various elements in actual project operations, the more convincing and useful its testimony. The art of inducing and guiding social and economic change is not yet highly developed. Data that show how things are going can be immediately helpful in improving project operations in midstream. And what can be learned about the results of completed projects can be most helpful in planning subsequent projects.

A good deal of this manual is therefore devoted to methods of collecting and analysing data of a psychological, educational, sociological, economic or political character, which can be used in planning and in evaluating development projects. The methods described are, essentially, the methods of the social sciences, applied to the practical problems of development administration. They have been successfully used by many who, though not academic social scientists and not having large resources available, have nevertheless wished to undertake careful, objective investigations. Such persons will find very useful the practical examples of data collection and analysis reported in the *International Social Science Bulletin*, 1955, Vol. VII, No. 3, on 'Evaluation techniques'.

Changes significant in development

Economic and social development is a complex process involving many kinds of change. Some changes occur early, others later. There is a high degree of inter-relationship among the various changes that occur, but it is usually very difficult to discover how, and how greatly, any particular changes affect other changes.

Both logical analysis and empirical investigation, however, point to certain kinds of change as being of strategic importance during the period when a traditional social system is transforming itself into one characterized by a self-sustaining process of technological and material advance. These strategic kinds of change may be grouped under three major headings, as they affect primarily individuals, social relationships and social capital.[1]

1. See E. E. Hagen, 'The process of economic development', in *Economic Development and Cultural Change*, 1957, Vol. 5, pp. 193-215, from which much of this listing was taken.

Changes in individuals: their information, skills, attitudes

1. Increasing literacy, scientific knowledge, scientific and technological training and a scientific or engineering approach to problems.
2. Increasing dissatisfaction with traditional levels of living, traditional status relationships, traditional economic activities.
3. Increasing belief that economic and social advancement can be obtained through new devices and new economic activities.
4. Increasing belief that economic and social advancement can be obtained through individual competence and effort (rather than through preference based on bribery, political favour, kinship, caste or social status, national origin, race, religion or sex).
5. Increasingly specific definition of economic rights expected from society, and economic obligations to society (rather than the vague, open-ended security often expected from the village or family, and the indefinite but broad economic obligations often felt towards the village or family).
6. Increasing respect for honesty in business and government and for the use of contracts in economic relationships.

Changes in social relationships and institutions

1. Increasing interpersonal communication (through adoption of common language, growth of literacy, increased media of communication, social mobility, travel, etc.).
2. Increasing economic opportunity (through decreased monopoly, increased availability of credit and greater occupational mobility).
3. Increasing rewards for economic activity (through land reform, tax changes, governmental aid of various kinds).
4. Increasing power of groups participating in these changes, and diminishing power of groups resisting change.
5. Increasing governmental activity in public service and in economic and social development.

Changes in social overhead capital

1. Increasing investment in education, public health and sanitation, water supply.
2. Increasing investment in transportation, communications, power, irrigation.
3. Increasing competence (likely to involve investment) in public administration, civil police, and the military establishment.

Measurement and evaluation

Many development projects are intended to bring about increases in the production or distribution of goods or services. Such increases may be temporary, ceasing with the end of the project. Well-designed projects do not limit their goals to changes that may be temporary, but seek also to bring about changes of the strategic kinds listed above, which are likely to persist for long periods of time. Both kinds of changes can be measured.

But obtaining measurements of project results and costs is not in itself 'evaluation', even though it provides information needed for evaluation. Measurement, i.e., concretizing and quantifying phenomena as far as possible, is only one step in evaluation. Contrary to the popular saying, facts seldom 'speak for themselves'. They have to be interpreted, and interpretation is the component that distinguishes evaluation from mere measurement.

For example, a literacy rate of 30 per cent may be judged 'very satisfactory' at present, but 'very poor' ten years hence. This means that the mere fact of 30 per cent literacy has to be placed in a context involving judgements about what is desirable and what is attainable. This example also illustrates the fact that standards may shift from one period to another.

Unfortunately, situations are rather rare where contexts and standards are fixed and clear without any interpretation being required. It may even be very difficult to find a good basis for appraisal or a satisfactory measuring device at all, especially with activities which are intended to bring about psychological changes.

Therefore evaluation is firstly—and necessarily—a body of concepts and practices which have proved their usefulness (in the field of social sciences) and which are applied in such a way that they can contribute to the improvement of practical activities. Thus a manual useful in evaluation could hardly be a timetable, a checklist or some other simple device. Rather, it needs to present applicable concepts and methods, as this manual does. The more experienced an evaluator is, however, the quicker will he lay hands on proper standards and devices and be in a position to help the development administrator.

First-approach and socio-psychological evaluation.

Closely connected to a correct understanding of evaluation is the distinction made in this heading, which may be illustrated by an example.

In a particular development project, fertilizer may be provided (at a subsidized price). The increased harvest may be looked upon as the primary measure of accomplishment of the project. A more fundamental measure

of accomplishment is the change in the willingness of the farmers to buy fertilizer at commercial prices in subsequent years. Still more fundamental is the change in their interest in trying out other improved agricultural techniques. Most fundamental of all is the change (if any) in their willingness to support, or to co-operate in, the development of new agricultural techniques.

In appraising economic development, attention has often been focused on superficial changes in production, consumption or saving, perhaps because they are relatively easy to measure. But significant changes are the ones that occur more profoundly in the minds of people. Often—especially when some prediction of future conduct is involved—it is only there that such changes can be investigated.

The more an evaluation remains at the first level in the example (i.e., using a simple but rather superficial criterion), the more it is first-approach evaluation. The more it comes to the last level in the example, the more it is socio-psychological evaluation—because it implies the use of socio-psychological procedures. The art of evaluation is to utilize first approaches whenever possible, but to go on to the socio-psychological level when necessary. It is possible to obtain data on any required level. Methods of collection are discussed under 'Step One' below.

Programme results and project results

This manual discusses ways of measuring and analysing both the results of development projects and the results of the individual operations comprising those projects. It does not, however, discuss the results of development *programmes*.

Development programmes usually comprise a wide variety of important activities, in different regions, over different time periods. Measuring the results of development programmes is, of course, very important. Governments cannot determine their policies, allocate their funds or administer their agencies without information about the results obtained. To measure these results, they call upon the whole statistical apparatus of censuses, records of production and trade, index numbers, input-output analyses, national income accounting, etc. Such over-all measurements and analyses are outside the scope of this book.

Uncovering the results of individual projects is simpler and less costly and can often be carried on by the persons directly involved in the project. Moreover, project results show up most clearly at the 'level' where a particular group of individuals comes directly in contact with particular project operations. The fewer the complicating factors present, the clearer will be the link between cause and effect. As governments improve their methods

of collecting data for use in evaluating their development programmes, however, a valuable by-product may be the increased availability of data which can then also be used in evaluating the results of particular development projects.

The fact that national development programmes are left out of consideration here does not, however, mean that this book is intended only for use with extremely simple projects. Nor is it intended only for projects for which a first-approach evaluation is undertaken. Even a limited and clearly defined development activity may present complicated aspects and may require socio-psychological investigation in order to provide administrators with the insights they need for evaluation and decision.

Steps in indicating the results of a project

The logical process of identifying (and measuring) the changes that occur and then attributing them to a development project may be outlined as follows:

Step one. Describing the development project and specifying the changes which are its goals.

Step two. Deciding what data to use to indicate changes in the directions desired (and also to indicate side-effects).

Step three. Collecting the data—before, during and after.

Step four. Analysing and interpreting the findings, and reviewing these with interested groups.

All the steps but the third are relatively simple ones—dictated by logic and demanding only the time and effort to think them through and reach agreement. These steps are little different in small, inexpensive projects from what they are in large, elaborate ones. They can be carried out without requiring additional staff or funds.

'Step three' can be done simply, without additional staff or funds. On the other hand, if the project is of such size or significance as to warrant it, this step can be developed in as great depth and detail as desired. The bibliography at the end of the manual lists references for use in carrying out such more intensive studies, together with reports of a number of field studies carried out in Asia and Latin America. It may be that a centre for social research exists in a local university or in some government department and can be called upon to carry out a detailed, scientific study of the changes induced by a project.[1] Here it is assumed that little if any additional

1. An illustration of the excellent social research that such centres may be willing to carry out is contained in A.F.A. Husain's *Human and Social Impact of Technological Change in Pakistan*, Dacca, Oxford University Press, 1956 (two volumes).

staff or funds are available to obtain and analyse data about project results. The procedures recommended are believed to be practicable for most operating agencies.

This manual a closed circle

At its optimum, a completed evaluation which followed these steps would conclude that a particular development project had contributed to one or more of the 'changes significant in development' which are listed above. Often, however, an evaluation—especially when it applies to a limited activity and is done without much professional sophistication—is inadequate to demonstrate convincingly that such a contribution has been made. Still, interpretation of the demonstrated results in relation to the more fundamental changes needed in a developing society is a valuable exercise. An example of such interpretation is given at the end of this manual, thus closing the circle by emphasizing the relationship of even the most limited development activities to the larger goals sought.

This relationship can well be kept in mind while planning an evaluation that proceeds by the steps now to be described. Even if the project is limited and methods of investigation must be simple, there is great value in having a broad view of its context and making a logical evaluation plan. Moreover, rigorous objectivity in handling data is a necessity. Each well-conducted evaluation can make an important contribution to the understanding and forwarding of a nation's economic and social development.

Step one   Describing the development project and specifying its goals

Before one can discuss results, it is obviously helpful, if not indeed a logical necessity, to ask 'The results of what?' Implied are questions both about what is done in a project and about why the project was undertaken—what were intended to be its results. The answer to these questions will then suggest what data will be germane as indicators of project results. The more detailed the statement of operations and objectives, the wider the range of possible indicators suggested.

Describing the project

What aspect of economic and social development is involved? Agricultural technology? Industrial production? Education? Health? Or what?

What organization or individual is responsible for the project?

How many and what kinds of people are engaged in it? What supplies, facilities and equipment are available for use in the project? What and whose funds are allocated to it?

Where will the project be located? For how long will it last?

The answers to such questions give the dimensions of the project. Project administration and evaluation are both facilitated if these dimensions are specified in writing before the project is initiated, or very soon after activities have begun.

Some dimensions may change during the project's life. If an explicit, dated record of such changes is kept, it will be most helpful in later interpretation of project results. This all sounds very simple. But one will be astonished how often it occurs that one or more of these questions cannot be answered (readily or at all).

Specifying the operations involved

Selection of indicators and interpretation of results are greatly assisted by detailed reports of the actual operations performed—by whom, what, for and to whom, where, when, and for how long? It is usually desirable to include at least as much detail as in the following examples:

Advice by (named) experts to (named) government officials in particular government ministries, concerning particular (named) problems.

Participation of (named) expert as voting member in development planning board, during the preparation of a development plan for the petroleum industry.

Drafting legislation for the organization and establishment of a central bank, working in consultation with the ministry of finance.

Training of a particular group of teachers in improved teaching methods, this training to include lectures (in what languages?), discussion groups (with interpreters?), specified reading, supervision of assignments and practice teaching sessions.

Development and presentation of a series of radio programmes of specified length, at specified intervals, etc.

Distribution of anti-malarial drugs at specified intervals to all families (estimated number) in a particular area (specified).

It is helpful if the exact operations which are to be undertaken are specified as early in a project as possible, ideally before the project itself is approved. This makes possible early collection of data about the groups which will be the objects of the operations.

It is not always possible, however, to specify all operations before the experts or administrators involved are on the ground and have a chance to study the situation in some detail. Moreover, changes may subsequently need to be made as particular activities turn out to be more or less effective than expected. If operations are specified in writing as soon as decisions are made, however, collection of data can begin early and will make possible later comparisons. For example, assume that, during a home economics educational programme, it is decided to stop distributing information pamphlets directly to private homes and, instead, to train local school-teachers who are then expected to pass the new information on to house-wives. There are now two sources of data on the progress of the project: namely, the teachers (whose knowledge and communication activities can be studied) and the housewives. The earlier that this change in operational strategy is known, the earlier the collection of pertinent data can begin.

Stating objectives

In deciding what kinds of indicators are needed to show results, it is obviously helpful to state with some precision just what the project is intended to accomplish. This then directs attention to the kinds of data that will indicate extent of accomplishment. Because objective, quantitative indicators carry most conviction, it is usually desirable to point up those changes that are objectively measurable by methods that can be duplicated by others. If a statement of desired changes is formulated at the beginning of a project, it will be possible to obtain more 'baseline data' as the first step in 'before and after' measurement.

The results of development projects—the standards by which they are judged successes or failures—may be considered in terms of six major dimensions:
1. The kinds of results sought (including their permanence and spread).
2. The side effects produced.
3. The area of activity and the area of influence where these results appear.
4. The time period within which results of specified size are obtained—the pace of change.
5. The cost of bringing about results of specified size—the efficiency of the development project.
6. The amounts of desired results brought about by factors other than the development project.

These dimensions are discussed separately below.

Definition of kinds of results sought

The major kinds of results usually sought through developmental projects may be classified as follows:

Impersonal results. For example, legislation is drafted, the remedy for a disease is discovered, a survey is completed.

Administrative results. For example, a formal organization or institution is established, enlarged, reorganized; policies or procedures are changed; budgets are altered.

Results involving persons

1. Changes in information, attitudes and skills (for example, X per cent know the relation of mosquito to malaria; Y per cent are willing to have their cattle inoculated; Z per cent can read material of designated difficulty; there are 75 accomplished X-ray technicians).

2. Actions and habits of individuals or groups (for example, X per cent take anti-malarial drugs regularly; Y per cent take their cattle to be inoculated; Z per cent go to parent-teacher meetings; a mobile X-ray machine visits specified villages once a year).
3. Changes in social structure or relationships (for example, lower caste persons participate in activities previously reserved for higher castes; the recommendations of a parent-teacher association are taken into account in changing a school curriculum; the village headman is no longer the sole important source of news).

Changes in economic and social conditions, resulting from actions taken by individuals or groups. For example, rice production is up by X per cent; malarial spleen rate is down by Y per cent; the cattle population is larger and healthier; employment expands; the occupational distribution changes; the consumption pattern shifts; the price index is stable.

Changes in capital resources. For example, a bridge is built; savings and investment increase; oil is discovered.

Selecting the changes to be measured. Under the heading, 'Changes significant in development', page 14, a list was given of changes believed to be especially significant for economic and social development. On the same page attention was directed to the different levels of investigation and evaluation, from 'first-approach evaluation' to 'socio-psychological evaluation'. An important aspect of the 'art' of evaluation is selecting the particular changes to be studied, in the light of the significance of the changes and the availability and reliability of data that can be used to measure those changes.

Some data may be available but unreliable for the purpose desired. For example, assume that a campaign has been undertaken to promote the use of chemical fertilizer in rice culture. Data may be available concerning the year-to-year changes in amounts of rice marketed, but these data may be considered unreliable in judging the effect of the fertilizer campaign, because the amount of rice marketed is affected also by weather variations and by irregular marketing. Much more reliable data can be obtained by examining records of fertilizer sales or by visiting a sample of farms to find out how many additional farmers are using fertilizer (and how much is being applied per hectare). Here, the project would be evaluated in terms of fertilizer sales and utilization, rather than in terms of rice marketed.

Some changes have a more fundamental significance than others. For example, assume that the object of a development project is to mobilize

village labour in co-operative, self-help activities. To judge the success of the project, one might obtain information over a period of time concerning: (a) the number of persons taking part in such activities; or (b) the number taking a share of the responsibility of running such activities; or (c) the number taking the initiative in interesting others in such activities. An increase in actions of the third type represents a much more significant change than a similar increase in actions of the first type.

Some changes may prove very difficult to measure, and ingenuity may be required to identify other but related changes which are easier to measure. For example, it may be impossible, without exorbitant expense, to seek out a random sample of the graduates of a school of industrial arts, two years after graduation, and ask them how good a training the school had provided and how well it had inculcated the idea of the dignity and worth of mechanical work as a vocation. Data indicating the information and skills of graduates can, however, be obtained at the time of graduation. Data on the attitudes and competence developed by the school can subsequently be obtained by asking new applicants for training if they have been sent there by former students or by the employers of former students.

Some changes are both very significant and reasonably easy to measure, but the evaluator may wish to get below the surface and find out how these changes came about. Knowledge of the intervening variables—of the information, skills and attitudes without which the actions would not have taken place—may be of great value in planning and administering other development projects.

Sometimes, the evaluator will wish to study changes that indicate long-term results.

The immediate, concrete results of a development project are often less significant than later, less tangible results. A bridge built, some cattle inoculated this year, a health mystery solved, a school of industrial arts established —these are not always considered the ultimate tests of the success of development projects. The ultimate test to be applied may be whether or not the project generates its own continuing and expanding activity. Have the new methods and training been incorporated into the normal life of individuals and communities, and do they persist and spread? Can the area now build its own bridges, and will it continue to do so? Will inoculations of cattle spread to the whole country, and perhaps lead to other improved practices of animal husbandry? Have methods of scientific investigation been inculcated, and the desire to use them whetted, so that health mysteries of new kinds are solved without new special projects? Will the new school attract good students and will it continue and expand, even when the government falls and a new minister of education takes office? Have the people of the area generally come to believe in progress and in their own capacity to bring it

about? These are often felt to be the real criteria of success.[1] Yet data of this kind may not be available for years after the end of the development project, and it will become increasingly difficult to disentangle the effects of a particular project from the effects of the myriad other changes going on in the meantime.

It is possible, however, to throw a good deal of light on the depth and enduring character of the project's impact, and on the likelihood that its effects will continue and expand, by determining people's attitudes during and at the end of the project. These attitudes can be uncovered, usually by using the interview techniques described in 'Step two' under 'Asking people about projects and results', page 41 (and described in more detail in the references listed in the bibliography). Attitudinal information obtained verbally must, of course, be used with care. People may say they are willing, for example, to be vaccinated, but may slip away when the time comes. There are various ways of getting below the surface of such verbal responses.

Another significant measure of the likely permanence of 'project-induced' changes is the extent to which project activities have been institutionalized. Has an appropriate organizational unit been established and been assigned competent staff, with a regular budget, to administer and expand this activity? Have communities incorporated this activity into their own local institutions? Do individuals practise these new methods within informal social or work groups? Experience in many areas has shown that creating or adapting institutions to provide administrative and technical machinery, capable of implementing new techniques and of encouraging the continuing growth of new ideas, is fully as important as providing technical advice and development assistance to the initial project.

Defining the results sought from each separate operation. Many development projects consist of a single operation each. Others, however, comprise several. For example, a campaign to promote the use of fertilizer may include subsidizing the price, training extension agents, teaching adults who are in literacy classes, setting up demonstration plots, and broadcasting radio programmes. These are separate operations, and each can to some extent be evaluated separately. If there is to be such separate evaluation, it is obviously desirable that some explicit advance statement be made that specifies the effects expected to result from each distinct operation. If this is accompanied by the rationale of these expectations, including the assumptions on which they are based, it will be easier to select data to measure the effects of specific operations. These effects may include not only the intended

---

1. See Morris E. Opler, *Social Aspects of Technical Assistance in Operation*, Paris, Unesco, 1954, especially pp. 69-70.

direct effects but also expected indirect effects, both desirable and possibly undesirable. At the same time this takes us to a last aspect of evaluation to be mentioned here.

The intention of the project must be kept in mind. For example, if i tis expected that the distribution of pamphlets, as part of a health extension programme, will simply spread information, this suggests obtaining one kind of data on results. If the pamphlets are expected to influence attitudes, or to bring about actions, a different kind of data are needed.

Thus, if advance information is provided concerning the separate operations to be performed, and concerning the changes expected to result from each of these operations, plans for collecting data can be set up that take fully into account the entire range of expectations of the project, and that include collection of data on the relations of the expected effects to the exact operations performed.

Side effects

It frequently happens that development projects produce unanticipated side effects, desirable and undesirable. Because the effects were unanticipated, baseline data may not have been obtained, or may not be already available from existing records, to permit 'before and after' comparisons. It may therefore be difficult to demonstrate convincingly that these apparent effects did in fact result from the project.

On the other hand, unanticipated effects are sometimes so important for administration and for subsequent planning that it may be desirable to make special efforts to determine their relationships to project operations. This may involve interviews after the fact with some members of the groups affected and with informed observers, in which they are asked to reconstruct from memory the situation existing before the project, and to trace in terms of their own remembered experience or observations the apparent relationship between the project operations and the unanticipated effects. It may be desirable to study data which were recorded before the project began and which throw light on the earlier situation, if such data are available. Even if these records are incomplete or provide only indirect evidence about the factors being studied, they may be very helpful in suggesting possible relationships between project operations and important changes apparently resulting from these operations. Similarly, comparison of the project area with a 'control area' may throw light on the possible cause of side effects.[1]

Not all side effects can be anticipated, of course, but many can, especially if there is early consultation with social scientists and others familiar with the

1. See 'Step Four: Analysing and interpreting findings', page 81.

people and culture where the project is to be carried out. If attention is given to prediction of possible side effects, this will: (a) permit planning of stand-by measures to be taken should undesirable side effects appear, or permit planning other measures designed to maximize desirable side effects; (b) help experts and administrators to be on the alert for indications that side effects are being produced. If they know where to look and for what, they can make the modifications needed at an early stage; (c) permit plans to be made for the collection of data on these side effects, for the guidance of experts and administrators.

The area of indirect effects is one where periodic collection of data throughout the period of project operation may be particularly valuable. If important indirect effects begin to appear, experts and administrators may wish to arrange for collection of data bearing directly on the extent of these effects and on their relationships to project operations. These new data can then be used to modify project activities promptly and in this way to maximize the benefits and minimize the harm from unanticipated effects of the project.

Definition of area of activity and area of influence

A development project has not only a locus of activity and a group of people who are touched by it directly. A project usually also has an 'area of influence', which includes all who are significantly affected by it. In order to decide what indicators to use to show the results of projects which are intended to influence people (i.e., most projects), it is usually desirable to specify: (a) the particular persons or groups on whom the project operations are to be performed (the 'area of activity'); (b) the total population on which the project is expected to have some discernible effect, and which is therefore also to be studied to uncover the magnitude of results of the project (the 'area of influence').

The most significant and the most easily discovered effects of a project are the direct effects within the 'area of activity', and, if resources are limited, collection of data will usually be most rewarding there. In planning collection of such data, it should first be specified what class of individuals are to be reached directly by the operations which are to be performed. The 'area of activity' might, for example, be 'rice farmers who live along main or secondary roads in Area A, and who planted during April 1955', if these farmers do in fact comprise the group with which the project will have some direct contact.

Indirect effects of the project could appropriately be studied in the whole 'area of influence', defined to include all rice farmers who might have been influenced directly or indirectly, by the project activities. These might be 'all farmers in Area A'.

