# Approaches to communication planning

Edited by John Middleton



# Monographs on communication planning

1

ļ

1

I

# Approaches to communication planning

Edited by JOHN MIDDLETON



Published in 1980 by the United Nations Educational, Scientific and Cultural Organization 7 place de Fontenoy, 75700 Paris Printed by Imprimerie des Presses Universitaires de France, Vendôme

ISBN 92-3-101801-9

© Unesco 1980 Printed in France

# Preface

Planning for communication institutions and systems has long been part of Unesco's programme, but it is only during the last five years that it has been approached systematically, as a coherent field of operations and research. Over this period a range of activities—including training, methodology studies and project development—has focused upon the needs and problems of integrated planning for communication systems, especially at the national level.

One of the strands of the programme has been the development of materials, and this anthology is the first of several publications to emerge. As such it is deliberately introductory; in its chapters, a number of specialists from a wide range of disciplines discuss the applicability of their field to the communication sector, and the problems posed by adaptation.

A companion volume, to appear shortly, will propose an operational framework for communication planning for development, derived from an analysis of earlier planning experiences and evaluated both in the laboratory and in the field.

These books should, however, be related to the overall programme of