27

If, in drawing up the original description of a development project, the 'area of activity' and the 'area of influence' are clearly defined, this will help both in planning the project and in deciding on the kinds of data to be collected as indicators of project results.

The pace of change

A full statement of the objectives of a development project might well include some indication of the time period within which specified amounts of the desired changes are expected to come about. It makes a great deal of difference whether substantial results are sought in a short time, or over a long period. This is important both in planning what is to be done and in evaluating results. If review of objectives, including pace of change, indicates that the operations planned are not believed likely to bring about changes of the magnitude desired within the period specified, it may be appropriate to plan additional operations or to reconsider the desirability of undertaking any.

The cost of the results obtained

The cost of a development project is itself one of the results of a project. Just as the unintended, undesirable side effects of a project must be considered, so also must its financial cost be weighed against its benefits. This appraisal usually includes both: (a) whether the project is achieving its objectives at the lowest possible cost, and (b) whether the benefits are recognized by the responsible organization and by the individuals affected by the project as greater than the costs.

A particular development project may be considered a key activity, necessary for continued progress in public health, education or agriculture. But there is a point beyond which it may be preferable to turn to more economical methods of accomplishing the same objectives. Could a different project be designed to yield the same effect, but at less cost?

For example, the desirability of training sanitarians may be fully recognized. But which would be cheaper and more effective—to establish a local training course with imported trainers, or to send the required number of trainees abroad?

'Costs', like 'results', are of many kinds. For example, a forestry demonstration might indicate that preparing the land for planting seedlings 'costs' less if bulldozers are used than if the work is done by hand. This calculation should also weigh the foreign exchange 'costs' involved in the purchase and operation of mechanical equipment. In areas of underemployment, it should also weigh the social 'cost' of not employing available labour.

Many development projects are pilot projects, intended to demonstrate

what is possible and feasible, with the expectation that the improved techniques demonstrated will be adopted on a much larger scale by organizations or individuals. An initial subsidy to the demonstration is often necessary to encourage participation and to avoid having the risk of failure fall on the participants. After the demonstration period, however, and outside the demonstration area, the new method can hardly be expected to be utilized unless its cost is within the financial capacity of the individual (or the community or government), and is not so large as to outweigh the added cash income and the improved well-being and any intangible benefits that may result.

It is desirable that early project plans include cost estimates in relation to anticipated achievements. In addition to rough estimates of expected total expenditure, there might well be a budget giving the amounts planned for major classes of expenditure (for example, international personnel, counterpart personnel, other host country personnel, travel, supplies, etc.). Simply stating realistic probable costs may force reconsideration of a project before it starts. Going into a project with open eyes also reduces the possibility of sudden withdrawal of support. Moreover, comparing later actual expenditures with initial cost estimates may be very revealing when seeking reasons for project success or failure.

Results of non-project factors

It may well happen that some of the results of the kinds sought through the development project are in fact produced by other factors. Attention to this possibility at an early stage is necessary to make sure that pertinent data are collected in time.

Who should describe the project and set its goals ?

The value of early formulation of an explicit and detailed statement about the nature of the project and its objectives has been suggested above. But before these dimensions and objectives can be stated, someone must decide what they are. Unfortunately, such decisions are sometimes not made; if made, they may not be explicitly stated; and if stated, they may be stated differently by different agencies or persons concerned with a given project.

There are frequently several interests involved in a single development project, the interest of the responsible organization (government agency, business enterprise, private institution, etc.) being, in theory, paramount and usually also in practice. If intergovernmental organizations or foreign governments or private foundations also participate in the project, they may be assumed to have their objectives to achieve, and these may or may not be

identical with those of the responsible organization. Moreover, individual experts, although nominally and usually governed by the goals of employing organizations, sometimes have their own and different objectives (such as testing out a pet theory, or demonstrating the superiority of a particular method or product). The people likely to be affected by a development project also have their own goals, which may or may not be reflected in the policies of their government. A government may change, too, or a new official be appointed during the life of a project, and as a result different goals may be established.

Unless the objectives of a development project are stated explicitly and explained to the persons involved, people having different interests may perceive the project quite differently and as serving different goals. This can, of course, lead to serious misunderstandings, to working at cross-purposes, and often to failure of the project from the point of view of at least one of the parties interested.

The success of a project can be furthered if one person or office takes responsibility for developing an explicit statement of the nature and objectives of the project, to be accepted and implemented by all parties interested. This statement is also useful in deciding what data to collect as indicators of project results. Unless responsibility for working out such a statement is clearly assigned to, and accepted by, one office or person, however, it is quite likely to fall by the wayside.

This agreed statement can best arise out of early discussions among all interested parties. Here a kind of circular process takes place. Given certain statements of objectives, certain indicators may appear appropriate. Yet, when these are suggested as indicators of project results, the need for different or fuller statements of objectives may become evident.

The resulting written memorandum, specifying project nature, operations, objectives and indicators of results should be available to all interested parties, including the public. The difficulties and time likely to be involved in working out such a memorandum should not be understated, however. A protracted period of successive approximation and clarification may be necessary. The arrival of experts from abroad may help this process along, when they make suggestions from their experience and ask questions to clarify their understanding of their own roles. Moreover, as the project gets under way, as additional information is obtained, and as discussions about ends and means take place, clarification and modification of objectives inevitably occur. Attempting to formulate a written memorandum helps this process along, and insistence that such a memorandum be developed as early as possible, despite the difficulties involved, can make a major contribution both to project administration and to the planning of data collection intended to help measure project results.

Deciding what data to use to indicate project results

This section is limited to types of data which are already being collected, or which could be collected without requiring substantial additional staff or funds. Those organizations which can afford substantial additional funds and personnel for collecting and analysing data may wish to consult the more detailed discussions of methods of observation, interviewing, question-naires, content analysis, and other procedures given in the references listed in the bibliography. The major types of data discussed here are: (a) already available data (registration and census data); (b) additional records that organizations or individuals might agree to keep; (c) data obtained by systematic observation of behaviour; (d) data obtained by systematic questioning.

Even though this section is limited to methods of data collection that do not require research experts and substantial appropriations, it may at first seem formidable. This is because it attempts to suggest a wide variety of methods, and this might be interpreted to mean that all methods should be employed in connexion with every project. Such, of course, is not the intention.

Few, if any, projects would justify collecting all the kinds of data described here. It is hoped, however, that those responsible for even minor projects will find here feasible suggestions for improving or extending the collection of data to indicate the extent of their accomplishments and the reasons therefore.

Staff time required. Although it is not suggested here that great time and effort are required for the collection and analysis of data along the lines suggested below, it is obvious that one cannot get something for nothing. Even when some other agency has already collected the needed data, time is involved in obtaining the data for evaluation purposes, in examining their

usefulness, and in rearranging and interpreting them for a specific purpose. If direct observation or interviewing is necessary to obtain data on project results, a good deal more time is required. For this reason, it is desirable to plan the whole evaluation procedure in advance. Enough time must be allowed for both collection and analysis, and all pertinent data must reach the stage of analysis and interpretation. Usually, the data collection alone will take a month or more of the time of at least one member of the staff. Even more is required when the staff member is inexperienced and has little guidance.

Often the time of senior persons is required to consider the whole plan and decide what should be done and how much time can be made available to do it.

Using available data as indicators

A great body of statistical data about human populations and behaviour is collected by government, business and private agencies.[1] Even in countries where statistical services are not well developed, there are likely to be substantial resources of data that can be tapped to measure project results, yet these are often ignored. The principal limitation of such data is that definitions of their scope, method of collection, etc., are determined by agencies which have their own purposes to serve, and it is often difficult to sort out the data needed or to arrange for modification of collection so as to provide information on project results. Yet such data provide unique records of past conditions and social situations, and they may also throw considerable light on current conditions and social situations which are otherwise difficult or expensive to measure.

There are two main categories of such data: registration data and census data.

Registration data. These consist of records made at the time of occurrence of an event, in accordance with legal or administrative regulations, business practice, or other rules. Illustrative (but by no means exhaustive) of the activities often covered by available registration data are:
1. Vital events: births, deaths, marriages, divorces.
2. Illness: hospital and insurance data, incidence of registrable diseases.
3. Education: school attendance, grades, numbers of graduates.
4. Fiscal: tax collections of different kinds, social security payments and benefits.

1. Cf. *Report on International Definition and Measurement of Standards and Levels of Living.* UN.E CN. 3'179. United Nations, March 1954.

5. Crime: crimes known to police, arrests, court actions, prison records, parole records.
6. Draft and army service.
7. Business activity: payrolls, production records, absentee records, sales (total, and of individual commodities), bank deposits, cheque clearances.
8. Transportation: automobile registrations, bridge and road toll receipts, freight car loadings.
9. Formal organizations: membership, office-holding, committee participation.

In recording many of the above activities, collateral information is often obtained. For example, records of school enrolment sometimes include data on the nativity of the parents, occupation of the father, place of previous residence, results of psychological tests, school marks, etc.

Census data. These are usually collected periodically by canvassing from house-to-house, school-to-school, business-to-business, etc. Sample censuses are being increasingly used to supplement or elaborate complete censuses, because the former cost less and can be more rapidly processed for use, and because it is easier to train and supervise a limited number of investigators, and hence to ensure the accuracy and completeness of the information obtained.

Existing registration and census data are often fragmentary, however, particularly in poorer countries. The extent of possible error is indicated in Table 1, in which the accuracy of a health centre's records of infant deaths

TABLE 1. Accuracy of infant death reporting in five Egyptian villages during 1950[1]

Village	Deaths under one year reported to the health centre	Actual deaths under one year of age (as shown by canvass)	Degree of under-reporting at health centre (percentage)
Sindbis	19	23	17
Quaranfil	5	24	79
Barada	26	60	57
Aghour El Sughra	10	36	72
Aghour El Kubra	24	›95	75
	—	—	—
All villages .	84	238	65

1. Source: Weir, Wasif, Hassan Salah el Din Moh Kader, 'An Evaluation of Health and Sanitation in Egyptian Villages', *Journal of the Egyptian Public Health Association,* 1952. Adapted from page 404 of the article by Louis Moss, 'The Evaluation of Fundamental Education', *International Social Science Bulletin,* 1955, Vol. VII, No. 3, pp. 398-417.

(registration data) were tested by making a complete house-to-house canvass.

There are other examples. In many countries, income statistics based on individual tax declarations tend to be underestimated. Similarly, statistics concerning illegitimacy are less accurate than those on births in wedlock. Anyone using such data should consider whether the respondents can be expected to give the information requested (both able to give it, and motivated to do so) and whether the recorder has a strong motivation to secure accurate information. The user of such data should know something of the special definitions and contexts that may affect the nature of such data as he may wish to use. Moreover, registration data and census data are often not collected according to the same method in different countries, and sometimes not collected consistently at different places within a single country. Finally, if radical changes in the method of collection by the collecting agency, or in definitions used, occur during the period studied, the statistics obtained may be worse than useless.

However, the probability of errors does not invalidate statistics based on records. Especially if there are only minor changes in the type and frequency of error during the period studied, recorded data may be quite reliable indicators over time. They can thus be of great value as before-and-after measurements of the effects of development projects which affect a given community or group of communities.

A first step in deciding what data to use to measure the results of development projects is to find out what potentially useful data have already been collected or are continuously or periodically being collected. This requires visiting various organizations, officials and professional people and learning what kinds of records they keep. Courts, schools, churches, police stations, agricultural and health agents, town clerks, hospitals, doctors, undertakers, business concerns, political parties, farm organizations, youth organizations, and women's organizations are all likely sources of data. Selectivity will, of course, be necessary in view of the limited time that can be devoted to gathering data.

Such data as are available may be used directly as indicators of project results. They may also be used to corroborate (or contradict) reports obtained in interviews or published in the press. In an interview 'study of the psychological impact of long-term unemployment', for example, unemployed men said 'that they felt much worse at the onset of the unemployment than after three years', in spite of the fact that their economic condition actually continued to deteriorate throughout the three years.[1] This 'shock' effect of

1. Marie Jahoda, Morton Deutsch and Stuart W. Cook, *Research Methods in Social Relations*, New York, Dryden Press, 1951. Examples taken from M. Jahoda-Lazarfeld and H. Zeizel, *Die Arbeitslosen von Marienthal*, Leipzig, Hirzel, 1932.

unemployment found confirmation in examination 'of such records as the accounts of the local grocer, which showed a sudden drop in sales in the months immediately after the onset of unemployment, followed by a slight recovery and a steady decline thereafter'. Similarly, a record of fertilizer sales might be obtained to check against other information on the effects of a project to increase the use of fertilizer.

A careful thinking-through of the possible direct and indirect consequences of a developmental project may well suggest the kinds of available data that can be used to indicate its results and the ways in which such data can most tellingly be used.

Such data may also be useful even where effects have not been anticipated, and in consequence arrangements to collect baseline data have not been made at the beginning of the project. In such cases, records obtained for other purposes may be examined to uncover conditions and relationships which existed before the project was initiated.

For example, an unexpected result of the introduction of hybrid corn in a certain district was a substantial increase in the consumption of imported wheat flour in that district. It was discovered that shop-keepers there regularly kept monthly records of the sales of flour which could be consulted for both the project district and a control district for the whole period since the hybrid corn project began. Analysis of these records indicated that the high productivity of the hybrid corn fields had made it economically possible to buy imported wheat flour, which was ordinarily bought by preference whenever cash income was high enough. Interviews revealed further that the taste and appearance of the hybrid corn meal were judged inferior to the customary strain, and this was a second factor influencing people there to shift from corn meal to wheat flour. No such changes had taken place in the control district.

Arranging for organizations to keep additional records

If organizations or individuals do not collect the kinds of data required to indicate project results, they may nevertheless be willing to collect them, provided that they are aware of the constructive purpose this may serve for the community or nation. This approach has not been fully explored, but it deserves to be. Even if it is paid for, collection by such means may cost less than by other methods. Moreover, the form and completeness of the data can be better controlled if some payment is made.

If a local doctor is asked to keep a current record in a prescribed form of the ailments of all the patients he sees, this may give very valuable indications over a period of time of the effectiveness of a health campaign. If the

campaign is successful, the relative incidence of certain ailments would be expected to decrease. Arranging for a school-teacher or village official to count the number of children in several age groups who are and who are not in school is another example of this method.

In Ceylon, a consumer survey was carried out by paying families a small amount to keep track of all purchases over a certain period of time, an inventory of all commodities on hand being taken on the occasion of three sucessive visits made by the interviewer to collect the information recorded.[1] Storekeepers can usually also be persuaded (for a small fee) to keep track of the amount of fertilizer they sell, before, during and after a campaign to promote fertilizer sales. This general method is capable of being adapted in many different ways.

Another relatively inexpensive method of collecting data can be 'built-in' the development project itself. For example, a technical expert engaged in a development project can learn a good deal about his own effectiveness and can obtain data which will be helpful in planning his future activities if he keeps systematic records of how he spends his own time. A simple time slip, filled in each day, as is done by lawyers, engineers and a great many other professional workers, can be used to record this information and perhaps eliminate discursive reports summarizing the week's or month's activities.

If the expert uses a car for official business, he can use weekly or monthly speedometer readings to obtain a simpler and more accurate indication of the extent to which he gets around than he can reconstruct from his unaided memory.

A secretary (not necessarily his) can retain the old pages of his appointment calendar and summarize for him the number of conferences held or attended, the kind of people he met, etc. If he keeps (as he should) a complete chronological file of his outgoing letters and memos, a summary can be made of their number, length and general subject fields. This will give him valuable insight into the amount of writing he does, the number and kinds of contacts he has by way of the written word, and the subjects that demand most of his writing time (which may be considerable).

These kinds of data, though apparently superficial, in fact tell an expert a good deal about himself and help considerably when he decides from time to time to see what results he is achieving. The expert who started out by seeing the minister of agriculture once a week, and who now sees a liaison official three levels down only once a month, really knows that something has happened!

Other kinds of data on the results of expert activity are well known— the scientist's or engineer's diary of his step-by-step accomplishment, or

1. *Survey of Ceylon's Consumer Finances*, Colombo, Central Bank of Ceylon, 1954.

the series of drafts of legislation prepared by the expert. These are familiar and useful as concrete evidence of results.

Other kinds of possible 'built-in' data, valuable as indicators of project results, are the following: number of patients admitted to a clinic, and chief symptoms they report; number of students appearing for class each day; number of parents coming to each parent-teacher meeting; number of a school's graduates who reply to mail questionnaires, or who return to reunions; number of new students who state in a first-day questionnaire that previous students influenced them to enter school; number and proportion of families refusing to have their homes sprayed with DDT; number of tickets sold for toll bridge, or bus line, or railway; number of requests for pamphlet offered during radio programme.

In all these cases, the keeping of records adds little work and is helpful to the operation itself, as well as in measuring its impact. Many opportunities occur to build different kinds of record keeping into project operations and thus provide a continuing guide to these operations as well as evidence that can later be used to prove what results, if any, have been obtained in the meantime.

An obvious and very important kind of data collection, which can well be 'built-in' the development operations, involves records of employment, expenditures and activities. What kinds of persons were on the job, when? How much was spent for them and for other purposes? Whose money went

TABLE 2. Comparison of budget and expenditures[1]

Items of expenditure	Budget provision for 1952-53	Expenditures incurred 1952-53	Expenditure as percentage of budget
Project headquarters	63.2	65.6	104
Agriculture, animal husbandry and de-demonstration equipment	12	1.2	10
Education	2	0.5	25
Communications, petrol	7	3.6	51
Workshop equipment	3.8	0	0
Housing (repairs)	20	0.2	1
Others	11	0	0
Total	119	71.1	60

1. Source: *Evaluation Report on First Year's Working of Community Projects.* Government of India Planning Commission, Programme Evaluation Organization. August 1954. Adapted from page 402 of the article by Louis Moss, 'The Evaluation of Fundamental Education', *International Social Science Bulletin*. 1955, Vol. VII. No. 3, pp. 398-417.

for what ? What activities went on how long, on what scale and at what cost ?

Records of this kind are in any case essential for an organization, although they are not always matched up against the results to permit calculation of benefit/cost ratios, and, where results are disappointing, to uncover possible causes. Such comparisons may provide valuable insights.

Many development projects, for example, include staffing plans for workers of all levels. If, after six months or a year, there are serious gaps or excessive turnover ih the staff, this in itself may suggest why results are not greater. On the other hand, if staffing appears adequate in numbers at each level, and if results are not as great as anticipated, the reasons may be sought in the inefficiency or unsuitability of the staff selected, in unanticipated difficulties in the problem or group for which the project was designed, in inadequate financial resources, in poor local co-operation, etc.

Similarly, a matching up of expenditures against budget provisions may be very revealing. In an Indian project, for example, the expenditure data shown in Table 2 were obtained.

Needless to say, the project had not achieved the results anticipated from

TABLE 3. Plan and accomplishment. Project: Mohammed Bazar-Agriculture[1]

Activity	Target for years 1953-54	Accomplishment April to September 1953
Distribution of improved seeds		
Aus paddy (maunds)[2]	20	20
Aman paddy (maunds)	150	151
Wheat (maunds)	50	—
Potato (maunds)	300	—
Distribution of fertilizers and manures		
Sulphate of ammonia (tons)	100	59
Paddy fertilizer mixture (tons)	100	531
Bonemeal (tons)	90	20
Village compost (maunds)	1 000	420
Sann hemp (maunds)	100	38
Short-term loans (rupees)	50 000	—
Storage godowns (rupees)	8 000	—
Afforestation (acres)	500	—

1. Source : *Evaluation Report on First Year's Working of Community Projects*. Government of India Planning Commission, Programme Evaluation Organization, August 1954. Adapted from table in mimeographed draft of the article by Louis Moss, op. cit.
2. Denomination of weight current in India and Western Asia, the standard maund is now 100 lb. troy or 82.30 lb. avoirdupois.

it. When this record was examined and the reasons for underspending investigated, it became apparent that in certain fields there had been inadequate forward planning, excessive centralization of operational decisions, and unrealistic expectations about local circumstances. Keeping the records on which this table was based was a normal aspect of operations. Yet these data were very valuable in explaining the lack of project results and in suggesting steps to be taken to make the project more successful.

Records of project activities, in comparison with plans, are likewise of value. Again, let us take as an example these data from India (see Table 3).

Interpretation of this table requires knowledge of seasonal factors, of course. Disregarding that, it would appear that some targets were too high and some too low, or that some activities were less well administered than others (assuming that the targets were realistic), or that there were gaps in the reporting system.

Again it should be noted that these are data of a kind necessary for ordinary administrative control and yet of great significance for interpreting results and guiding future project operations. None of these three—staffing, expenditures, or activity records—provides final answers, but they all suggest lines of inquiry to be followed in collecting data of other kinds.

Observation

Observation of the effects of activities goes on all the time, of course. With relatively little effort, existing methods of observation can be refined to provide objective and often quantified data about the results of development projects. The value of such observational data will usually be greatest if: (a) the purpose for which the observations are to be made is clear; (b) they are planned systematically rather than made haphazardly; (c) they are systematically recorded and related to other events or activities; (d) they are subject to checks and controls with respect to validity, reliability and precision.

Observation may deal either with spontaneous behaviour in typical life situations (before a development project is initiated) or with behaviour somehow influenced by new elements (during and after a development project). In either case, observation has the advantages that it does not depend on the person observed being able to report his own behaviour, and it likewise does not depend on his willingness to report. Moreover, the observer may sometimes, although not always, remain outside of the situation he is observing. The method of obtaining data by questioning individuals, as described under 'Asking people about projects and results', page 41, does not have these advantages.

Because different observers often give quite different reports about the same situation, especially when they are observing group behaviour, various 'controls' may be introduced. These controls include: (a) careful definition of the units to be observed (for example, all individuals ?, all adults ?, all men ?, school classes ?, voluntary groups ?) and of the information to be recorded; (b) selection of pertinent data for observation (for example, use of specific tool in the field, drawing water at approved well, using screens on windows); (c) standardization of the conditions of observation (for example, prescribed day of week, time of day, place, persons who will act as observers); (d) whenever possible, the use of mechanical devices, tests and other aids to accuracy.

Observations can be quantified, but qualitative analysis is usually desirable to complement the quantitative.

Quantification may take the form of frequency counts (how many people draw water from the village well each day) and may be supplemented by grading or scaling devices (number of persons drawing water being divided into classes such as those simply drinking there, those taking away water to drink, those bathing and taking water away to drink, etc.).

Where the purpose of a development project is to introduce and promote improved techniques, such as the use of metal ploughs or the installation of screens on houses, systematic observation of sample areas at predetermined intervals will give invaluable information.

Where the purpose of a project is to improve and encourage the use of clinics or schools, counts can be made of attendance.

Where the purpose is to build up voluntary organizations, the size and frequency (and character of attendance) of their meetings can be observed.

A simple turnstile with a cheap mechanical counter attached, or an observer with such a counter in his hand, can be used to count attendance at meetings, or at exhibitions, or crossing a new bridge, or patronizing a store. Electric devices, some of them quite simple, have been widely used to count the number of vehicles passing a given spot.

Photographs can be good indicators of project results. Indeed, still photographs, before and after, may tell 'more than a thousand words'. Motion pictures taken of group behaviour can also be shown to observers individually, after which their reports on the film can be compared and a consensus report arrived at. Motion pictures of a main village street, before and after the impact of a development project, may show much about changes in the life of the town, illuminating and supplementing data collected in other ways.

In projects affecting many communities, or many households or persons in a smaller area, it is rarely practical to observe more than a small fraction of those expected to be affected. The selection of those to be observed needs

some attention if the observations are to be even roughly representative of the whole group. This process of selection ('sampling') is briefly discussed on pages 47 and 48, and is discussed in more detail under 'Sampling' in the appendix, page 94.

Asking people about projects and results

Recorded data and summaries of direct observations are of great value in themselves as indicators of project results. They may be even more revealing and even more convincing if corroborated by reports from the people personally involved in or affected by the development project. Moreover, some results, such as changes in information, attitudes or opinions, can usually be discovered only by obtaining such reports.

Going directly to people to ask them what has happened is a familiar procedure. Informal methods of questioning are constantly in use by the visitor in a new environment, the reporter on his beat, or the diplomat at a dinner party. The formal methods of clinical interviews, opinion polls, market research and census taking have been the subject of much study and refinement. If adequate funds and staff are available, and if a particular development project justifies the effort, highly scientific methods of obtaining data by questioning people may be employed. These methods are described in detail in the references listed in the bibliography.

Many development projects cannot justify, nor have they the funds and personnel available to carry out, these highly refined procedures. Certain principles developed in connexion with scientific interviewing are, however, applicable to the limited and informal questioning which may be undertaken without extra funds and staff, and which may repay many times over the modest effort and time involved. This section discusses these principles under the schematic headings of 'What?' 'Who?' 'Whom?' 'How?'. It is assumed that the questioning takes place during the course of the project, and that its purpose is to aid in carrying the project forward. If questioning takes place after the project has ended, the tenses of the verbs used below must, of course, be changed.

WHAT TO ASK?

It is often profitable to ask people several quite different kinds of questions about projects. Six major kinds of questions are suggested here.

What do they know about project activities?

1. Do persons intended to be affected by the project know of its existence ?

Have they observed or heard about any project activities ? Have they participated or been invited to participate in any activity ?
2. What do persons carrying out the project report has been done so far ? As suggested in Tables 2 and 3, operations often do not take place according to the time schedule or in the magnitudes provided for in the original plan or budget. Results can hardly be appraised without knowledge of what is actually taking place, and often the only way to learn this is to ask people.

Moreover, what people report about a project differs greatly according to whether the person questioned is an official at the central office, an official at a field office, a project worker in the field, or a person affected by the project. These discrepancies in themselves help a good deal in interpreting project results.

What do they think are the purposes of the project ?

Here again, there will probably be great discrepancies between the answers given by persons at different 'levels' in a development project. Some may believe, for example, that a project is intended to raise production of wheat ; others may consider it primarily as a method of training rural leadership ; still others, as a way of increasing general interest in scientific farming.

What the purposes are understood to be will obviously affect the way participants behave. Some of them may already be collecting data to indicate the project's results, but they may be collecting the wrong kind of data, or only one kind of desirable data. For example, for some purposes, raw statistics of production may be adequate indicators ; for others, attendance at agricultural extension meetings may be much more significant.

If the various participants themselves have different ideas of a project's goals, they may be working at cross purposes some of the time, and none of the purposes will be served as effectively as they might be. Information indicating present understanding of purposes may therefore throw a good deal of light on the success being achieved by the project. It may also suggest putting some effort into obtaining better understanding of purposes as a means of maximizing project results.

A telling illustration of the extent of misunderstanding about purposes that may exist, particularly at the beginning of a project, comes from a village development project in India, in which representatives of the American Friends Service Committee participated. A survey of the people in and around the village affected was made by an Oriyan anthropologist to learn what the villagers thought were the intentions of the AFSC people, then in temporary residence on the fringe of the village.

He found: '(a) Some of the people say that we will acquire the best lands of the area and start a farm. (b) We will change the caste system and make

everybody equal. We will change the religion of the people. (c) We have acquired the area of Barpali from the Government of Orissa as it could not pay off the loan taken from the American Government for the purpose of constructing the Hirakud Dam Project. We are permanently settling here and will rule them. (d) We are very high officers delegated by the Government of Orissa to settle and rule the people of this area.If this is not a fact then the Government people would not have built buildings for us and the great officers such as deputy commissioners and members of the Board of Revenue would not have come to us to pay homage to us. (The visits of high Government officers to us make them believe that we are concerned with the Government and make us higher in position than those who visit us.) (e) Some believe that we are going to change the way of life from individual families to community living. Under the conditions of community living we will ask them to have a common kitchen regardless of caste, creed and nationality, and they will have to live under one roof. They need no longer take care of their children as their children will be taken care of by us in our maternity centre. As soon as a child is born in the village, the parents will be asked to hand it over to us. (f) Some of the villagers are under the impression that we will not allow the leper patients to live in the area of Barpali. We will either shoot them or take them away from this area and keep them in a leprosy hospital. (g) We have come here to influence the people politically in order to gain votes.'[1]

How important do they themselves feel these purposes are?

This question is addressed primarily to the persons affected by the development project. Their reactions are important for its success, for unless they feel that the project serves purposes important to them they have no motive to participate wholeheartedly in making it a success.

The replies to this question therefore indicate: (a) whether or not the necessary popular support exists; (b) how much additional effort needs to be put into explaining the purposes of the project; (c) whether and what changes need to be made so that agreement may be reached on what is to be done to achieve what results.

Although the following example required a greater expenditure of funds and staff time than can usually be justified for smaller development projects,

1. Howard M. Teaf, Jr., 'Origins of a Private Improvement Project', in: Howard M. Teaf, Jr. and Peter G. Franck (eds.), *Hands Across Frontiers*. The Hague, Netherlands Universities Foundation for International Co-operation. 1955, and distributed in the United States of America by Cornell University Press, Ithaca, N.Y. The quotation is from 'A Brief Report on General Survey' made to the Orissa Project of the American Friends Service Committee by Nityananda Patnaik in October 1952.

let us consider the simple but highly illuminating interviews carried out in India by the Programme Evaluation Organization working with the Community Development Projects. These interviews were carried out in 16 villages at an early stage, some six months after the projects were started. The first questions were about what was being done ('What do they know about project activities?', page 41). The villagers were asked if they had heard that the community project was being started in their area, and how they had heard about it. Nearly a third had not even heard of the project, and in one village 85 per cent had not. The projects were much better known among small farmers than among agricultural labourers or non-farm workers.

The villagers were also asked if they had been personally invited to take part in some of the project activities in their villages. Only 37 per cent replied

TABLE 4. Project purposes reported and activities considered important by respondents[1]

Programme	Respondents reporting					
	What project is trying to accomplish		Most important need of village		Most important need of respondents	
	No.	%	No.	%	No.	%
Improved agricultural practices . .	386	33	240	20	209	17
Improved seed	237	20	14	1	21	2
Manure/fertilizers	161	14	44	4	73	6
Compost pits	140	12	1	—	—	—
Irrigation	238	20	267	22	191	15
Communications	650	56	199	16	100	8
Veterinary aid	193	17	—	—	—	—
Medical facilities	238	20	22	2	9	1
Sanitation	459	40	102	8	46	4
Education	431	37	73	6	45	4
Cottage industries and employment	16	1	60	5	105	8
Loans	20	2	20	2	72	6
Co-operative societies	54	5	20	2	14	1
Land allotment/land reform . . .	—	—	—	—	153	12
Total number of respondents reporting	1 159		1 215		1 235	

1. Source: *Community Projects—First Reactions.* Government of India Planning Commission, Programme Evaluation Organization, August 1954. This table is taken from page 410 of the article by Louis Moss, 'The Evaluation of Fundamental Education', *International Social Science Bulletin,* Vol. VII, No. 3, 1955, pp. 398-417.

that they had, and in some villages only 11 per cent so replied. Again the small farmers (44 per cent) had been reached much more effectively than the agricultural labourers (21 per cent).

The villagers were then asked what they thought the project was trying to accomplish; which activities they thought most important for developing the village; and which activities they thought most important for the benefit of themselves and their families. The replies are shown in Table 4, p. 44.

Note the discrepancies in this table. For example, 56 per cent of the respondents knew that the project was intended to improve communications; yet better communication was not a 'widely felt need'. Only 16 per cent considered that improved communication was 'most important for the development of their village' and only 8 per cent that is was 'most important for their own benefit'. Similarly, the need for medical facilities and sanitation was apparently felt much less generally than project planners may have expected.[1] On the other hand, the most widely felt needs of the villagers were for irrigation and improved agricultural practices, but far less (20 and 33 per cent respectively) knew that the project also had these purposes.

These data were of great value in administering the projects, for they indicated what changed emphases were necessary and how much time had to be allowed for the process of community development. Moreover, analysis of differences among villages indicated the relative effectiveness of different workers and different educational methods.

What changes, if any, have they observed or do they believe have occurred during the course of the project? (For example, changes in other people, in production, or in social or economic conditions).

If records or other available data give satisfactory information on changes occurring during the project, little purpose is served by asking this question. In many situations, however, reliable data indicating changes are not available, or are available for only a few of the possible changes in which project personnel are interested. A limited number of interviews with people likely to be informed about or actually affected by the project and about the changes associated with it will usually be most worthwhile.

---

1. It should be noted that the percentages in the second column are based on total responses to a question which permitted multiple responses, and these in fact averaged nearly three per respondent; while the percentages in the fourth and sixth columns are based on total responses to questions which permitted only one response per respondent, and only 88 and 84 per cent respectively gave any response to those questions.

How have the persons in the project area themselves changed, if at all, during the course of the project? (Their information, skills, attitudes, behaviour, social relationships).

To some extent, observation without asking questions will give adequate information about changes in skills (for example, literacy), behaviour (for example, having their cattle inoculated) and social relationships (for example, forming a parent-teacher association and holding meetings). If direct observations are to be made, however, some interviews might well be made at the same time. They take little more time and add significance to the observations. Moreover, only limited information on some important changes (information and attitudes) can be obtained without questions being directed to the persons expected to evidence these changes. In some development projects the most important results sought are changes in attitudes, such as willingness to do manual work, willingness to save and invest, willingness to take risks in private or public enterprises. It is of course possible to obtain clues to the extent of such attitude changes without questioning the people involved, but asking about these changes can bring out a richness of experience and detail that adds greatly both to understanding and to conviction about what has occurred.

Which of these changes (the two immediately preceding questions), if any have occurred, do they believe have resulted from particular project operations?

Even when the 'hardest' facts about changes are available, as when reliable production data have been collected in the development area and in a 'control' area, sceptics may question whether these results have been brought about by the development project. Corroboration by personal interviews with persons affected by the project, or informed about it but not engaged in it, can be very useful. Moreover, administration of the project as it goes along can be aided greatly by data indicating that certain operations produce certain results. It may be that the greatest results are being obtained from operations that cost the least.

In an example given above, shopkeepers' records of flour sales were drawn upon to help explain the reason why cultivation of hybrid corn resulted in increased consumption of imported wheat flour. If such shopkeepers' records had not been available, fairly convincing evidence could still have been obtained by: (a) asking shopkeepers in the two areas if they had noticed a change in purchases of imported wheat flour during the period in question—and if so, what they thought had brought it about (perhaps its price had dropped) ; (b) visiting a random sample of households and getting information on which to base estimates of *per capita* consumption of wheat

flour in both areas after the hybrid corn project had been in operation for some time (this would be only suggestive, not conclusive, as the differences found might have existed before the project began); (c) asking family members in both areas if they remembered a change taking place. At the same time, they could be asked to give the reasons why they felt that the change, if any, had taken place.

WHO SHOULD ASK?

Except where it is feasible to employ a trained staff of interviewers the process of asking questions about project activities and results is likely to be most productive if undertaken by persons having direct responsibilities for the project, that is, by experts or officials of the responsible government or private agency. If a foreign or intergovernmental organization is co-operating by furnishing technical or economic assistance, its experts and other representatives may also be interested in doing some of this work. Of course, there is always the possibility that some bias will be introduced if persons carrying out a project themselves ask questions about project results. On the other hand, there are also disadvantages if other persons are brought in to collect data (see pages 66-67). An agreed schedule of questions to be asked and an agreed division of labour will enable all interested groups to obtain the most information at the least cost.

A trained staff of interviewers is, of course, a great asset in obtaining the information desired, if a large-scale survey is to be made. Even where such a staff is available, however, the same general principle about responsibility for the inquiry seems appropriate. The staff should ask questions on behalf of the responsible agency (or agencies) and in close co-operation with the experts and administrative staff directly involved in the project. Otherwise, the asking of questions takes on the aspect of an outside investigation, rather than a tool of internal good management.

WHO SHOULD BE ASKED?

Since most development projects seek to bring about changes in or through people, testimony concerning these changes is logically sought by directing questions to the people personally affected—the members of a training class, officials of an extension service, farmers in a particular area, workers on an irrigation system, etc.

If the group directly affected by the project is small enough, it may be possible to question all its members, although usually a sample must be selected. Sample size and methods of selecting samples to avoid bias are described under 'Sampling' in the appendix, page 94. Locating this topic in

the appendix, along with other detailed statistical topics, does not mean that it is unimportant, however. The sample used must be large enough, or conclusions drawn from it will be unreliable. Beginners often jump to erroneous conclusions because they rely on samples that are too small. In fact, if one wants to subdivide the sample for refined analysis, the sample needs to be quite large. Another factor determining the size of sample needed is the size of the percentages obtained on particular questions. Simple random variation can result in percentages around 50 per cent differing by several points from one sample to the next if the samples are not quite large. Percentages around 90 or 10 per cent (i.e., towards either extreme of 100 or 0 per cent) are less likely to vary much. Precise definition of the 'universe' to be sampled is also important.

Finally, it should be pointed out that a frequent source of bias in results can be avoided if every member of the entire group has an equal chance of being selected in the sample.

For example, if a list of the names of all members of a training class, or all officials of an extension service is available, a sample of twenty or so names may be drawn by chance. Or the names may be written down (in any order) on a list, and every tenth, or twentieth, or fiftieth chosen for the sample. Both methods give every member an equal chance of being selected.

In a village or larger area, it is desirable not to interview just those living along the main road, or just those in town on market day, or just those selected by the village headman. It is much better to make a map, even a rough one, of the area and to spread even such a small sample as twenty interviews among the various communities (if possible, in proportion to density of population). At any location in the area, the individual informant should, if possible, be selected by some random method—avoiding concentration on the most or least prosperous looking, the most or least interested looking, the most or least accessible, etc. Some distortion through sample selection is likely to enter into any group of responses, but attention to the principle of equal chance of selection will minimize this distortion.

It is not only those affected by a project who may have valuable information to give. As suggested above, it may well be worthwhile to ask questions of various other people involved in the development project or likely to be acquainted with it. At the local level the village headman, schoolteacher and other community leaders can give useful information, and their attitudes towards the project will be significant. They may report results which might not otherwise be observed, and their attitudes, as leaders, may be very important determinants of the success or failure of the project. On the other hand, they may have developed a personal interest in the success or failure of the project, and interpretation of their replies should take into account their possible bias.

Experts and administrators involved at various levels in the project (in the field, at regional offices, and at headquarters) are also prime sources of information on project purposes, activities and results. Again, their replies may reflect their own special interests but are none the less valuable.

Particularly for large projects or for projects located in urban areas, a number of individuals may be well informed about the project although not connected with it—officials of government agencies not directly responsible for the project, members of a university faculty, businessmen, political figures, reporters and editors, diplomats, and persons concerned with other development activities, such as foundation representatives, and United Nations or other international personnel. Those having political interests or functions may also have an axe to grind, of course, and this needs to be taken into account in interpreting their replies.

It is not merely the information and attitudes discovered in this way that are important; it is also the discrepancies among various groups of replies. These can be brought out by summarizing separately the results of questioning each group. When the special interests of the different groups of informants have been taken into account, the discrepancies among groups may well be used to suggest: (a) the need for further checking of reports ; (b) the need for fuller agreement on purposes ; (c) the need for better communication about purposes, activities and results ; (d) the need to make changes in actual project activities or in their administration.

HOW SHOULD THE ASKING BE DONE? [1]

Obtaining information and insights by talking directly with other people is, of course, a continuing activity of all persons helping to carry out development projects. To the extent that this often spasmodic and rather subjective process can be made more systematic and objective, the information obtained will be the more valuable to those directly concerned and will carry greater conviction.

The situation where one person questions another is not, however, a simple one. What is the purpose of the questioning ? How did the respondent happen to be selected? What will be done with his replies? What will happen if he doesn't reply? Add to these such additional factors as differences of status, political stresses and sometimes the presence or actual

---

1. Much of this section has been adapted from Charles F. Cannell and Robert L. Kahn's ' The Collection of Data by Interviewing', in: Leon Festinger and Daniel Katz (eds.), *Research Methods in the Behavioral Sciences*, New York, Dryden Press, 1953, Chapter 8, pp. 334-61, by permission of the authors and the Dryden Press, whose copyright it is. Those desiring a fuller and very excellent treatment are referred to Robert L. Kahn and Charles F. Cannell's *The Dynamics of Interviewing*, New York, J. Wiley & Sons; London, Chapman & Hall. 1957.

participation of a foreigner, and it is obvious that the situation is quite complicated.

Still, a good deal of experience has been accumulated in ways of asking questions systematically and successfully. Moreover, certain general psychological factors are known which can be used to improve results.

Procedures in asking questions

Introducing the series of questions. Ordinarily, the questioner's first tasks are to :
1. Explain the purpose and objectives of his questioning.
2. Describe the method by which the respondent was selected, and indicate how many other people (usually without giving names) in similar circumstances are being interviewed.
3. Identify the sponsor or the agency which desires the information.
4. State the anonymous or confidential nature of the questioning.
5. Establish by his manner, his tone of voice and his sensitivity, a genuine rapport with the respondent.
6. Motivate the respondent to communicate fully and frankly.

It is of the greatest importance to convince the respondent that the questioning is not a method of checking up on him or his activities, nor a way of uncovering facts reflecting on his superiors, on his friends, or on particular individuals participating in the project. He must recognize it as a means of guiding and improving the operations of a development project, and thus as having a solely constructive purpose. In the introductory part of the interview, the questioner tries to sense the respondent's attitutes towards the project and to explain to the respondent how the interview relates to these attitudes. For example, he may well point out that the interview provides an opportunity for the respondent to make his views known to those in authority who sincerely want to learn the general public attitude and who want to get from the public specific suggestions for improvements.

Asking the questions. The ideal procedure is to ask exactly the same questions of each respondent in identical fashion and in identical order. This requires careful wording and pre-testing of questions to be sure that they are understandable to all respondents.

Such standardization is often impractical, of course, but it is desirable to keep the form of questions as constant as possible, primarily to avoid the personal emphasis or distortions which may creep in during improvisation.

The sequence of questions may well be varied to fit the natural development of the discussion. Even in an informal interview which takes the form of a discussion, it is desirable to have at hand a prepared list of questions as

a reminder of the precise form of questions and of subjects which might unintentionally be omitted during a wide-ranging conversation.

Stimulating complete responses. In many interviews, some questions evoke responses which are incomplete or unclear. It is often desirable, when such replies are given, to supplement the prepared questions with 'probes' which (a) elicit from the respondent further information required and (b) clarify or make more specific information he has already given.

In probing for further information such phrases might be used as, 'Would you tell me some more about that?' or, 'I'm interested in what you say. Could you tell me a little more about that?' or, 'I see what you mean. Can you tell me a little more about how you feel on this question?'.

Probing for clarification may use such a phrase as, 'Now let me see if I have it straight. As I understand it, you feel...', followed by a summary of what the respondent has said. Or, 'I would like to read my notes back to you to see whether I have your point of view straight'.

Probing also helps a good deal to develop the permissiveness and warmth which are so important in keeping the respondent interested and answering questions fully.

Recording the responses. The only accurate way of reporting an interview is to record the responses at the time, either by mechanical methods or by taking notes. Much valuable detail disappears if recording is left until later. There are situations, of course, where it is embarrassing to jot down notes during an interview. On the other hand, an opening statement about the seriousness of the questions and the importance of reporting correctly the views of the respondent so that follow-up action can be taken will usually pave the way for taking notes in almost any situation.

Psychological factors in asking questions

Winning co-operation from those questioned. Why, in the first place, should a person agree to be questioned? This is not usually a problem for leaders and other informed persons at the village, urban or national level, because many are likely to be already familiar with the purposes and plan of a projected or active development project and may know personally, or know about, the individual or agency doing the questioning. The persons directly affected by the project, however, especially if they are uneducated (as in many rural areas), may at first be suspicious that the questioner represents the taxing authorities, or the law, or a political group. There may be suspicion even if the questioner speaks their language and is obviously their country-man (note the example on page 42). On the other hand, among all groups

curiosity and traditional politeness play a part. Moreover, most people customarily try to please persons in authority, and the interviewer may well be perceived as an individual, or as a representative of an agency, possessing authority and commanding respect from the respondent. Thus, if a questioner can persuade a regional official or village leader to vouch for him, he can usually win co-operation even in the most isolated areas.

How can the respondent's co-operation be maintained throughout the interview? The motives described in the previous paragraph may ensure a hearing for a few opening statements and questions, but they may not be strong enough to maintain the co-operation for very long. What additional motives may be appealed to, through early and proper use of statements and questions designed as 'rapport builders'?

The questioner will usually get the best co-operation if he states clearly and openly why he is seeking information. These opening statements of the purpose of the interview and the ways in which the data will be used should, if possible, relate them to existing motives or goals of the respondents.

The motives or goals of respondents that are favourable to co-operation include:
1. Approval of the general objectives of the development project.
2. Curiosity about the development project and specific planned operations.
3. A desire to play the respondent role successfully, to be able to answer the questions posed. (The questioner can help motivate the respondent by making clear the type of response expected, thus reassuring the respondent about his ability to give an appropriate and face-gaining reply).
4. A desire (usually) not to hurt, nor to cause to lose face, a person who is obviously of some consequence.

As the interview proceeds, two other major motives often come into play:
5. The desire to influence the questioner, and through him the development project or persons connected with it. It was common during World War II, for example, when surveys were being conducted in the United States of America on government programmes, for a respondent to preface his responses with, 'You tell those people in Washington that I said...'.
6. The desire to express oneself to another person.

There is considerable satisfaction to be gained from talking with a receptive, understanding person about something in which the respondent is interested and perhaps emotionally involved. This satisfaction is increased if the questioner is seen as warm and responsive, expressing a genuine interest in the respondent and acceptance of him as a person, entirely permissive and tolerant of all expressions of opinion, attitude and feeling; and the respondent must be free from any type of pressure or coercion and the questioner refrain from intruding his own wishes, reactions or biases into the interview

situation. The perception of the questioner as capable of understanding or of being completely tolerant of the respondent's point of view depends far more on the questioner's attitude and the sort of relationship which he establishes with the respondent than on such external factors as the questioner's dress or appearance, although these externals may provide some initial (and possibly inhibiting) cues for the respondent.

Establishing this friendly type of rapport should, of course, be kept in proper proportion to the real purpose of the occasion—in other words, there should be an unspoken agreement to 'get on with the job' of completing the interview.

Unfortunately, certain factors may inhibit communication or distort the content of the information given by the respondent. For example, he may not be in sympathy with changes sought through the development project, or with the political party currently in power and hence identified with the project, or with the idea of foreigners in general or foreigners from particular nations playing an important role in introducing such changes in his nation. He may be afraid that frank replies, if they imply any criticism of anyone, will be detrimental to his own interests, or to those of his friends, or will indicate some lack of appreciation of the efforts of well-meaning foreigners.

The respondent may, of course, refuse or restrict communication if he feels that there is a great difference in education or economic status between himself and the questioner and that the questioner is therefore incapable of understanding his family circumstances or of tolerating his widely different, even possibly deviant, opinions and attitudes. He may be right, of course. His replies may be distorted by the questioner in the light of the latter's stereotypes or pre-judgements.

Language. The language used in asking these questions should be as simple as possible while retaining precision of expression. When asking questions in a language other than the respondent's native tongue, the wording used should take into account the likely range of the respondent's vocabulary. This is likely to be an important consideration even when talking with highly educated government officials, who may know several foreign languages but not all of them equally well. Differences between British and American usage may be important, the former being the more familiar in many underdeveloped countries. Interviewing in connexion with rural projects should usually be carried out in the language of the common people, often different from the classical language of the country's literature.

Frame of reference. Equally important, the phrasing of questions should take into account the 'frame of reference' of the respondent. Each topic

53

should be introduced in a form adapted to the perceptions of the respondent and consistent with his notions of what is and is not obviously related to the topic under discussion.

The importance of frame of reference became evident in a United States Bureau of the Census survey of employment. During this survey the following question was asked, 'Did you do any work for pay or profit last week?' In reply, young people attending college and women keeping house generally replied, 'I am a student' or 'I am a housewife', even if they also did some part-time work for pay, because they saw their primary role as the more important. These responses resulted in many persons being classified as 'non-workers', although the fact that they were employed part-time, which was to them less significant, meant that they met the census definition of 'workers'. When the sequence of questions was changed so that the respondent first had a chance to explain his or her major occupation (as a 'non-worker'), and was then asked if, in addition to this major occupation, he or she also did any work for pay, many more 'workers' were identified, and the official estimate of employment in the United States was consequently raised by more than one million persons.

The respondent's 'frame of reference' may also be important in determining whether or not he will be willing to communicate a given piece of information. If this information does not seem to him to be clearly related to the topic he understood to be under discussion, the abruptness of change of topic (as he sees it) may cause a block, diverting the respondent's attention, shattering rapport and possibly even leading to termination of the interview.

It is often desirable to ask questions about several quite different topics during the same interview. An effective way of bridging different topics is through transition statements or transition questions. For example, 'Well, we've been talking about malaria control; now we want to talk a bit about fertilizers. How do you feel about using artificial fertilizers on your paddy?'

Information level. Questions asked should be limited to those which respondents can reasonably be expected to answer. Moreover, to protect those who may be unable to answer, questions should be asked in a way which will save face for the respondent. Asking a question outside the respondent's competence will produce 'Don't know' replies, which are not very helpful, and invented replies, which are positively harmful. Moreover, the failure of the respondent to do what is obviously expected of him by the questioner (whose very role invests him with some authority, in the eyes of the respondent) often results in embarrassment and resentment. Not only will no meaningful reply be forthcoming but motivation for communication will be decreased during the remainder of the interview.

On the other hand, proper phrasing can take into account the possibility that substantial numbers of respondents will not know the answer, and thus save face even if the respondent is unacquainted with the subject asked about. 'Many people haven't yet heard much about the project the government is starting up at X, but some have picked up information about it. Do you happen to know ... ?'

Social acceptance. A similar problem is posed by questions where certain replies are socially or politically unpopular. Deviant opinions and deviant behaviour are difficult to investigate because it is hard to convince many respondents that their replies will be kept anonymous, and even when they feel secure about their anonymity they may fear the disapproval of the questioner.

This problem can be met somewhat by: (a) special effort on the questioner's part to establish his complete tolerance and iron-clad guarantee of anonymity; (b) phrasing questions so that socially or politically unacceptable replies do not seem exceptional. For example, 'Surveys last year indicated that nearly half of the people in this area practise family planning in various ways. Do you think that ... ? or, 'Several of the people I've talked with feel that this project is doing as much harm as it is good. What do you think are its bad points and what are its good ones ?'

A similar problem arises when people become aware that they are expected to have acquired certain attitudes—such as the willingness to be vaccinated or to use improved seed—and yet are basically unconvinced. They will often give socially acceptable verbal responses, but never follow through with actions. Where this problem is anticipated, it may be helpful to use 'projective' questions such as, 'Do you think most of your neighbours are willing to be vaccinated ?' The attitudes they attribute to their neighbours are likely to be their own, 'projected' consciously or unconsciously.

Leading questions. Questions should, of course, be so phrased that they contain no suggestion as to the most appropriate response.

Obviously loaded questions are easy to avoid. Not so obvious, perhaps, but likely to cause bias is wording like the following, 'Would you say that you are in favour of the new teaching methods in the local school ?' Replying 'yes' here is easier than replying 'no', for saying 'no' seems to contradict the questioner.

A story is told of an obviously important traveller trying to find his way in a poor and status-conscious area. He repeatedly asked, 'Is this the way to Y ?', and repeatedly received, 'Yes, sir', for an answer. He never reached Y, for the answers were incorrect. The respondents knew it ; they weren't trying to mislead him ; but who were they to say 'No' to such an important person ?

The first question above could be freed of bias by wording it thus: 'Would you say that you are in favour of the new teaching methods in the local school or are you opposed to them?'

Loaded words can also be the cause of biased replies. Examples are, 'Asiatic' versus 'Asian', 'native' versus 'national', 'foreign expert' versus 'international expert', 'neutralism' versus 'independent policy', etc. Familiarity with local usage will help to eliminate words which have acquired favourable or unfavourable connotations.

Another way in which the wording of a question may distort findings is by linking one of the possible responses to a universally popular result. For example, 'Do you favour or oppose higher taxes to pay for universal primary education?' Because education is widely favoured, its popularity will influence the replies to the question on taxes. Such questions should be avoided.

The single idea. Findings are clearest if each question is limited to a single idea or to a single reference. If a question is asked, 'Do you favour building a central market and paving the rural roads to increase sales of the village's chief crops?', a reply may refer to either or both of the two projected improvements, and it will therefore be difficult to assess the popularity of either alone.

Question sequence. Questions are most readily understood if their sequence seems logical to the respondent. If the progression from item to item is easy and natural, the respondent can often anticipate the next question or immediately sense its pertinence.

The 'funnel approach' is usually helpful. This refers to the procedure of asking the most general or the most unrestricted question first and following it with successively more restricted questions. This sequence permits learning something about the respondent's frame of reference at an early point and reduces the danger that early questions will influence those that come later.

The following sequence illustrates the 'funnel approach'.

Question one. What do you think are the most important problems facing this village?

Question two. Would you favour or oppose asking the central government to send someone to work in the village and help solve its problems?

Question three. If the central government could send someone to work on one of these problems, which problem do you think it should be?

Question four. Would you favour or oppose establishment of a co-operative credit society here?

If the last question in this series had been asked first, it might well have distorted the responses to the other questions. Moreover, the first questions

make it possible to sense the attitude of the respondent towards change in general, and towards the central government in general, before pinning down replies to specific proposals.

The form of the response. There are two main types of question commonly used in sytematic questioning: first, the 'open' or 'unrestricted' question, which permits the respondent to give a wide range of responses and to reply in his own words. Questions one and three above are 'open' questions. Secondly, there is the 'closed' or 'restricted' question, which requires him to select from pre-assigned categories the response coming closest to his own opinion. Questions two and four above are 'closed', as the only replies sought are 'Favour', or 'Oppose'.

The open question has many advantages, owing to the fact that the respondent is encouraged to structure his answer as he wishes. A greater range and variety of information is elicited than with the closed question. The respondent may state his own frame of reference, or make this clear by the responses he gives. His level of knowledge or degree of expertness will often be revealed by his reply to the open question, and the questioner can then easily discover whether or not the respondent clearly understands the question being asked.

On the other hand, open questions often take up more interviewing time than do closed ones and it is much more difficult to combine the replies obtained and to summarize them in a statement such as, 'Thirty-five per cent of the group felt that . . .'.

The closed question is well adapted to situations in which: (a) there is only one frame of reference from which the respondent can answer the question ; (b) within this single frame of refererence, there is a known range of possible responses ; (c) within this range, there are clearly defined choices which accurately represent the position of each respondent. A closed question is quite suitable in asking about marital status. In monogamous societies, one can simply ask, 'Are you at present single, married, divorced, separated or widowed ?'

On the other hand, a closed question like the following is not well chosen, 'At the present time, do you think personal income taxes are too high, too low, or about right ?' Many people can express their attitudes satisfactorily on such an issue only by relating their answers to the progressive character of personal income taxes (perhaps too high on low incomes, too low on middle incomes, etc.) and to the level of governmental expenditure or the stage of the business cycle.

It is often practical to combine closed questions with follow-up open questions. For example:

Here is a list of measures sometimes used to control mosquitoes. Please tell me if you have taken any of these measures during the last six months.

	Yes	No
Screens on house	. . . .	. . . .
Copper sulphate on stagnant water	. . . .	. . . .
Draining stagnant water	. . . .	. . . .
DDT spraying of outer walls	. . . .	. . . .
DDT spraying of underbrush, etc., near house	. . . .	. . . .
Others, please specify. . . . .

I note that you do not do . . . .
Could you tell me why not ? . . . .

Tabulation and analysis are easier if questions can be answered with a check-mark. Interviewing time is shorter and response rate on questionnaires higher when such questions are used. They require most careful wording and pre-testing, however, to make sure that all words are fully and correctly understood, that the issues covered are really well enough crystallized in people's thinking to warrant check responses, that all suitable alternative answers are listed and do not overlap (except where intended to do so), and that bias is not produced by the wording or sequence.

USING A PREPARED INTERVIEW SCHEDULE OR QUESTIONNAIRE

Attention is given above to the care needed in designing questions and to the desirability of standardizing their form, to building rapport, to using a sequence of topics that seems logical to the respondent, and to using the 'funnel approach' from the more general to the more specific. As this is a good deal to keep in mind in extemporaneous interviews, it may be worth while to plan the questioning carefully and to write the questions down beforehand on an interview schedule or questionnaire, even when only a limited number of people are to be asked about a project.

An 'interview schedule' is usually kept by the questioner to remind him of the scope, form and sequence of questions. He may show it to the respondent, but he himself records the responses. Some interview schedules are merely notes to guide the questioner; others are so laid out that the answers obtained can be recorded directly on the schedule, often utilizing check-marks or figures when 'closed' questions are asked.

'Questionnaires' are essentially standardized interview schedules, usually composed of relatively simple, straightforward questions which can be answered briefly, perhaps by a check, a number or a single word. Giving or sending them to respondents to fill out by themselves is usually impractical, as relatively few replies will be received except from respondents who are both highly motivated and quite well educated. Questionnaires may be

filled out during a personal visit, however, in which case the questions may either be read aloud (and repeated verbatim when necessary) and the answers written down by the interviewer, or the respondent may be asked to read the question himself (if he can read) and to write down the reply (if he can write). Where the group questioned is literate, questionnaires may be administered to large numbers of individuals simultaneously, or may be left at offices or homes and later picked up or returned by mail. However, experience with self-administered questionnaires in underdeveloped countries has not been on the whole satisfactory.

The impersonal nature of a questionnaire—its standardized wording, its standardized order of questions, its standardized instructions for recording responses—ensures some uniformity from one situation to another. The 'do-it-yourself' nature of a questionnaire also has some value, as respondents feel that they are actively participating. Questionnaires returned by mail or administered to a group can be answered without revealing the name of the respondent, and this anonymity may be an important factor in obtaining full and candid replies on some topics. Moreover, when the respondent is given plenty of time to fill out the questionnaire, he can weigh his reply carefully, not being under the social pressure of a waiting interviewer (or an impatient animal) to reply quickly.

On the other hand, the interview situation does permit more flexibility (rephrasing questions, changing their order to follow the respondent's flow of ideas, probing) and more opportunity to check on possible misunderstanding by asking additional follow-up questions. Questionnaires are obviously difficult to use with illiterate respondents, and societies that are unfamiliar with them may give quite unsatisfactory results if questionnaires are self-administered. Complicated questionnaires which require much writing can be used with only a very small percentage of the population, of course. For these reasons, questionnaires are less useful than interviews, but there are cases where their advantages are significant enough to justify trying them.

Every interview schedule or questionnaire is improved by pre-testing. No matter how carefully thought out and discussed, it will in practice be found less than perfect. Some of the 'best questions' may fail to elicit the type of data needed. They may not be understood at all by respondents, or may be misunderstood in terms of a different frame of reference. The sequence may disturb respondents, or the use of some words may cause embarrassment or resentment.

Pre-testing in the field is the only remedy and a most important safeguard. Relatively few interviews are needed for a pre-test, but some should be made by the kind of persons who will be using the interview schedule or questionnaire in the actual study, for interviewers occasionally misunderstand their

own schedules and instructions. The pre-test results will usually dictate a number of changes.

CONTROL GROUP

Attention must also be drawn to the usefulness of a control group. In Step Four, when analysis and interpretation are discussed, comparison appears to be of great importance. The possibility of comparison may come from the comparable study (other area, other time, similar project). Another possibility is the conscious creation of a control group. This is a group of persons unaffected or nearly unaffected by the action. Such a control adds to the work but may contribute more to a good interpretation of results than many a refinement (which also means more work) in the action study proper. A control—if feasible—broadens the 'baseline'.

Examples of data used to indicate project results

It is rarely possible to collect all the kinds of data described above without devoting substantial time and effort to it. Collection of data of some of these types will usually be feasible for almost any project, however. Imagination and ingenuity will be valuable if even a modest effort is made to obtain more complete, more systematic and more objective data indicating project results.

Several examples are given on the following pages of typical 'minor' projects and the easily collected data used to indicate their results.

Malaria control

Project purpose: To kill mosquitoes and reduce incidence of malaria in village A.

Project operations: One spraying of all homes with DDT in March 1955, and distribution of cartoon books showing symptoms, causes and treatment of malaria.

Expected results		Unexpected results	
Results	Kinds and sources of data	Results	Kinds and sources of data
Increased consumption of antimalarial drugs.	Sales of antimalarial drugs (data already available, before and after, in records of local dispensary).	Markedly reduced annoyance from insect bites (bedbugs and lice as well as mosquitoes).	Report made by village headman in informal interview.
Reduced incidence of malaria.	Spleen rate among all patients at local clinic (data already available, before and after, in records of clinic).	Erection of screens at windows of a few houses.	Observations made while driving through village, supplemented by interview with head-man.
		Formation of village committee to request and raise money for a second spraying.	Report made by regional official of public health service, in interview.

Fertilizer application

Project purpose: To obtain higher yields through wider, more intensive and more scientific application of fertilizer to chief crop of district B.

Project operations: Sale of fertilizer at half price for two years, personal encouragement by agricultural extension agent, dissemination of information through school classes and literacy classes, filmstrips, distribution of pamphlets.

Expected results		Unexpected results	
Results	Kinds and sources of data	Results	Kinds and sources of data
Increased sale of fertilizer.	Records of shipments of fertilizer to principal market town in district B (already available at Ministry of Agriculture, before and after).	Changed pattern of land use resulting from higher profitability of fertilized crops versus other crops.	Records of amounts of other crops marketed at principal towns in districts B and C.
	Number of new applications from retail dealers wishing to handle fertilizer (data available on request from wholesaler).		Interviews with agricultural extension agents in districts B and C.
	Number of requests for additional copies of pamphlet on fertilizer use (data available from extension agent).	Increased pressure for change in land rent, and reduced resistance by landlords to this change.	Interviews with journalists, legislators, foreign observers.
	Extent of discussion of fertilizer use initiated by members of school and literacy classes (interviews with teachers).		
Increased yield.	Records of amount of grain marketed at principal towns of district B and of district C (a control area).		
	Interviews with several farmers concerning yield of specific plots, before and after.		

Transportation

Project purpose: To facilitate the movement of merchandise between a major town X and a large agricultural area.
Project operations: Building a highway bridge across a large river.

Expected results		Unexpected results	
Results	**Kinds and sources of data**	**Results**	**Kinds and sources of data**
Lower cost of transportation across river.	Interviews with two leading farmers in agricultural area.	Increased deposits in savings bank in town X.	Bank records, before and after.
Increased marketing of agricultural produce in town X.	Data on sales from records of government agents at markets, before and after.	Increased attendance at secondary schools in town X, by students from across river.	Records of school enrolment, plus interview with a school principal.
Increased sales by merchants in town X.	Data on sales, recorded by merchants for a small fee, before and after.	Increased circulation of newspaper published in town X.	Records of newspaper sales.
Increased transportation of freight across the river.	Estimates of weekly movement of freight by ferry and small boat before bridge built (from observations at the time).	Increased interest in national affairs by people in agricultural area.	Interviews with village leaders in area and with political leaders in town X.
	Estimates of weekly movement of freight by vehicle over bridge (by actual count of vehicles during one week and estimate of amount of freight of each tenth vehicle, made by a local observer hired for purpose).	Establishment of three new stores in agricultural area (based on its increased income).	Interviews with village leaders, and visit to one new store.

Deciding what data to use to indicate project results

Mechanical arts training

Project purpose: To establish, equip and staff an institute of mechanical arts, and train two hundred industrial mechanics a year.
Project operations: Training ten teachers (three of them abroad), setting up organization, constructing and equipping the buildings, developing curriculum and teaching materials, recruiting and selecting students, giving training courses.

Expected results		Unexpected results	
Results	Kinds and sources of data	Results	Kinds and sources of data
Four buildings constructed and equipped.	Observations and photographs made at institute; interviews with director and chief mechanic.	Establishment by graduates of two new garage and repair shops.	Location, size and quality learned by interview with director.
Ten teachers trained.	Extent of training and excellence of performance learned from interviews with director and from records of four training agencies.	Heightened prestige of mechanical employment.	Improved quality shown in records of entering classes. Popularity increase learned in interviews with principals of two schools from which students drawn.
Organization of staff.	Structure, procedures and effectiveness discussed in interviews with director, four teachers, librarian and chief mechanic.		
Students in training.	Number and quality of students, and reasons for attending, learned from interviews with director, four teachers, chief mechanic, librarian, five students and with principals of two of the primary schools from which students came.		
Competent industrial mechanics.	Records of graduating classes. Competence indicated by records of examinations and by interviews with employers of several students. Number of graduates learned from records of graduating classes.		

Collecting the data—
before, during and after

It is usually desirable to centralize the guidance of data collection, but to decentralize the actual collection. As suggested above, some data may be obtained from existing sources, or arrangements may be made with such various sources as local shopkeepers, professional people, or research centres in universities or government agencies to collect desired data. Other data may be collected in interviews carried out by government officials, by project experts, by officials of international agencies, or by trained crews of interviewers. But the greatest efficiency will be attained if one person or office explicitly takes reponsibility for controlling the collection of the desired data. This may perhaps be the person who works out an agreed statement of objectives (Step One, 'Stating objectives', page 22), or the office which keeps records of project activities (Step Two, 'Arranging for organizations to keep additional records', pages 35-39), or someone who does a major share of the observing and the asking of questions (Step Two, 'Observation', page 39, and 'Asking people about projects and results', page 41). But the knowledge about what is being attempted in the project, and the flow of data to indicate what the project is accomplishing, need to be brought together at some one point.

Throughout the process of collecting data another principle appears to be important, that data collection should as far as possible be handled by the people actually involved in the process of change. One set of interested persons are officials and experts of the responsible organization, often in association with international experts. Their records can be the source of much data, and their record keeping can often be expanded to include a considerable range of additional data. Information intended to help operations is most likely to be pertinent and to be used if it is collected by those responsible for the operations.

Another set of interested persons are those affected by the project

(the trainees, the counterpart personnel, the farmers, the families adopting new health practices) who want to know what is going on, and can themselves provide information on project results. Both providing information and getting information are important to them. Unless people are able to understand what is happening to them, and what, in turn, they are causing to happen, much of the value of development may be lost.[1] Responsible participation in the political process can take place only if information is freely available, freely discussed, and utilized in reaching decisions as to policies and programmes and as to the people who are to carry them out.

Helpful as this may be to local persons directly involved in a project, participation in data collection may be even more valuable to experts or administrators who come from outside the project area, particularly if some of them come from abroad. Collecting data in the ways suggested is one of the best ways for them to learn what needs to be done and how the task may be accomplished. Briefing sessions to discuss the objectives of the project and the kinds of data needed for defining objectives and evaluating accomplishments may contribute greatly to the effectiveness and morale of the personnel of the project and to the efficiency and impact of its operations.

On the other hand, there are some dangers inherent in an expert or administrator from headquarters or from abroad playing a major part in collecting data on project operation and results. Despite the fact that this 'outsider' may be the person having the greatest competence, interest and even time for collecting such data, others may be given the impression that he is studying the local people as an aloof anthropologist might study a strange tribe, or that he is conducting some kind of experiment on people who are, after all, simply subjects for his experiment, or that he is looking for shortcomings in the operation of the project, or that he wants facts to enable him to reach his own conclusions and pronounce his own judgements. Any suggestion of such motivation is, to say the least, unfortunate and can well undermine morale and efficiency. The 'outsider' would do well to take positive steps to avoid creating such impressions. To this end, he might consult with local personnel about what information should be collected, and for what purposes. He might ask help from local personnel in the collection of data and make available full reports on all information uncovered. He might also review his tentative findings and interpretations with local personnel (see also Sep Four, 'Analysing and interpreting findings', page 85), for his position is an extremely delicate one, requiring the utmost sensitivity and tact.

The suggestion has often been made that outside research experts be

1. For elaboration of this point see Louis Moss, 'The Evaluation of Fundamental Education', *International Social Science Bulletin*, Vol. VII, No. 3, 1955, pp. 398–417.

brought in for the specific purpose of collecting data on the results of development projects. The notion of outside *investigators* is incompatible with the view we have presented that improved data collection should be regarded as a management tool to be developed and utilized through the co-operation of all interested parties. It is perhaps true that a person working on a project can rarely maintain complete objectivity, both because he identifies himself with the project and because of (possibly silent) pressures from his associates to endorse their activities. It may also be true that a 'fresh view' will reveal new questions and implications; and that an outsider has greater freedom of movement in an agency, partly because he has a certain ambiguous and yet authoritative status, partly because people often talk in confidence to a stranger, and partly because he is not identified with any faction among the project personnel. On the other hand, attempts by potentially hostile strangers to collect data may meet barriers to frank disclosure, time may be wasted in getting acquainted with the project and the people concerned, and the report produced may be simply ignored.

If a 'fresh view' or a more 'objective' one is wanted, or if advice is sought on more effective and economical techniques of collecting and analysing data, it may be worth while to bring in outside experts to *advise* on the process of collecting data. Such experts are usually most successful if they become part of the team carrying on the project, and operate in such a way as to be acceptable and helpful to the people and government of the host country. They are likely to be less successful if they function, or are perceived as functioning, as the long investigative arm of headquarters organizations, or as inspectors from the foreign governments which are helping to finance development activities.

At various stages in a development project, new needs for data may arise, and different methods of collecting data may be feasible and considered worth the effort. These differences are discussed below in temporal sequence, as follows: before the project is initiated; during the course of the project; after the project comes to an end.

Before, during and after, in real life

That every project has a before-phase, a during-phase and an after-phase seems obvious. But in actual practice, even for minor projects, it is not always easy to distinguish these phases.

For example: in 1952 a physician at A started a private health project on his own. He even erected a centre for it. In 1954 legislation was enacted providing that similar projects were to be established all over the country. In 1956, the first project, at A, was made official, and the facilities there

were to be expanded. In 1958 three physicians were working at A, but no trained nurses were available. In 1960 the nurses were there but two of the physicians had dropped out and had not yet been replaced.

Question: what in this project is before, what during, and what after? Any of the years mentioned might be used to mark the 'beginning' of the project. Making such a decision depends on circumstances (e.g., what information is wanted; what records are available) and on the choice of operations to be evaluated (e.g., the physician's original centre, the professional staff at A, the whole government programme). Moreover, in a literal sense there will be no after, as the project is intended to continue indefinitely. What can be done here is select for evaluation one activity or a couple of years of activities from the whole, although selecting too limited an activity or period may not be very revealing. In fact, it may not be worth the effort and expense required for a systematic evaluation unless substantial results can reasonably be expected from the activity or within the period. If the project is expected to affect people's attitudes, substantial change can hardly be anticipated except after considerable time.

Thus, the terms 'before, during and after' will in the following pages be used to mean what has been chosen for good reason as being before, during and after. Moreover, while an ideal evaluation would cover all three phases, a careful evaluation study can be of much value at any phase. Even if such a study takes place before the project begins, it will be valuable in providing base-line data and indicating the nature and magnitude of problems to be solved.

Before the project is initiated

Prior to the initiation of the project (or, if possible, even before it is agreed upon), it is most helpful to carry out six types of inquiry:

1. *Detailed review of other similar projects.* Projects in the same country, or in other countries with similar conditions, should be reviewed to find out what was done, what effects and side effects were produced, which specific operations gave the greatest (desired) results at least cost (including undesired results), what data were collected to indicate project results, how the responsible officials and the people directly affected now feel about the project, etc. This review may be based on carefully prepared studies; it may involve visits to these other projects; or, more likely, it may simply draw upon the experience and interpretation of informed individuals who are available nearby or who can be brought in as consultants. In any case, it can be very valuable both for the guidance it gives in planning the new project and—

which is its pertinence here—for the leads it provides in deciding what kinds of data about what kinds of changes it will be feasible and desirable to collect.

2. *Detailed discussions of the proposed project.* Detailed discussions should be arranged with persons informed about the kinds of changes sought and the kinds of people who will be affected by the project. Local community leaders, social scientists and officials of government agencies or private organizations may be most helpful in such discussions. They would presumably cover, among other subjects, the likelihood of producing various side effects, and the kinds of data that it will be feasible and desirable to collect.

3. *Exploratory interviews.* Some of the people likely to be affected by the project, or at least a few leaders in communities likely to be affected, should be interviewed. If the proposed project is discussed at an early stage with such people, those interested in doing so can collect some baseline data (see pages 70-71) on existing conditions and attitudes; can obtain significant insights into possible effects and side effects; and can learn about already available data, or about individuals or organizations which might be willing to collect simple data that would be helpful in indicating the amount and pace of changes brought about by the project. Moreover, the very questions asked will help clarify local understanding of project purposes; and pointing out that local people will be consulted in the future will both heighten present interest and prepare the way for future co-operation in data collection.

4. *Collection of already available data.* These data should indicate conditions as they exist before the project starts. Such baseline data are often very helpful in project planning as well as in making possible before and after measurement. Moreover, knowing what is already available and what is already being collected greatly facilitates decisions about what additional arrangements need to be made.

5. *Arrangements for individuals or organizations to collect additional records.* Such arrangements should obviously be made as early as possible.

6. *Systematic observation and systematic questioning.* If these are to be undertaken in connexion with the project, it is helpful to initiate them before the project begins, again to get baseline data.

The first three suggestions describe activities that contribute as much to project planning as to measurement of effects. They are usually undertaken

in connexion with every development project. What is suggested is not a large additional expenditure of funds and staff time, but a slight adaptation of standard planning processes to include the planning of data collection. It is possible that this adaptation may also result in improving the planning of the project activities themselves.

The other three suggestions involve some additional staff time and perhaps additional funds. A major project may justify a major effort along these lines. Even a minor project will usually justify some attention to them.

The third, fourth and sixth suggestions indicate the importance of obtaining baseline data. This point deserves some elaboration.

Even if the only data available are those collected after the project has come to an end, these are usually adequate for at least partial evaluation. Evaluation is more informative and convincing, however, if baseline data have been collected at the start. In this way, the range of possible before and after comparisons is greatly increased. Moreover, collection of baseline data helps in the establishment of continuous project records. Such records are particularly valuable if, as frequently happens, there is considerable turnover of personnel during the course of the project.

The concept of baseline data used here, incidentally, is much more limited than that used in connexion with some development projects, where a baseline report may cover relationships with local government units, adequacy of physical facilities, number and competence of co-operating personnel, etc. (see page 73). Here the term baseline is used to refer to data obtained from observations, records, interviews, etc., made before the project starts, these data being limited to those that throw light on the situation the project is intended (or expected) to change.

Baseline data designed to permit later measurement of project results are most valuable if collected early, before any operations are initiated. In fact, some basic economic and social data should, if possible, be collected even before final plans are crystallized. However, the selection of the location of the development project and the choice of control or comparison areas should usually precede the collection of baseline data. Otherwise, the data needed may fail to be collected in some of the most important places.

A significant by-product of obtaining such baseline data is that they help define the dimensions of the problem to be solved. Baseline data show where we are now, as a starting point for where we want to be at some time in the future. They make possible a full and detailed statement of objectives, the need for which was discussed above in Step One. By showing how much illiteracy there really is, or how much malaria and where, they may dictate changes in the plan for project operations previously drawn up.

Moreover, in collecting baseline data factors may be uncovered which

were previously unknown, such as local practices or resistances to change. The proposed innovation may be culturally unacceptable, or may involve small but intolerable risks. Discovery of such factors may well influence the final plans for operations.

For example, a handicrafts expert was sent to an Asian country by an intergovernmental agency to help develop and apply improved techniques in a number of handicraft and cottage industries. His work established which of the industries chosen could be successfully introduced, and also showed which industries it would be worthless to pursue, either owing to material difficulties or because the workers did not accept them for various reasons. Some of the reasons for disinterest among the local people in certain industries were that the work was too widely dispersed over a large area; that no material incentive was offered; that there was a certain rivalry between the old inhabitants of the area and new colonists; and that there was some psychological resistance to handicraft work, considered to be a 'lower caste' activity. In this case, discovery at an early stage of these psychological resistances provided very useful guidance in subsequent promotion of particular handicraft industries.

Conversely, psychological barriers which had apparently not been fully anticipated limited the success of a venereal disease control centre which was established in another Asian country, with several staff from another intergovernmental agency. This project was generally successful except for health nursing and contact tracing, which was impeded by local customs and traditions. The traditional shyness and reluctance on the part of the people to co-operate in the control of venereal diseases emphasized the need of health education of the public before any real success could be hoped for. It is possible that full prior appreciation of these difficulties might have dictated a different balance and timing of project activities. For example, if an extensive public health education campaign utilizing radio, posters, etc., had been the first activity undertaken in this project, these psychological barriers might have been reduced.

Finally, baseline surveys may show that psychological barriers assumed to be present are really not there at all. For example, Japanese studies have shown that recent dissemination of contraceptive information was welcomed by the population, to the astonishment of administrative authorities who had believed that Japanese farmers desired large families.

The kinds of questions which may profitably be investigated before drawing of final plans for a development project are indicated by the following list (drawn up in connexion with a proposed change in material technology):
1. What, if anything, will the introduced trait replace?
2. What other tools and techniques are likely to be modified as a result of the introduction?

3. What other tools and techniques will have to be modified if the new trait is accepted?
4. For what new tools and techniques is there likely to be a demand as a result of the introduction?
5. Who in the society will have to abandon or change their occupations, if there is replacement? Who will be likely to modify their occupations and who will be given new occupations?
6. Who within the society will benefit immediately from the introduction? Will the benefits be in terms of economic advantage, increased prestige, or what?
7. Who, if anyone, is likely to suffer immediately? In terms of real or fancied threat to economic security? In terms of social status?
8. Will shifts in occupation affect the division of labour between men and women?
9. What are the formal and informal social organizations which those affected participate in and consider important?
10. How are these social organizations likely to be affected? Will their power or social position be enhanced or lowered?
11. Is there a possibility of the introduction opening up new forms of co-operation? Of conflict?
12. Do the individuals and group leaders affected understand the nature of the introduction?
13. Who has participated in the planning of the change? Who has not participated?
14. What customs, other than the technology, are likely to be affected? Food habits? Relations of young and old? Marriage customs? Ceremonials? Religious beliefs? Major values?
15. Does the change reinforce these customs or conflict with them?
16. What are the attitudes toward the innovator (the field worker) as a person? Toward the ethnic group of which he is a member?
17. What is the recent history of the relations between the introducing group and the people?
18. What is the history of similar introductions to this group?[1]

In this connexion, it should be noted that reactions to a new project are frequently unfavourable, as any change tends to be uncomfortable. Reactions of this kind may indicate the need for better psychological preparation and fuller explanation during the project, and for repeated assessments after the project has been in operation for some time.

Resistance to change commonly appears also if the change: (a) is perceived

---

1. Edward H. Spicer. *Human Problems in Technological Change*, New York, Russell Sage Foundation, 1952, pp. 91-2.

as threatening basic security; (b) is not understood; (c) is forced upon people.

Consequently, it is very useful, especially in the early stages of a project, to ask questions which will bring out whether or not the people affected: (a) feel threatened by the project (for example, threats of loss of status, disintegration of the family, reduction of employment or income); (b) understand the project (its objectives, operations, expected by-products); (c) feel compelled to make changes without adequate opportunity to consider them, discuss them, and influence the decisions finally taken.

Other questions which may be investigated to throw light on popular co-operation or resistance are the following:
1. Have the cultural linkages been discovered and utilized in the procedure?
2. Has work been carried on through the existing social organization or have social organizations been set up which conflict with those previously in existence?
3. How are the innovators' purposes and ways of behaviour regarded?
4. Have the innovators misinterpreted the responses of the people through cultural bias?
5. Has the maximum possible participation been encouraged and allowed to develop?
6. Have the needs that are felt been obscured by opposition to methods of introduction?[1]

Some organizations, which provide experts to assist in development projects, require each of their experts to prepare what is called a baseline survey or a work plan shortly after his arrival in the field. This is usually an analysis, developed in collaboration with responsible government officials and others, of the existing situation in the project area and of the activities to be undertaken and completed within a given period. It usually includes a clarification of the needs, demands and expectations of the responsible government agency, and is directed towards a specific plan of action.

Such a procedure has real advantages. Every expert, of course, makes some preliminary assessment of the field situation within the means at his disposal. To introduce the work plan as a formal responsibility requires the expert to make a more thoughtful and orderly exploration of the prevailing situation than might otherwise be the case. Further, the procedure provides the expert and the agency administrator with a set of benchmarks, however subjective, in relation to which progress can be reviewed. One possible danger of this idea is that the agency administrator may require the expert to gather types of information which are not readily accessible, and which it

1. Edward H. Spicer, op. cit., p. 294.

may involve considerable work to obtain, thus detracting from other work by the expert. Some type of benchmark survey or work plan is, nevertheless, part of the process of getting the project started, and to introduce it as a formal procedure may prove of value both to the expert and to the agency administrator.

During the course of the project

It is usually desirable, during the course of the project, to discuss operations and apparent results, from time to time, with persons informed about the progress of other projects, particularly ones involving similar activities. Information on results being obtained elsewhere may well direct attention to kinds of data that might profitably be collected during the project in question.

A number of agencies have found it helpful to obtain periodic reports, partly to keep responsible officials informed of progress (and difficulties), partly to indicate to headquarters officers how they may help field personnel do a better job, and partly for the educational value that preparing such reports may have for the field personnel themselves. Such reports are concerned with questions which are continually in the minds of field personnel, and periodic reporting provides an opportunity for them to organize their thinking and to review systematically the current status of the project. The development of standard questions to be covered in such periodic reports, especially if the data required are of kinds easily collected and summarized, can be a major step towards better measurement and understanding of project results. Quarterly reports, one being drafted as an annual report, are frequent enough for most purposes.

One international agency instructed its experts to submit quarterly reports organized under the following headings: (a) project objectives; (b) physical conditions of work; (c) personnel; (d) liaison activities; (e) description of work; (f) public information; (g) miscellaneous information: (h) evaluation of the project; (i) projected work plans. These topics are expanded somewhat below.

(a) Project objectives

To contain a precise statement of the short-term objectives of the project. These objectives may change from time to time according to local conditions and may differ markedly from the concept of objectives held at headquarters. In many cases a precise formulation of objectives must be worked out by the expert himself operating under field conditions.

(b) *Physical conditions of work*

To include an account of physical plant, transportation, housing, equipment, financial arrangements, and so on, and a discussion of problems relating thereto.

(c) *Personnel*

To include an account of changes in personnel, the number and effectiveness of counterpart personnel and trainees, arrangements regarding fellows, and related personnel matters.

(d) *Liaison activities*

To include an account of the nature of liaison and relations with government officers, technical assistance experts employed in the area, visiting experts or officials, any individuals, groups, or organizations with whom it is necessary to establish contact for the proper conduct of the mission, and related pertinent information.

(e) *Description of work*

To include a precise description of the work done on the project and the actual operations performed during the period, excluding items listed under *b, c* and *d* above.

(f) *Public information*

To include an account of publicity given to the project. Newspaper clippings, photographs, accounts of public lectures, and related matters may be included.

(g) *Miscellaneous information*

To include any items of information not included under the above headings which the expert feels are important and of particular interest. This may include an account of specific work methods or techniques which have proved to be particularly effective, or ideas and suggestions for the improvement of the project.

(h) *Evaluation of the project*

To include an appraisal by the expert of results achieved during this period in relation to the stated objectives of the project, and descriptions of actions taken to evaluate results. Information in this section should include not

only the subjective appraisals of the expert regarding the effects of the work in progress, but also where possible concrete factual information, sometimes in statistical form.

(i) *Projected work plans*

A factual and precise description of work to be done during the next period.

Periodic reporting of physical conditions, the personnel involved, the expert's own activities, and the actual operations carried out (items *b*, *c*, *d* and *e* above) is particularly important, because actual operations often differ widely from what the original plans called for, from what the headquarters staff believe are going on, and from what retrospective accounts describe. These differences arise from misapprehensions or wishful thinking on the part of the staff, differences of interpretation about the real objectives of the project, conflicts between agency conceptions and local law or custom, unavailability of particular personnel or resources, etc. Because differences between planned and actual operations often arise, project results cannot be intelligibly assessed without accurate, factual, first-hand information on what is actually being done under the project. This is what makes periodic reports from field personnel so valuable.

Reporting the cost of each operation is important also. Excessive detail is rarely needed, but enough detail should be given to distinguish major categories of activity. For example, how much time is spent by various personnel on different functions, and what out-of-pocket costs are involved? Translation of expert and administrative time into money costs can usually best be undertaken at headquarters, to ensure consistency. For field use in supervision and for decisions on project administration, records of actual time spent are helpful.

Some development projects have had substantial effects on the people they reached, but actually reached so few people that the cost of producing these effects far outweighed the benefits. Unless at least rough accounts are kept of time spent and all major costs, conclusions about the value of such projects relative to costs may be difficult to make.

In addition to reporting on what they themselves do or observe, field personnel should report data obtained in other ways. Exploratory interviews may be made from time to time with some of the people affected by the project, or at least with some leaders in the communities affected; with persons engaged at various levels in project operations; and with other people informed about the project. Such informal questioning is usually undertaken in any case to see 'how things are going'. If questions about desired and undesired project effects can be included, ways of collecting

additional significant data may be suggested. Reports may also include already available data and utilize additional records kept by organizations. Finally, they may be based upon systematic observation and systematic questioning.

The collection of data on project results may result in changes in project operations, usually the primary purpose of collecting such data. When project operations change, it may be appropriate to collect different data on project results, or to apply different procedures.

Data should be collected and analysed as quickly as possible, for stale data lose their interest for operators, as well as their value as indicators of what changes are needed in operations. Because timeliness is so important, the amount and kinds of data collected during the course of a project should be held to the minimum that will provide adequate guidance. This tends also to limit cost.

Cost can be further decreased if it is possible to collect data on more than one project at the same time. For example an interview with a government official or other informed person might cover several projects. There are opportunities in such interviews to get from each person interviewed comparative appraisals of the success of the different projects, as well as their ideas about the reasons why one project is more successful than another. Similarly, if two or more projects affect a given group or community, information on all can be obtained at the same time. Cases where this is possible are rare, but may well permit some economies of time and effort.

Methods of collecting information should be as quick and simple as possible, and the data collected should not be concerned only with the effects of the project as a whole. Data on results are most useful in administration if they indicate the specific effects of specific operations. For example, has increased use of fertilizer by rice farmers in a particular district come about as a result of the distribution of pamphlets, of the work of demonstration farms, or of the efforts of extension workers who made personal calls on farmers? Questions will have to be asked directly of some of the people affected by the project in order to assess the results of specific project operations. Data obtained without questioning may, however, be adequate to indicate what effects the project is having as a whole.

An interesting example of tracing the impact of a specific operation comes from rural development work in India. At the Etawah Pilot Project in Uttar Pradesh, it was decided to try to introduce green manuring on a large scale. A supply of sanai seed was obtained for the co-operative store, and a vigorous campaign was carried on to educate the farmers. Day after day passed, but very few farmers arrived to buy the seed, and those who bought it came almost solely from within a mile of the store. In an effort to advance the project, the staff workers took the seed from the store and delivered it

to individual farmers who were willing to buy it, and left a modest quantity in every village with one of the respected and responsible villagers, who were willing to sell it at any time of the day and on any day of the week. A hundred and seventy-five bushels were distributed to farmers in this way and were sown in the fields. The resulting green crop was turned under while standing, and the subsequent wheat crop yielded about 40 per cent more than previously. The next year at the time for the sowing of sanai, there was a great rush to the store for this seed. More than five times as much seed was sown as in the previous year.

A limited amount of direct questioning revealed that the advantages of the use of the seed recognized in advance by the farmers were just about balanced by the disadvantages. Suspicion and dislike of government officials in general was traditional, they feared the risk involved in devoting money, effort and land to growing a new crop, even for manuring; and most of them had to travel a long distance to the store. The single operation of bringing the seed to them, together with the symbolic significance of government workers showing a real interest in the farmers, tipped the scales.[1]

After the project comes to an end

There is usually a keen demand for data, after the project comes to an end, to show what results have been obtained. Headquarters offices and supporting organizations want to know 'what they got for their money'. There is wide interest in the lessons learned that may be applied in other projects, yet the opportunity to collect data is at this stage very limited.

If earlier arrangements for record keeping have not been made, and if certain kinds of data have not already been collected, it may be too late to obtain entirely convincing evidence of results. Some useful new steps may still be taken, however, at this stage.

For example, it may prove quite profitable to review reports of similar projects undertaken elsewhere, to find out what results were obtained, and what data and methods of analysis were utilized to provide the evidence of these results. Similarly, discussion of the project with those planning other projects may well suggest data that should, if possible, be collected, to help elsewhere. Such reviews and discussions may indicate the desirability of collecting many new types of data, even though project activity has come to an end.

At the end of a project, even if interviews have not been carried out

1. Adapted from Edward H. Spicer, *Human Problems in Technological Change*, New York, Russell Sage Foundation, 1952, pp. 55-67, 'Case 4—The Village Level', by Rudra Datt Singh.

earlier, it may be very helpful to do an 'experience survey', asking persons who participated in the project, as well as persons presumably affected by it, what they remember was done, what happened as a result, what was most significant, and so on. This is, in fact, the way most data on project results are now obtained.

While it is too late to arrange for new record keeping by individuals and organizations after a project has ended, it is not too late to make a careful investigation of records that have been kept. Intensive review of the evidence available in the records kept by project staff, and visits to the other individuals and organizations which might be expected to have kept records pertinent to project results, may be valuable in turning up existing data that may now be analysed.

In particular, this is the time to prepare a definitive record of exactly what project operations were actually carried out. Usually this record will differ considerably from the plans originally made and it is therefore essential in order to understand and appraise results.

It is especially at the end of a project, and sometimes a good deal later, that side effects tend to receive attention. If these side effects have not been anticipated, observations or interview data bearing directly on them have probably not been obtained during the project. If they have appeared only at a late stage in the project, data collected earlier may not show them as being significant in amount or prevalence. Observation and questioning about these side effects are therefore particularly important at the end of a project.

The following example indicates the variety of side effects that may be produced by an apparently simple project operation (providing one wagon to an isolated American Indian village). Data concerning these effects and their relationship to the project were obtained by interviews carried out after the project had ended.

The modern technique of wagon transportation was introduced to the Papago Indians of the village of Choulik, in Southern Arizona, U.S.A., by the simple means of providing them with a wagon and showing them how it worked. The direct and indirect effects were manifold. There was a decline in the handicraft of harness, pannier and pack-saddle making, partly because the wagon harness was purchased and partly because transport by pannier and pack decreased. A local man learned how to work iron, in order to repair the community wagon. The making of small earthenware water containers declined, being largely replaced by larger metal containers which, though they had to be bought, held much more water. The wagon made it feasible to transport them. Better roads were built, and the village people in the process developed engineering skills and a new pattern of group labour. Trading trips could do so much more business that fewer trips were made,

with less persons involved, so that face-to-face contacts with neighbouring tribes declined, despite a larger volume of trade resulting from the increased cash sale of wood and crops made possible by lower transport costs. A new communal activity developed in planning and administering the use of the wagon, which was treated as a community resource and often let out in turn to each family, and in planning and directing the building of roads. All this resulted from a single project 'operation', in the sense used above.[1]

This example suggests a final point. While it is usually desirable to collect many kinds of data immediately after a project has ended (because effects may be obscured by subsequent change, records may be lost or destroyed, individuals may move away or die, and memory may fade), a strong case can be made for collection of some data at much later dates. Just how much later is a matter of judgement, related to a variety of factors. In many projects however, data might well be collected from one to three years after the end of the project, and then a second time from five to ten years later.

In the period from one to three years later, certain indications of permanence and spread of changes induced by the project may be obtained. In this period also, certain side effects, like those described above, may appear.

But only after five or ten years can convincing evidence of permanence and spread be obtained. Because permanence and spread are major criteria of the success of development projects, any 'final' conclusion that a project was worth the effort and funds devoted to it can hardly be drawn until some such period has elapsed.

---

1. Adapted from Edward H. Spicer, *Human Problems in Technological Change*, New York, Russell Sage Foundation 1952, pp. 23-32, 'Case 1—In the Wake of the Wheel', by Wesley L. Bliss.

Analysing and interpreting findings

Analysis

Analysis takes time, of course, especially when many types of data have been collected. This can be shown by a simple mathematical law. If one collects two sets of data, one can make only a single intercomparison, comparing one with the other; if one collects 3 sets, one can make 3 comparisons; for 4 sets it is 6; for 20 it is 190; and for n data it amounts to $[n(n-1)]/2$.

Of course, not all of these intercomparisons (in pairs) will be meaningful, but some of them are. Therefore, one should only collect many different kinds of data if it is going to be possible to analyse them adequately. (In addition, there are also ways of analysing multiple relationships rather than simple comparisons in pairs—but this requires some specialized skill and for most purposes can be ignored.)

In this section, attention will be paid to certain general aspects of analysis. More detailed aspects (proper classification and coding, proper tabulating, etc.) are dealt with in the appendix. As pointed out earlier, deferring these details for discussion in the appendix does not mean that the topic is unimportant. On the contrary, the difference between a good evaluation report and a poor one is frequently in the soundness of the statistical analysis.

Let us assume that it has been possible to collect considerable data showing that changes of the kind sought in a development project have in fact taken place during the course of the project. Let us also assume either that baseline data were obtained at the outset of the project, or that available records give a reasonably satisfactory picture of conditions at that time, or that the memory of informed people can be relied upon to give such a picture. It is feasible, then, to make some estimate of the magnitude of change during the period in which the project has been in operation.

Does this mean that the change so estimated has resulted from the project, itself? Or is it possible that it has resulted from some other cause?

Many changes are so obviously a result of project activities that this question does not arise. On the other hand, sometimes there are valid questions about apparent cause-and-effect relationships.

For interpretations about cause and effect to carry conviction, it is necessary to do more than simply measure changes of specified kinds in a specific activity, area or group. It must also be shown, by some logical process, that these changes are related to, or result from, the particular operations performed in the development project.

There are two principal methods of showing such a cause and effect relationship. The first is by use of a 'control' group, area, or activity. This is a group, area or activity separate from but comparable to that which the development project is intended to change, and not subject to the operations of the development project. If both the 'control' and the development target show some change, this can hardly be attributed to the operations of the development project. But if the development target shows the change and the 'control' does not (or else shows a significantly smaller change), and if the only observed difference in circumstances likely to produce the particular change is the development project itself, a good case can be made that the latter is responsible for the different development of the two groups.

The second method is similar. In this, the 'control' is really the nation, or else some group or region or activity which includes the development target and yet is large enough so that the development project could not be expected to influence more than a small part of it. If both the development target and the more inclusive entity show some change, again this can hardly be attributed to the operations of the development project. But if the development target shows the change and the larger entity does not (or else shows a significantly smaller change), and if no *other* development activities have been impinging on the development target but not on the larger entity, a good case can be made that the development project is responsible for the difference in observed change, as between the target and the larger entity.

Having good base-line data for the development target, and for either kind of 'control' discussed here, obviously makes measurement of changes easier and attribution of effects more convincing.

Comparisons without baseline data are less satisfactory because they must be based on the assumption that the situations now being compared were not significantly different before the project operations began. This assumption is not always valid but it can be controlled and validated; it is indeed preferable to make this kind of comparison rather than to make no com-

parison at all between activities, areas or groups which have been affected by the project and those that have not.

Suppose that areas X and Y are in general comparable agriculturally and are far enough apart so that little intercourse takes place between them. Suppose also that a one-year campaign to promote the use of commercial fertilizer has taken place in area X. If records of fertilizer sales by dealers are available, a table like the following might be drawn up:

TABLE 5. Annual sales of commercial fertilizer, in tons

Area	Year before campaign	Year following campaign
X (had campaign)	1 000	1 500
Y (no campaign)	800	1 000

If the data in the first column (the baseline data) were not available, the figures for the year after the campaign might be interpreted to show that, since area X and area Y were on the whole comparable, the campaign had increased fertilizer sales by 50 per cent in area X. If the data on area Y (the control area) were not available, the before and after figures for area X alone might be interpreted to show the same magnitude of result. But knowledge about the concurrent change in area Y suggests that some general factors (perhaps improved education, or lower prices for fertilizer, or higher crop prices, or changed legislation concerning land rent, etc.) were operative in both areas at the same time that the campaign was going on. Under the influence of these factors, area Y increased fertilizer purchases by 200 tons, or 25 per cent, even without a special campaign. If these factors were equally operative in area X, sales of fertilizer there might have been expected, even without a campaign, to increase by 25 per cent likewise, or from 1,000 tons to 1,250 tons. Thus only half of the increased sales in area X, or 25 per cent of the pre-campaign sales, could be attributed to the effect of the campaign.

Of course, if area X had been affected more than area Y by the general factors operating during this period, an even smaller percentage of the change in sales in area X might be attributable to the campaign. For example, perhaps there were more sharecroppers in area X, and a nation-wide reduction of land rents therefore benefited area X more than area Y, both by increasing its income and by increasing the economic return from investment in fertilizer. Or adult education classes might have been expanded faster in area X than in area Y, although no link with the fertilizer campaign had been

intended. Cause-and-effect analysis in a period of active development can be very complex.

Finally, it is often illuminating to compare two similar development projects that differ in only one or two significant ways. Comparing the results of the two projects having the same purposes and generally similar project operations may throw a good deal of light on the effectiveness of the particular operations that are carried out in only one of them.

Where considerable masses of data have been obtained on project results, refined statistical procedures may be used for analysis. Such procedures are described in the references in the bibliography. For many purposes, quite simple methods of treatment are entirely adequate.

One method is graphic presentation. It is often easier to interpret results, and the interpretations are often more convincing, if the data are presented in graphic form. For example, the data on fertilizer sales presented above might look like this on a bar chart.

Annual fertilizer sales
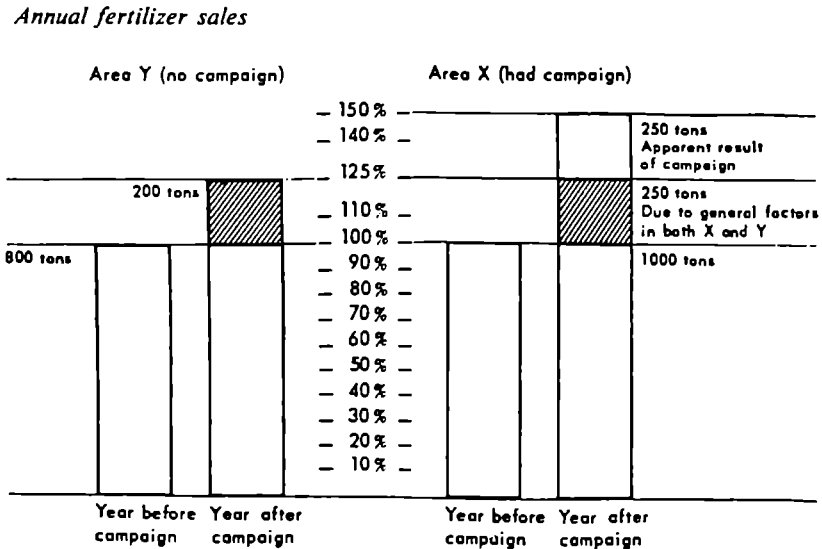
Such a chart often conveys more about relative amounts, particularly to visually minded people, than does the table on which it is based.

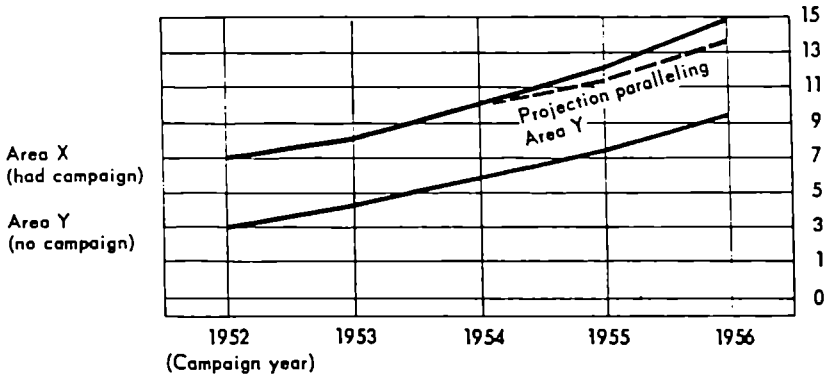If to these data are added annual sales for other years, the records of area X and area Y might be compared over time, as follows:

Annual fertilizer sales

Area X
(had campaign)

Area Y
(no campaign)

1952
(Campaign year)   1953   1954   1955   1956

Clearly evident here are: (a) the yearly trend upwards in fertilizer sales in both areas; (b) the 1955-56 spurt in sales in area X, where the campaign took place; (c) the size of the spurt in sales, relative to the growth that would have been expected if the growth of area X's sales had simply kept pace with the growth of sales in area Y.

Of course, it is often impossible to obtain good statistical data indicating project results. Moreover, even when statistical data are adequate to measure certain changes, other and perhaps more significant changes may be evidenced only by replies to informal and spasmodic (although carefully thought out) questioning. How can cause and effect be discovered in such cases?

The logical process of comparing changes in two groups, or comparing current conditions in two groups, one affected and the other not affected by the project, is the same for data obtained by questioning as it is for recorded data. The possibility of generally operative factors producing the observed change or difference must just as rigorously be ruled out.

One way to test whether particular effects are the results of particular project operations is by discussion with the various groups of people concerned with or affected by the project. The several experts and administrators connected with any project are continually forming and checking their own estimates of project results, of course, in order to guide them in their own project activities. They are in an excellent position to confirm or deny the likelihood that particular changes have in fact resulted from project

operations. Moreover, any information concerning project results they can get from other persons is helpful to them.

The participants in the project (the teachers being trained, the village development workers, etc.) and the broad range of persons affected by the project (the families whose children are taught by the new methods, people in the villages where change is being fostered, etc.) are also directly concerned in the results of development projects, and interested in interpretations of cause and effect. Reporting results and interpretation to these people is very important. Such reporting builds up their interest and thus reinforces the impact of the project; and it gives them the information they need if they are to play their part as responsible citizens in a democratic society.

If such results are to be understood and utilized by them, they should of course be presented in simple and direct form. Opportunity should be given for full discussion of results—how they were obtained and analysed, and what they may imply. In this way the process of reporting and analysing can itself stimulate awareness, discussion and common effort, and can help to establish democratic ways of living and solve actual problems. Moreover, the people involved can often throw new light on the data obtained and on inferred cause-and-effect relationships. They have lived the project, and their perceptions of what has been going on and what produced certain changes can be of great value as a check on the data and interpretations of those responsible for the project but perhaps less intimately connected with it. If these data and interpretations do not make sense to the people most directly involved in a project, or to those best informed about it, the case for re-working the data and re-thinking the interpretations will be strong.

Interpretation

It is apparent above that interpretation of findings is deeply inter-related with analysis. Both are involved in the writing of the report. But interpretation goes beyond statistical analysis to include other factors of great importance, such as the specific operations undertaken (described in Step One), current economic and political events, and the general circumstances surrounding the project and affecting the society.

Interpretation and evaluation may be limited rather narrowly to the project itself, the circumstances surrounding it, similar projects undertaken, and 'controls' employed. Interpretation and evaluation may, however, range much more widely, taking into account national and international

conditions and trends. But even when rather limited interpretation and evaluations are undertaken, the more general conditions of society cannot be neglected altogether, as the following example shows.

Interpretation and society

A study of the effect of literacy classes for adults revealed that, after a lapse of some years, the adults who had taken these classes lost much more of their active literacy than had been expected.[1] (This is, by the way, a common phenomenon—that changes do not persist in full. It is also common to expect greater changes than actually take place.) In interpreting this finding, it is natural to ask if the literacy courses given were long enough, if the teaching methods were adequate, and what follow-up measures might be devised. But broad social factors should also be considered. For example, to what extent does the society itself reinforce the skills of literacy, once learned? Is literacy important and pervasive in the society? Is it necessary in daily work (written instructions, printed advice, printed public notices, etc.); in consumption and private life (advertising, government bulletins, health notices, etc.); in relaxation (newspaper in every home, printed schedules of radio programmes, travelling by modern means, etc.)? Literacy is not just a product of a differentiated and wealthy society; it is an essential tool for getting ahead.

In a society already somewhat developed, a man who cannot read and write cannot have a good job, run a farm well, manage even a small business, keep up with the news and converse with his neighbours, and so on. Therefore, everybody in such a society wants to read and write. In a society that does not yet use literacy widely, there is little compulsion to become literate. In fact, there may even be barriers to literacy, in that 'wasting time' on reading and writing may be considered a form of idleness, and hence frowned upon.

Thus, when one evaluates literacy education projects, it is of top importance to consider what role literacy plays in the society. Longer courses, better teaching methods and better follow-up will certainly help extend and maintain literacy skills. But compulsion and reinforcement by society are fully as important, if not more important. Expectations of success in literacy courses should take this social factor into account, and thus be realistic. And the literacy courses should be seen as parts of a much bigger and more general development process, affecting the whole society.

1. Presented at the Cairo seminars on evaluation in 1963 by Mr. Abdel Aziz Kamel Sakr.

Closing and shaping
the circle

In this final chapter, the first topic is one that logically might have been placed at the beginning of the manual, but that now becomes more meaningful in the light of the series of steps discussed. This topic concerns the planning of an evaluation study, particularly with respect to its time dimensions.

It may well happen that an evaluation study is still in the collection phase when time begins to run out. Obviously, insufficient time is then left for thorough analysis and thoughtful interpretation, as well as for the preparation of the report. This can be avoided by setting up a time schedule, at the beginning of the evaluation, for the different steps of the study. Thus the evaluation circle can be shaped into a well-conceived form.

Planning an evaluation

Suppose that an evaluation study must be completed and a report prepared in a year's time. The best plan, especially for a beginner, is to aim to finish a draft of the report in eight months at the most, or preferably even two months earlier. Drafting the report makes apparent what conclusions can already be drawn and what is still lacking. It also makes it possible to estimate how much additional time will be needed to carry out a more thorough analysis of the data (or even to collect some missing types of data).

The first evaluation that anyone undertakes naturally takes longer than later ones. At a minimum, it is likely to take several months, if a project of some importance is being studied. Later studies take less time.

Partly, this is a matter of learning from experience. Partly, it comes about as evaluation is integrated into routine procedures. But responsible evaluation always requires the allocation of considerable staff time and other expenditures. In the example given below, the evaluation study required

two civil officers to spend three months each on it. Clearly, as considerable costs are involved in evaluation studies, one should undertake only those studies for which the required staff time and other expenditures are available and can be justified.

The project as a contribution towards significant change

In the Introduction, there was given an outline of changes considered especially significant for the processes of economic and social development. Interpretation and evaluation can be of greatest value if project results are seen in the light of their significance for such changes. Thus, the evaluation circle can be closed.

Normally, interpretation at this level requires rather specialized knowledge and competence, and would therefore be most appropriately undertaken by social scientists and professional evaluators. But even the part-time evaluator should realize its existence and relevance. An example will indicate why.

An evaluation study was made of classes for artisans that were held in the Gezira region of the Sudan.[1] In this developing region, a need was felt for craftsmen who could help the population, which was originally semi-nomadic, to build houses and to furnish them. For this purpose, classes for artisans were established a number of years ago. These were two-year courses, in carpentry and in building, both at an elementary level and close to the present needs and customs of the population. The evaluators wanted to know whether or not these classes had been a real success. Their principal approach involved the questioning of three samples of the population: former pupils, the fathers of former pupils, and leaders in the local communities. It appeared that the former pupils were earning their living well, that they had a certain prestige in the community, and that the community was conscious of their usefulness. Therefore, interpreted in the light of the direct intentions of the classes, the evaluators could draw the conclusion that the classes had been successful.

But how significant were the results in broader terms of the society's development? The study disclosed that the pupils had come from all sectors of society in the Gezira region, including the sector with highest prestige (farmers with the larger tenancies). Moreover, the leaders in the community (often farmers, but also others) stated that they would certainly not object to having their sons attend the classes, that they would have their building and furniture-making done by the pupils, and so on. It was nearly

1. Report presented at the Cairo seminars on evaluation by Mr. Abdel Rahman Al Sheikh and Mr. Zeid Musa.

unanimously felt that the classes should continue, that the number of pupils should be increased, and that the curriculum should be expanded. Very few disagreed. From the point of view of development, these findings seem very important. The population of the region of Gezira (and the population the region attracts) is in the midst of re-orientation from semi-nomadic life to a sedentary life with modern characteristics: a society with considerable professional and other differentiation. In the future, this modern society will surely develop, along with professional differentiation, its own social and prestige differentiation. It may reach a stage where leaders will express the opinion that they would prefer that their sons not become artisans. But this opinion was not expressed in this study (although it is a view often expressed in traditional societies), and this is important because it may indicate a flexibility favourable to development. If such a view had been expressed, it could have indicated the existence of a social cleavage which would almost certainly hamper development by directing youths and re-sources into education for 'jobs with higher prestige, thus for example crowding the white-collar occupations and neglecting manual occupations.

This interpretation of the results does not imply, however, that the Gezira region is in fact developing harmoniously. The single study is too limited to justify so broad an interpretation. Still, the study does give indications that are favourable for development, and this is a good example of a concrete and detailed activity having broad significance.

Benefits from improved information

Is better information about most projects really needed? Many a man engaged in development activities says: 'I know what happened. I was there'. He is right, of course. Yet his knowledge and understanding depend on the amount and quality of data available to him. Any increase in the amount and quality of data bearing on the changes in which he is interested can add significantly to his knowledge and understanding. If this booklet helps him get better information than he is now using, it will have served its purpose.

If better indicators of project results can be obtained, they will produce the following direct benefits: better guidance for the community affected, for the expert, for the administrator, and for the responsible organization. Moreover certain indirect benefits are also worth noting here: clarification of goals (which is a prerequisite to the appraisal of results), a better morale of all engaged in development because a better understanding leads to more realism in expectations and so to less disappointment, more relevant reporting in normal administration, and more realistic development of science and teaching.

One of the fears often expressed about any effort to get better data on project results is that it may divert time, energy and funds away from operations. The fact is, however, that well planned record keeping and collection of other data may reduce total administrative costs and increase the efficiency of administration. Current records of operations are essential for good management. Concise summaries of objective data on results are far preferable to the descriptive (and often discursive) reports frequently prepared. They reduce administrative costs relative to results.

Recapitulation of the four steps

For those wishing a 'check list' of the principal actions suggested above, as ways of moving towards better measurement, understanding and administration of development projects, the following recapitulation is offered. As pointed out in the preceding section, it is necessary to plan ahead of time to fit the whole series of steps within a specific time schedule.

Step one

Describe the development project and specify its goals in a written statement covering:
1. Nature of the project and specific operations to be carried out.
2. Kinds of results sought—including impersonal, organizational, personal (information, attitudes, skills, actions, habits, social relationships), and changes in conditions and in capital resources.
3. Side effects anticipated.
4. Area of activity and area of influence.
5. Pace of change.
6. Costs anticipated.
Review this written statement with interested parties, to the extent practicable.

Step two

Decide what kinds of data will be used to indicate project results and how these data are to be obtained:
1. Find out what registration and census data are already available or

already being collected, and decide which can be utilized and how they are to be obtained.
2. Decide what individuals or organizations are to be asked to collect additional pertinent data, including individuals and organizations directly involved in the development project ('built-in' data collection); and what periodic reports are to be requested.
3. Decide what data are to be collected by systematic observation of affected and control groups, and plan procedures for such observation.
4. Decide what data are to be collected by systematic questioning of affected and control groups, and plan the questions to be asked. Prepare standardized interview schedule or questionnaire, if one is to be used.

Step three

Collect the desired data:
1. Beforehand—to help plan the project.
2. During—to help administer the project.
3. After—to help plan future projects.

Step four

Analyse and interpret the findings, and review findings and interpretations with the various interested parties.

Some statistical concepts and elementary procedures

This appendix is provided for those who are interested in the statistical concepts and procedures involved in selecting samples of persons to be interviewed, and in summarizing and analysing the data obtained, and who may wish a somewhat more extended treatment of these topics than given above. Fuller discussions can be found in the references in the bibliography.

Universe

The total of what shows or could show the phenomena considered is called the universe. For malaria this is probably all the population of a region or a country; for a unique professional school this may be a very limited number of trainees and ex-trainees only.

Small universes should—for reasons stated below—most often be studied and analysed in their entirety; large ones can be represented by samples. But the question of what constitutes a universe has its pitfalls. Suppose one is studying specific attitudes towards innovations in cattle-holding and breeding. A sample of thirty from a village may seem very acceptable. But if it can be expected that these attitudes are affected by owning or not owning cattle, and if it is important to find this out, this sample can easily turn out to be inadequate. It is unlikely that two subsamples within a total sample of thirty will show differences which can be considered reliable. Each subsample must be looked upon as a sample of its own universe, and should be large enough to represent that universe.

Subdivisions in small universes are often too small in any case. Suppose the above unique professional school has finished its first course; twenty-five have taken the diploma and five have dropped out during the course. Those five do not lend themselves to statistical analyses. But it may be very worth while to include them in the final interview concerning the course, because they—magnified as it were—are likely to give valuable clues for the interpretation of what is found out in interviews with the twenty-five who finished the course.

Sampling

Many development projects are intended to reach considerable numbers of people. Obtaining information on project results by observing or questioning the entire group of people potentially affected would be wasteful. By observing or questioning only a fraction ('sample') of the group the desired information can be obtained more economically, more quickly, and very often more accurately than by including the entire group.

It is essential, however, that the sample be selected properly, that is, by relying on the laws of chance in the selection process. Otherwise, biased or erroneous conclusions may be drawn.

The two major characteristics of a sample are:

1. Its size. Large samples (over one hundred) are needed relatively infrequently in studying development projects. Many useful kinds of data can be obtained from samples of ten (persons known to be well-informed) or twenty (cross-section of a whole community or region).
2. Its representativeness. A sample cannot fairly represent the whole if the persons in the sample differ substantially from those not in the sample, and if these differences are related to the kinds of data being collected.

Size of sample

Samples must of course be large enough. The fallacies that can emerge from anecdotes of a single instance or from study of a very small group are well known. It is very difficult to avoid bias of some sort in any very small sample. On the other hand, because a development project usually affects a limited area or a limited class of people (for example, teachers in technical schools), and because the persons affected are thus a good deal more homogeneous than when a national survey is undertaken, a sample of twenty individuals (if effort is made to avoid bias in the sample) will provide a good deal of useful information, particularly if one is interested only in fairly marked changes. A hundred individuals, again with efforts to avoid bias in selecting them, will usually be enough to indicate fairly precisely the direct effects of a development project. After all, great precision in the measurement of such effects is unnecessary, because changes to be socially significant must be sufficiently marked to show up under even quite crude measurement.

The size of sample needed is obviously related to the variability of data obtained. If, in interviewing a representative sample of families, the first twenty all said that they spent between 10 per cent and 30 per cent of their income for rent (see Table 6), one could accept this as indicating the range of rent/income ratios for the whole community.

Similarly, if all, or all but one, members of a representative sample of twenty said that they were in favour of building a school, it would hardly be necessary to inquire further.

Where substantially identical or unanimous results are obtained from representative samples, therefore, the samples used can be quite small.

On the other hand, if there is great variation in replies, or if a large measure of

disagreement exists, then substantially larger samples are required. Even here, however, the samples need not be very large unless it is important (which is rarely the case in studying development projects) to get precise measurements or to identify or measure small differences between one group and another. Usually, findings that might be utilized in planning projects, and the project results themselves, are either large enough to show up in small samples, or are not important enough to justify the time and expense of using large samples.

For example, consider the results of a survey of working-class households in York, United Kingdom. The income of each family was determined, and then the rent each family paid was expressed as a percentage of income. Dividing the whole group into five income classes, the following table shows the percentage of income spent for rent by all households in each class. It also shows the percentage obtained from a sample comprising only one out of fifty households in each income class.

TABLE 6. Percentage of income spent for rent

Income class	Complete survey		Sample 1 in 50	
	Total number of households	Percentage	Number in sample	Percentage
A (highest)	1 748	26.5	34	27.1
B	2 477	22.7	67	22.6
C	2 514	19.8	46	18.0
D	1 676	15.8	35	16.9
E (lowest)	3 740	11.3	77	11.5

While the results from the sample differ in each income class by up to 2 per cent from the results of the complete survey, still the same factual conclusions can be derived from the sample as from the whole, to the effect that, in this community, the percentage of income spent on rent has a consistent and important relationship to income level, being more than twice as large in the highest income group as in the lowest. This finding was obtained by using a sample of from thirty-four to seventy-seven households in each income class.

Another important point is illustrated in this table. If the whole sample (in this case 259) is broken down into sub-samples, it is the size of each sub-sample which is at issue. Many survey samples in the United States and Europe are quite large (2,000 to 3,000), partly so that numerous sub-groups can be studied separately, and partly to permit identification of relatively minor differences among sub-groups, or between these sub-groups and similar sub-groups studied in other surveys.

The size required for samples under certain conditions is indicated in Table 7.·

To use this table, one may first 'guesstimate' the percentage to be used in entering the first column or, much better, one may actually obtain responses from a first sample of twenty-five or so. If this sample is representative, their responses will often provide enough indication of homogeneity to permit planning the size of the

TABLE 7. Sample size required when 'margin of error' of specified size is tolerable (calculated for questions requiring only 'Yes'-'No' or 'Favour'-'Oppose' reply)

If approximately the following percentage say 'Yes'	And if it is satisfactory to have a 19-out-of-20 assurance that the 'margin of error' percentage is not larger than	An adequate sample is
50	20	25
	10	100
	5	400[1]
	2	2 400[1]
70 or more, or 30 or less	20	20
	10	80
	5	300[1]
	2	2 000[1]
90 or more, or 10 or less	10	35
	5	140[1]
	2	900[1]

1. This precision and therefore this sample size is ordinarily not needed in guiding the administration of development projects.

final sample. These plans may, however, have to be altered later if it is found that the first sample was not representative.

There are no rules for entering the second column. This is simply a decision one must make, in the light of the importance attached to accuracy of results and the cost of getting enough interviews to reduce the margin of error.

Table 7 can then be used as follows. Assume that 70 per cent of a representative first sample drawn from a village of several thousand can, when asked, give the name of the local village extension worker. Even if the sample contains only twenty persons, one can be 'reasonably sure' (nineteen chances in twenty) that at least 50 per cent of the total population (which may be several thousand) know his name, and perhaps 90 per cent know it. (50 per cent is 70 per cent minus 20 per cent; 90 per cent is 70 per cent plus 20 per cent; the 'true' answer for the whole village is indicated to lie within 20 per cent above or 20 per cent below the 70 per cent figure obtained).

Conversely, if it is desirable to establish the fact that at least 60 per cent of the whole village know the village worker's name, the sample size should be increased to eighty. If 70 per cent of this sample of eighty knows the village worker's name, one can be 'reasonably sure' that this is not in error by more than 10 per cent.

If, in the process of getting these additional responses, the proportion of the sample knowing the name climbs from 70 per cent to 90 per cent, one can stop with a sample of thirty-five or less. With a sample of thirty-five and a response of 90 per cent knowing the name, one can be 'reasonably sure' that at least 80 per cent

of the whole village knows the worker's name (90 per cent with 'margin of error'. of 10 per cent).

As indicated in the example above, the required sample can be small when the results obtained do not differ greatly from person to person. The data on percentage of income spent for rent showed relatively little variation. Answers to a 'Yes'-'No' question vary much more widely, as an individual can really only answer 100 per cent 'Yes' or 0 per cent 'Yes', with no gradations between. Samples used for such questions generally need to be larger than where intermediate gradations of response are possible, although groups giving very high (90 per cent or more) responses in one category can be represented adequately by quite small samples. As indicated earlier, the general principle is that the more homogeneous the population the smaller the sample needed.

One way of reducing the size of the sample needed is to stratify the sample. This means breaking it up into groups, each of which is internally more homogeneous than is the group as a whole (this was done in the study reported in Table 6). A relatively small sample can then be taken from each sub-group, and the results are then weighted by the reciprocal of the fraction comprising the sub-sample before being added together. This method avoids unnecessarily large representation of particularly numerous sub-groups in the total sample.

Sample size does not have to be finally determined before undertaking a survey. A series of small samples (if each is representative) can be studied, and the results examined. If great homogeneity is present, or if successive samples give results which do not vary too greatly to be useful for evaluation purposes, the collection of data can be terminated at an early stage. Even if not terminated, the study of early small samples to provide a basis for the designing of subsequent samples will result in saving the total time and money necessary to obtain results of a prescribed precision.

Representativeness

How to avoid bias? The mistake to avoid is allowing a selection method to be used that depends on factors that are themselves related to the operations being performed or to the effects desired. If, for example, one is measuring the impact of a fertilizer demonstration programme, one should not limit his sample to the farmers who were already buying fertilizer from a dealer at the beginning of the programme. A later survey of the same farmers would miss all the newly converted fertilizer buyers. Similarly, one would not limit his sample of malarial patients to those who came to the clinic for treatment, ignoring those who were too ill, or not ill enough, or too poor, or too afraid, to come to the clinic. In both examples, 'self-selection' would be an important biasing factor.

.Representativeness, or freedom from bias, in a sample depends mainly on proper procedure in the drawing and collecting of the sample. The objective here is to give every person in the whole group, or in each sub-group if the sample is stratified, a known chance of selection. This usually means an equal chance of selection, at least within each sub-sample.

This objective is most important if precise results are sought from a sample of a large and heterogeneous population. When one is seeking only rather obvious

results, in limited and fairly homogeneous groups, not as much attention needs to be given to refined methods of getting a representative sample; but, of course, any action that it is feasible to take to improve representativeness will add conviction to the conclusions reached.

The following are the chief methods by which an equal chance of selection can be attained.

Simple random sampling. One can write on a separate slip of paper the name and address of every person or family who might be selected, mix all the slips together, and then draw the requisite number of slips at random. Sometimes the mix is poor, however.

Systematic random sampling. One can file all the slips prepared above or use lists already prepared for other purposes (if believed to be complete lists), and then count the number of names and decide on the size of the sample. If there are 5,000 names and one wishes a sample of 100, one can simply choose every fiftieth name from the file or list.

If there is any reason to suspect that a list has some repetitive pattern, such as one based on multiples of five, one should of course avoid intervals which are related to the pattern. In the example just cited, if one suspected a pattern based on a multiple of five, the whole list might be divided into 10 sections of 500 each, and within the first section one might take every fiftieth name, starting with the first; in the second every fiftieth name, starting with the second, and so on.

Stratified random sampling. As indicated above in discussing the size of the sample, substantial economies can be obtained by dividing the whole group being studied into sub-groups, each of which is internally more homogeneous than the whole. For example, families living in high income areas may be put into one group or stratum, those in middle income areas may be put into another, those in low income areas may be put into another. Or different communities located near a development project may constitute separate strata. Or the persons known to be favourable to an expert may be one stratum, and those critical of him another.

If the different strata differ in their responses to the questions asked, the results from a stratified sample will be more reliable than those from an unstratified sample. Within each stratum, of course, every person must have as much chance of being selected as any other person.

For many studies, particular groups or strata are more important than others. The whole population of a community might well be stratified; and then 100 per cent of the leaders might be interviewed, along with 50 per cent of the storekeepers, 25 per cent of the other urban dwellers, and 10 per cent of the farmers (assuming that the farmers were the most numerous group). The sample of each group to be studied must meet the tests for size indicated above, but it is not necessary for each sample to include the same percentage of its group. Of course, if the results are to be combined to give single percentages or averages applying to the whole community, appropriate weights have to be given to the results from each sample when calculating the combined figures.

Cluster sampling. The foregoing three methods limit the number likely to come from any one family to one or two persons. In studies in which it is desirable to interview the whole family, or the head of the family in every case, use can be made of lists of families (or other natural groupings) and then one or more individuals can be selected from each family. Lists of villages may be similarly used. This method requires expert statistical guidance, and is mentioned here only to indicate its availability and pertinence.

Area sampling. This is another technique that requires expert statistical guidance. If enough is known about the geography of the area primarily affected by the development project, it may be possible to lay out the area in a substantial number of separate sub-areas, preferably of approximately equal population size. If areas are unequal in population size, attempts should be made to estimate the relative population size of each. Samples can then be drawn from each sub-area, in proportion to its size. Within each sub-area, however, some rigid technique must be used so that the sample is not made unrepresentative by containing just those people living on main roads, or just those on the first floor of apartment buildings, etc. Blindly sticking pins at random into an aerial map of a sub-area is one technique sometimes used to force representation of less accessible families. Travelling specified distances in specified directions within each sub-area is another.

Judgement sampling. Samples which do not rely on the principles of chance are usually called 'judgement samples'. The person drawing the sample may consider that official X or community leader Y should certainly be included in the sample. If each selected person is considered to be a separate stratum, of course, and is given his appropriate stratum weight in the final result, this introduces no bias. If, however, such judgements result in certain persons being represented and certain others excluded without appropriate adjustment in the weighting process, the results will not be representative.

Moreover, if judgements are made that, for example, 25 per cent of the sample should be *entrepreneurs*, and if it is not known and cannot be reasonably estimated what proportion of the whole group are *entrepreneurs*, then the results for the group as a whole may be biased by a figure which it is not possible to measure.

If, moreover, quotas are thus established and are allowed to relieve interviewers of the necessity of getting representative persons (by some form of random selection) within each quota group, further bias, of unmeasured amount, may result.

On the other hand, such 'quota sampling' (really a form of judgement sampling) is usually better than no stratification at all, for it helps to ensure representation of different important groups, and permits separate study to be made of each such group.

Panels versus successive samples. In studying changes over a period of time, one may select successive samples of comparable composition, or in certain circumstances one may re-interview the same sample. In studying consumer behaviour, for example, as was done by the Central Bank of Ceylon, it may be desirable to select a sample of households and then interview them several times. In the United States, 'consumer panels' of this kind are often used, and are asked to keep diaries of purchases, reading, radio listening, etc., and to take stock (or allow an inter-

viewer to take stock) of the kinds and amounts of goods on hand at periodic intervals.

In some areas, the most feasible way of obtaining desired data on income, expenditures, employment, illness, etc., is by periodic re-interviews of the same sample. Some families themselves can and will keep the daily records required. In many situations, however, it is best to employ a local person to make the entries in consultation with the family.

When a panel is used, relatively small changes over time are often statistically significant, whereas changes of the same size might be ascribed to chance variation from one sample to the next when successive but different samples are interviewed. Thus, for trend studies, results based on relatively few interviews repeated several times with a constant panel can be as reliable as results based on interviews with many more people who are contacted only once. Moreover, more information can often be secured by a series of interviews with the same group, because the results of successive interviews can be cumulated, and because panel members become more articulate over time (partly because they get to know the interviewer, and partly because, between interviews, they think about the subjects studied). It may even be possible to get panel members to co-operate in a development project, for example, by keeping literature available in the house or by turning on certain radio programmes. The exposed families thus become the most likely ones to show changes.

The major disadvantages of repeat interviews are: (a) the considerable 'drop-out' of respondents from one round of interviews to the next (this is called 'panel mortality'), resulting from people moving, becoming ill, dying, or becoming unco-operative; (b) panel members may develop a 'critical set', and thus cease to be representative of the general public; (c) panel members may try to be consistent in actions or opinions, from one interview to the next, thus submerging the very changes the development project is intended to produce and the study to uncover; (d) village workers may concentrate their efforts on the sample villages or panel families they know are being surveyed, thus making them progressively more unrepresentative of the whole area.

Time slice samples. Where either the behaviour to be observed or the process of collecting data is continuous, a different kind of sample is often taken. This is the 'time slice sample', made up of periods of time of specified length, during which data are gathered.

For example, in time and motion studies in factories, an observer (perhaps with a motion picture camera) may select five-minute periods at random throughout the day, each period providing adequate data for a single worker. Care must, of course, be used to ensure that the choice of periods does not introduce a bias. This may be done by random selection (for example, drawing all possible periods out of a hat), or by systematic selection (for example, selecting one period every half-hour throughout the day).

The Dun and Bradstreet surveys of expectations of American businessmen, for example, use a systematic time slice sample of business concerns having a net worth of $75,000 or more. Dun and Bradstreet analytical reporters call on numbers of business concerns every business day (the selection reflecting a periodic need

for new data, not personal choice) to obtain financial information for credit rating purposes. During two weeks at the beginning of each quarter, each analytical reporter asks certain additional questions about the expectations (of sales, profits, inventories, etc.) of all businessmen called on. The sample thus comprises not every fiftieth businessman, but all businessmen called on during a particular period, or 'time slice'.

Combining and analysing data

The collection of data is only one step, though an important one, toward knowledge of project results. The amount of data collected may be enormous and extremely varied. If pointed, action-directed conclusions are to be drawn from them, substantial efforts are required to combine them, analyse their significance, and interpret them in terms germane to the action decisions under consideration.

Editing

When interview schedules or questionnaires are used to collect data, they must usually be edited before the responses obtained can be tabulated. This editing is intended to detect errors and omissions and get them corrected before tabulation is undertaken. The work of editing is to ensure the accuracy and reliability of the data obtained, their consistency internally (with each other) and externally (with data obtained in other ways), the uniformity with which the data were recorded, their completeness, their fitness for tabulation, and their arrangement so as to facilitate coding and tabulation.

In the editing process, it often happens that some of the data sought are missing, or that inconsistencies appear. If data are missing, it is often possible to go back to the source for the needed information, perhaps by a personal interview, a telephone call, a letter, regardless of how the data were originally collected.

It is also often easy to spot where 'years' was written though 'months' was meant, where one reply is inconsistent with another from the same respondent, where an error has been made in arithmetic, where 'jokes' or faked replies occur or where inaccurate information on age, income and occupation is given. Although it is easiest to catch such errors while the interview is going on, they can often be detected afterwards and the necessary changes made.

Classification and coding

Almost all answers to questions must be grouped into categories before they can be combined.

This process of classification may consist simply in listing as 'favourable', 'neutral' or 'unfavourable' all the comments made about a project, about an expert's work, etc. It may go far beyond this, however. For example, one might wish to classify much more elaborately the reasons given for favouring or opposing the introduction of a new examination system in the civil service.

The same reply may, of course, be classified several times, in relation to each

classificatory principle considered significant. For example, here are three ways of classifying possible replies to the open question, 'What is your opinion about using tests in selecting civil servants?'

Classificatory principles

Attitude towards tests:	Attitude towards civil servants:	Attitude towards inter-viewer:
Favourable	Favourable	Co-operative
Neutral	Neutral	Neutral
Unfavourable	Unfavourable	Suspicious
	None expressed	Antagonistic
		Other—please specify

Sets of categories like the three just illustrated should meet three criteria: (a) the categories within any one set should be exhaustive of all possible responses to the question, so that all responses can be classified; (b) the categories within any one set should be mutually exclusive, so that a given response cannot usually be classified in more than one category in that set; (c) each set of categories should be based on a single classificatory principle.

Classificatory principles (and the categories based on them) will in large part have been established before collecting the data, in terms of the kinds of effects expected to result from the development project being studied. Defining these categories precisely, however, deciding how to classify borderline replies, and providing for wholly unexpected replies can usually not be done except on the basis of data obtained in pre-tests, or the first data obtained in the survey. In any case, these processes must be completed before 'coding' can be carried out.

In analysing the data from any considerable number of interviews or questionnaires, it is usually desirable to classify the respondents as well as the responses. On questionnaires, respondents may be asked to classify themselves according to certain of the principles outlined below. In interviews, the interviewer usually does this but may ask the respondents to do so.

Principles often used in classifying respondents, and common categories under each principle, are:

Place of residence: Rural (farm and non-farm); urban (city of a million or more, 100,000 to a million, 50,000 to 100,000; 10,000 to 50,000; 2,500 to 10,000).

Skin colour: White, red, yellow, brown, black.

Nativity: Native born (of native-born parents, of foreign-born parents), foreign born.

Age: Under 5, 5 to 9, 10 to 14, etc., age unknown.

Highest grade of school completed: None, grade school (one to four years; five to six years; seven to eight years), high school (one to three years, four years), college (one to three years, four years or more).

Marital status: Single, married, widowed, divorced, married—spouse absent.

Families: Members and relationship of each member to the head of the household.

Housing and rental: Type of structure of dwelling unit, monthly rent or rental value.

Employment status: Employed, unemployed, not in labour force.

Class of worker: Non-agricultural (private wage or salary worker, government worker, self-employed, unpaid family worker), agricultural (wage or salary worker, self-employed, unpaid family worker).

Occupational groups: Professional and semi-professional, farmers and farm managers, proprietors, managers and officials, clerical and kindred workers, salesmen and saleswomen, operatives and kindred workers, craftsmen, foremen and kindred workers, domestic service workers, service workers (except domestic), farm labourers and foremen, labourers (except farm and mine).

Earnings: (Per week, month or year, before or after income taxes).

Family income: (Same variation).

Industry.

Coding is the technical procedure by which replies are transformed into numbers or symbols which can be tabulated. For example, one may assign the letter F to all favourable replies, U to unfavourable, and N to neutral. Marking such letters or symbols in coloured pencil in the margin of each interview schedule or questionnaire, in a uniform position, greatly simplifies and expedites tabulation.

Coding replies to closed questions is more or less automatic, although judgement is required even there, especially if one of the permissible replies is 'Other—please specify'. Where open questions are used, coding the replies requires great judgement and considerable training. For such questions, clear definitions of categories of replies are absolutely essential before coding takes place. Sometimes, these categories cannot be established until a good many schedules have been examined, so that recoding is necessary of all schedules examined before the final decisions were made as to the categories to be used.

Tabulating

Data from any considerable number of observations quickly lose form and meaning unless they are combined into categories (see previous section), counted by category, and arranged in tabular form.

Tabulation is usually done in one of two ways (it is assumed that tabulation will be done by hand). The categories of response to each question may be listed on a page and then the responses may be tallied along side, using some form of tallying which simplifies counting, such as ɰ or better, ⟨⟩. In each case the finished symbol represents five. The second method of hand tabulation uses small cards, usually 3 × 5 inches in size. One card is made for each interview or questionnaire and the responses are recorded, usually as numbers or symbols, in predetermined positions on each card. Each position is assigned to a single classificatory principle (e.g., age; attitude towards taxes), and each number or symbol stands for one category established for that principle. To tabulate the replies to a particular question, attention is directed to the position used for replies to that question, and the cards are simply sorted into piles, each pile containing cards marked with the same number or symbol in the same position. The cards in each pile are then counted. This can be done very quickly, even for quite large numbers.

The central feature of tabulation is the summarizing of results in the form of statistical tables. If the general nature of the desired tables can be envisaged before

data are collected and processed, time and money will usually be saved. For example, it will then be possible to eliminate questions for which replies cannot be categorized, to avoid recording irrelevant data during the interviews, and to shorten the work of coding and tabulating. Tabulation of pre-test data can be very helpful in working out the kinds of tables needed for the final evaluation report.

Tables are of two basic forms. First, tables may present simple counts of the numbers and percentages of respondents falling into each category of a set, or giving replies falling into each category of a set. Two examples are shown in Tables 8 and 9 below.

TABLE 8. Nativity of respondents (fictitious data)

Nativity	Number		Percentage	
Native born	170		85	
of native-born parents		150		75
of foreign-born parents		20		10
Foreign born	30		15	
Total	200		100	

TABLE 9. Attitudes towards introduction of mental tests (fictitious data)

Attitude	Number	Percentage
Favour	110	55
Neutral	44	22
Opposed	46	23
Total	200	100

Tables of a second basic form are used to present cross tabulations of data, thus bringing out possible relations among the factors investigated. This is illustrated in Table 10.

In this table, the (fictitious) data indicate a strong positive relationship between the two factors measured. A much higher percentage of the native-born respondents than of the foreign-born respondents were favourable to the use of mental tests.

Tables serve two major purposes, namely: (a) analysis; (b) presentation. Tables to be used in analysis, or work tables, present all the data, in their original categories, and in as great detail as can be conveniently obtained. Original unrounded figures are used, and percentages are calculated to an extra place. Many of these tables may reveal nothing new or significant, and may be reproduced only in an appendix, if at all.

Presentation tables are selected for their significance. In order to point out the significant facts they reveal, the categories used are often combined, figures are rounded, averages computed, and other clarifying processes carried out.

TABLE 10. Relationship of nativity of respondents to attitudes towards mental tests (fictitious data)

Nativity	Favour tests		Neutral		Opposed tests		Total	
	No.	%	No.	%	No.	%	No.	%
Native born								
of native-born parents .	100	66[1]	25	17	25	17	150	100
of foreign-born parents.	5	25	10	50	5	25	20	100
Foreign born	5	17	9	30	16	53	30	100
Total	110	55	44	22	46	23	200	100

1. This illustrates the principle of 'forcing up or down' percentages so that the total is 100 per cent. The number 100 is 66.67 per cent of 150 and 25 is 16.67 per cent. The forcing in this case is done on the largest figure, where it makes the least relative change. Usually, forcing is done by increasing or decreasing a figure where it does least violence to the impression given by the unforced set of figures.

Summarizing data

Regardless of the kind of data collected and tabulated, it is necessary to express simply and clearly what is contained in the tables. The statistical measures most commonly used for this purpose are described here briefly.

Counts or frequencies. A simple statement of the number of people or responses in a significant category is the simplest kind of measure. This may be the actual number or frequency obtained by counting the data, or it may be the estimated number or frequency.

If, for example, every tenth child in a school is asked what game he likes best, and if six children say they like badminton best, it can reasonably be estimated that sixty children in the school prefer badminton. Any estimate derived by an expansion of this sort has a 'margin of error'. The calculation of this margin is explained in standard statistical works (see bibliography).

Percentages. One of the most useful ways of summarizing data and establishing the relative frequency of different kinds of response is by calculating percentages. If one divides the number in category A of a particular universe by the total number comprising that universe, and if one then multiplies by 100, one will obtain the percentage in category A. For example, if seventy-four persons are asked their attitude toward using mental tests in selecting civil servants, and if sixty-five approve, the percentage approving is stated as 88 per cent.

While this calculation produces the figure 87.84 per cent, the last two figures are meaningless when the total is only seventy-four, and are usually dropped by rounding to the nearest whole per cent. Moreover, since one of the main purposes of converting to percentages is to provide a figure (proportionate to 100) which can be easily grasped, it is simply muddying the waters to report human behaviour, which is in any case variable, in tenths of a per cent.

Percentages are extensively used in making comparisons between groups, of course, because they provide a common measuring rod for groups of different size (see 'Differences and relationships based on data from samples', page 108).

Averages. If a large percentage of a group acts in a particular way, we may call this act 'typical' of the group. Other measures of what is 'typical' are the several kinds of averages.

The mean average (or simply the 'mean'). This is the most familiar kind of average. It is computed by adding up the scores (or ages, or incomes, etc.) of respondents and dividing by the number of respondents.
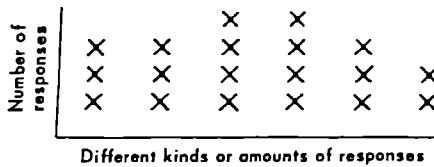
The modified mean. Because a few extremely high or extremely low scores may distort the mean average, a small number of the highest and the same number of the lowest scores (up to one-sixth of each) may be excluded from the calculation. This results in a modified mean, which may be quite useful, provided the method of calculation is explicitly reported. For example, in comparing two surveys at different times, the chance inclusion of a millionaire in one survey and none in the other would make it appear, when true means were calculated, that a substantial change in the average income of the group had taken place. Using a modified mean would avoid such an erroneous conclusion.

The median. The extreme case of the modified mean is the median. Here just less than the highest 50 per cent and just less than the lowest 50 per cent are excluded from the calculations. The median is determined by identifying the middle score (or calculating the average of the two middle scores), and is also used to describe the 'typical' score of the group. An extreme case above (or below) the median has no more influence on the median than any other case above (or below) it.
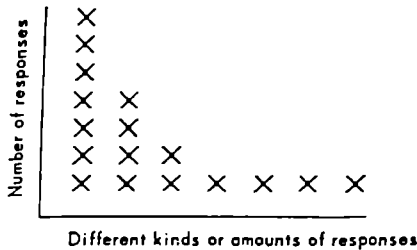
Variability. It is often desirable to describe briefly the variability of responses obtained, as a measure of the homogeneity or heterogeneity of the group studied. This may be done by stating, for example, that 80 per cent of the group gave one out of three possible responses to a question (great homogeneity) or that no more than 20 per cent gave any one of seven possible responses (great heterogeneity).

Where data are in the form of scores or amounts, we may use the range (from lowest to highest) to describe how different the individuals in the group are with respect to the characteristics so expressed. Here again, however, the occasional occurrence of extremely low or extremely high scores may make it desirable to exclude extremes. If the highest one-sixth and the lowest one-sixth are excluded, we have a modified range (comparable to the modified mean), which can be described clearly by saying, 'The middle two-thirds of the responses fell between X and Y'. This measure, if divided in half, is quite similar to the more precise measure called the standard deviation (one standard deviation above and one below the mean will in a normal distribution encompass 68 per cent of the responses).

Distribution over the range. If responses are charted as vertical bars, some distributions of responses look like the following diagram. This is something like a 'rectangular' distribution. Neither very many nor very few responses fall at any one point on the horizontal scale.
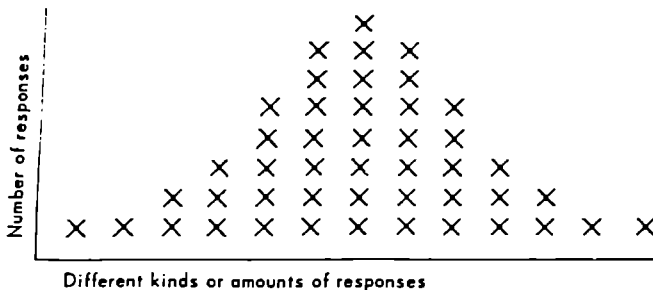
```
         X   X
X   X    X   X   X
X   X    X   X   X   X
X   X    X   X   X   X
```

Different kinds or amounts of responses

Other distributions look like this:

Number of responses

```
X
X
X
X  X
X  X
X  X  X
X  X  X  X  X  X  X
```

Different kinds or amounts of responses

This is a 'J curve' distribution. Many responses fall at one end, and progressively less fall at each point as one moves away from that end.

Other distributions look like this:

Number of responses

```
                  X
              X   X   X
              X   X   X
          X   X   X   X   X
          X   X   X   X   X
      X   X   X   X   X   X   X
    X   X   X   X   X   X   X   X
X   X   X   X   X   X   X   X   X   X   X   X
```

Different kinds or amounts of responses

This approximates a bell-shape curve, or 'normal curve'.

It is important to know the shape of the distribution of responses, in order to decide what statistical measures to use in summarizing and analysing the responses. Many statistical short cuts and special measures have been developed for use with 'normal curves' and are inapplicable to other types of distribution curves. If, for example, field data are to be analysed more elaborately at headquarters, it will help greatly to find out the shape of the distribution of responses so that applicable tests of significance or relationship can be employed.

Differences and relationships based on data from samples

In evaluating projects, the primary purpose of statistical analysis of data indicating

project results is to determine whether or not significant changes take place which can be attributed to the project or to its separate operations. As pointed out in Step Four, page 81, there are two logical ways of doing this: (a) by uncovering significant differences between the project situation before and the same situation after (so long as like differences do not show up in otherwise comparable situations which were not subjected to similar operations); (b) by uncovering significant differences between current situations which differ in the incidence of project operations. The latter method is inferior, as it assumes without having baseline data that the situations compared were not significantly different before the operations began. It is often all that can be undertaken, however, and it is much better than nothing.

Fairly complex statistical procedures are required to uncover significant differences and to calculate their extent and reliability, where the data used have been obtained from samples. Some rough rules of thumb can be suggested, however.

Suppose a campaign to promote the use of commercial fertilizer has been carried on for a year, and that surveys of fertilizer use have been carried out with four samples of farmers, as follows: sample A in area X (where campaign took place), before campaign; sample B in area X (where campaign took place), after campaign; sample C in area Y (where no campaign took place), before campaign; sample D in area Y (where no campaign took place), after campaign.

Suppose the four samples show the following numbers and percentages using commercial fertilizer at the time of the survey:

TABLE 11. Farmers using commercial fertilizer (fictitious data)

Area	Before campaign			After campaign		
	Sample and size	Number using	Percentage using	Sample and size	Number using	Percentage using
X (had campaign)	A (76)	20	26	B (85)	40	47
Y (no campaign)	C (90)	24	27	D (88)	30	34

It will be noted that the use of commercial fertilizer increased in both areas during the campaign. Perhaps some general factors were operating during this period to increase the use of fertilizer. A glance at Table 12 will quickly indicate, however, that the apparent increase in area Y from 27 per cent to 34 per cent was of a size which could have occurred as a result of chance variations from one sample to another. For sample D to have been considered significantly different from sample C, the difference (with the sample size around 100 and the percentage obtained around 30) would have had to exceed 9 per cent, whereas it was only 7 per cent, well within the variation which might have arisen by chance.

In fact, to be considered really significant, the difference in this case should be compared with the chance variability of not one percentage above but of the difference between the two percentages. Where the percentages compared have

about the same chance variability (as indicated in the third column of Table 12), the difference between them will have a chance variability about 1.4 times as great. In the above example, then, the difference between sample C and sample D should have been 13 per cent or more to be taken as evidence that any change had really occurred in area Y.

In area X, however, the difference between sample A (26 per cent) and sample B (47 per cent) was 21 per cent, a change extremely unlikely to have resulted from chance variations.

By using Table 12 (interpolating where necessary), and multiplying by 1.4 the larger of the chance variabilities of the two percentages compared, one can readily decide whether the apparent 'change' could in fact be explained simply as a chance variation that could have resulted from sampling. If the difference is too large to have been likely to arise by chance, it is reasonable to attribute its cause to the development project or to individual operations which are associated with the difference.

There are numerous other statistical methods of determining and measuring the extent of relationship. These are complicated, however, and for most field studies add little to the percentage difference method described above.

All relationships indicated must, however, be examined with the greatest care, to determine whether or not an apparent relationship between factors A and B may not in fact have appeared because A and B are both related to C, without

TABLE 12. 'Chance variation' above and below obtained percentage

Percentage obtained	Size of the sample	The 'chance variation' or 'margin of error' in 19 out of 20 similar samples (per cent)
50	20	±22
	50	±14
	100	±10
	200	± 7[1]
	1 000	± 3[1]
70 or more or 30 or less	20	±20
	50	±13
	100	± 9
	200	± 6[1]
	1 000	± 3[1]
90 or more or 10 or less	20	±14
	50	± 8
	100	± 6
	200	± 4[1]
	1 000	± 2[1]

1. This sample size, and therefore this precision, is not usually feasible or necessary to guide the administration of development projects.

any relationship actually existing between A and B. 'Spurious' relationships of this sort are common. Men living in large houses more frequently own automobiles than men living in small apartments. Does home ownership bring about automobile ownership? Not necessarily. Income is highly associated with both, and so is location of dwelling unit. Logically and then statistically untangling real from spurious correlation is of first importance for all studies of project results.

Finally, it is rarely worth while to apply elaborate statistical methods to differences or changes which appear to be small. If a project has not had effects which at least appear to have some significance, why spend time analysing data? On the other hand, differences often appear to be significant when in fact the proper statistical tests cast doubt on their actual significance. The tests suggested above are more likely to demolish claims of project success than to demonstrate success where none was claimed.

Bibliography

Methods of collecting and analysing data

ABRAMS, Mark A. *Social surveys and social action*. London, Heinemann, 1951.

AMERICAN PUBLIC HEALTH ASSOCIATION. Health program evaluation, *Public health reports*, vol. 70, no. 2, February 1955, p. 136-41. Report of the eighty-second Annual Meeting of APHA, 11-15 October, 1954.

ARNOLD, Mary F. Principles of evaluation, in: *The voluntary health agency—Getting community action*, p. 45-9. (Continuing education monographs no. 4) San Francisco, Calif., Western Regional Offices, APHA, 1963.

BATTEN, T.R. Some criteria for the evaluation of the training of extension workers in fundamental education programmes, *International journal of adult education*, vol. 17, no. 3, September 1956, p. 12-19.

BOGARDUS, Emory S. *Introduction to social research*. Los Angeles, Calif., New York, Suttonhouse, 1936.

BURGESS, A. Nutrition education in public health programs—what have we learned? *American journal of public health*, no. 51; 1961, p. 1715.

BURROUGH, G. E. R. Evaluation in fundamental education, *International journal of adult education*, vol. 17, no. 4, December 1956, p. 6-13.

BYRN, Darcie (ed.) *Evaluation in extension*. Washington, D.C., United States Department of Agriculture, Division of Extension Research and Training, Federal Extension Service, 1959.

DIXON, Wilfrid J.; MASSEY, Frank J. Jr. *Introduction to statistical analysis*. Second edition. New York, McGraw-Hill, 1957.

Evaluation techniques. *International social science bulletin*, vol. VII, no. 3, 1955.

FAO. *Report of the European seminar on evaluation of home economics extension programs, Vienna, Austria, 9-21 May 1960*. Rome, 1960. (Nutrition meetings report series no. 27.)

——. *Some guidelines for evaluating home economics extension programs*. (European Seminar on Evaluation of Home Economics Extension Programs, Vienna, 9-21 May 1960. Mimeographed. (FAO/NU/HE/60/5.)

FESTINGER, Leon; KATZ, Daniel (eds.). *Research methods in the behavioral sciences.* New York, Dryden Press, 1953.

FLECK, Andrew C., Jr., M.D. Evaluation as a logical process, *Canadian journal of public health,* vol. 52, no. 5, May 1961, p. 185-91.

FRENCH, David G. *An approach to measuring results in social work.* New York, Columbia University Press, 1952.

HANSEN, Morris H.; HURWITZ, William N.; MADOW, William G. *Sample survey methods and theory,* volume 1, *Methods and applications.* New York, J. Wiley & Sons; London, Chapman & Hall, 1953.

HYMAN, Herbert H. *Interviewing in social research.* Chicago, Ill., University of Chicago Press, 1954.

——. *Survey design and analysis.* Glencoe, Ill., Free Press, 1955.

——; WRIGHT, Charles R.; HOPKINS, T.K. *Applications of methods of evaluation: four studies of the incampment for citizenship.* Berkeley and Los Angeles, Calif., University of California Press, 1962. (University of California publications in culture and society, vol. 7.)

JAHODA, Marie; DEUTSCH, Morton; COOK, Stuart W. *Research methods in social relations.* New York, Dryden Press, 1951.

KAHN, Robert L.; CANNELL, Charles F. *The dynamics of interviewing.* New York, J. Wiley & Sons; London, Chapman & Hall, 1957.

KNUTSON, Andie L. Evaluating program progress, *Public health reports,* vol. 70, no. 3, March 1955, p. 305-10.

——. Human behavior factors in program planning, *Public health reports,* vol. 70, no. 11, November 1955, p. 1129-34.

KRUYER G.J. Evaluation-ondersoek, *Landbouwkundig Tydschrift,* September 1964, p. 615-28.

LEBRET, Louis J. *Guide pratique de l'enquête sociale.* Paris, Presses universitaires de France, 1951.

LERNER, Daniel; LASSWELL, Harold D. (eds.). *The policy sciences.* Stanford, Calif., Stanford University Press, 1951.

LIKERT, Rensis. 'Behavioural research: A guide for effective action', in: Rensis Likert and Samuel P. Hayes, Jr. (eds.), *Some applications of behavioural research.* Paris, Unesco, 1957.

LINDZEY, Gardner (ed.). *Handbook of social psychology,* volume 1, part 3, 'Research methods'. Cambridge, Mass., Addison-Wesley, 1954.

LUNDBERG, George A. *Social research, a study in methods of gathering data.* New York, London, etc., Longmans, Green, 1942.

MAGET, Marcel. *Guide d'étude directe des comportements culturels.* Paris, Civilisations du sud, 1953.

MCCARTHY, Philip J. *Introduction to statistical reasoning.* New York, McGraw-Hill, 1957.

MCKENZIE, J.C.; MUMFORD, Pamela. The evaluation of nutrition education programmes: a review of the present situation, *World review of nutrition and dietetics,* vol. 5, p. 21-31.

MERTON, R.K.; LAZARSFELD, P.F. *Continuities in social research: Studies in the scope and method of 'the American soldier',* Glencoe, Ill., Free Press, 1950.

MORONEY, M. J. *Facts from figures*. Baltimore, Md., Penguin Books, 1953.

OPLER, Morris E. *Social aspects of technical assistance in operation*. Paris, Unesco, 1954.

PARTEN, Mildred B. *Surveys, polls and samples*. New York, Harper, 1950.

ROBERTS, Beryl J. Concepts and methods of evaluation in health education, *International journal of health education*, vol. V, no. 2, April-June 1962, p. 52-62.

ROGERS, E. M. *Bibliography of research on the diffusion of innovations, 1964*. Available from Department of Communication, Michigan State University.

UNITED NATIONS ECONOMIC AND SOCIAL COUNCIL/UNICEF EXECUTIVE BOARD. *Possible ways to strengthen evaluation*. Note by the Executive Director. 4 November 1963. Mimeographed. (E/ICEF/486.)

UNITED STATES NATIONAL ADVISORY MENTAL HEALTH COUNCIL. Community Services Committee, Subcommittee on Evaluation of Mental Health Activities. *Evaluation in mental health: a review of the problem of evaluating mental health activities*. Washington, D.C., United States Department of Health, Education and Welfare, 1955. (Public health services publication no. 413.)

UNIVERSITY OF MICHIGAN. *The second national conference on evaluation in public health. September 12-14, 1960*. Published by the University of Michigan, School of Public Health. Mimeographed.

WARREN, Roland L. *Studying your community*. New York, Russell Sage Foundation, 1955.

WOODBURY, Robert M. *Methods of family living studies*. London, published for the International Labour Office by P. S. King & Son, Ltd., 1940.

YATES, F. *Sampling methods for censuses and surveys*. New York, Hafner Publishing Co., 1949.

YOUNG, Pauline. *Scientific social surveys and research*. Third edition. Englewood Cliffs, N. J., Prentice-Hall, 1956.

ZEISEL, H. *Say it with figures*. New York, Harper, 1950.

Examples of the collection of social data by scientific methods, outside Europe and North America

BEERS, H. W. Evaluation in Community development—The Indian experience, *International review of community development*, 1960.

BOSE, S. P. *Evaluation of rural extension*. International Agricultural Centre, Wageningen, The Netherlands. Twelfth International Training Centre on Methods and Programme Planning in Agricultural and Home Economics Extension, 6 July to 1 August 1964. Mimeographed.

DROBNY, Abraham. Evaluación de Programas de Salud, *Boletín de la Oficina Sanitaria Panamericana*, vol. LVII, no. 2, Agosto 1964, p. 112-15.

FLORES, Gerardo. 'A study on functional literacy for citizenship in the Philippines', *Fundamental education*, 1950, vol. II, no. 3.

GAMBLE, Sidney D. *How Chinese families live in Peiping*. New York, London, Funk & Wagnalls Co., 1933.

HAYDEN, Howard. *Moturiki; a pilot project in community development.* London, New York, Oxford University Press, under the auspices of the South Pacific Commission, 1954.

HUSAIN, A. F. A. *Human and social impact of technological change in Pakistan.* Two voltmes, Dacca, Oxford University Press, 1956.

JACOBY, E. H. Evaluation of the effects of agrarian reform programmes, in: N. V. Sovani and V. M. Dondekar. (eds.), *Changing India.*

LEACH, Edmund R. *Social science research in Sarawak.* London, HMSO, for Colonial Office, 1950.

LOOMIS, Charles P. 'Studies of Latin America', part III of *Studies in applied and theoretical social science at Michigan State College.* East Lansing, Mich., Michigan State College Press, 1950.

——; MORALES, Julio O.; CLIFFORD, Roy A.; LEONARD, Olen E. *Turrialba: Social systems and the introduction of change.* Glencoe, Ill., Free Press, 1953.

MASON BEERS, N. *Extension evaluation in a foreign culture: some important factors to guide the technical worker in rural development.* Ithaca, New York, Cornell University, 1961. (Comparative extension publications, mimeo release no. 5, February 1961.)

MCINTYRE, Alan J.; MCINTYRE, J. J. *Country towns of Victoria, a social survey.* Melbourne, Melbourne University Press, 1944.

MUÑOZ, Raúl; SERRA, Belen M.; ROCA, Angelina S. de. 'Research and evaluation in a program of community education', *The journal of social issues*, vol. 9, 1953, p. 43-52.

NISHIKIORI, Hideo. *Togo-mura, a village in northern Japan.* New York, International Secretariat, Institute of Pacific Relations, 1945. (Original Japanese edition published by the Tokyo Imperial University, College of Agriculture, Section on Agricultural Policy, 1936.)

PUERTO RICO. Division of Community Education. *The use of social research in a community education programme.* Paris, Unesco, Education Clearing House, 1954.

SPICER, Edward H. (ed.). *Human problems in technological change.* New York, Russell Sage Foundation, 1952.

TEAF, Howard M., Jr.; FRANCK, Peter G. (eds.). *Hands across frontiers.* The Hague, Netherlands Universities Foundation for International Co-operation, 1955, and distributed in U.S.A. by Cornell University Press, Ithaca, N.Y.

VERSLUYS, J. D. N. *Evaluation of community development projects*, in: *Indian year-book of international affairs, 1955*, Madras.

WEIR, Wasif, Hassan Salah el Din Moh Kader. 'An evaluation of health and sanitation in Egyptian villages', *Journal of the Egyptian Public Health Association*, 1952.

Community projects—first reactions. New Delhi, Government of India Planning Commission, Programme Evaluation Organization, August 1954.

Evaluation report on first year's working of community projects. New Delhi, Government of India Planning Commission, Programme Evaluation Organization, August 1954.

Report of a sample survey of social stratification and mobility in the six large cities of Japan. Tokyo, Japan Sociological Society, Research Committee, 1952.
Report on the survey of rural credit and rural unemployment in East Pakistan, 1956. Dacca, Dacca University Socio-economic Survey Board, 1958.
Survey of Ceylon's consumer finances. Colombo, Central Bank of Ceylon, 1954.

35740
9.2-71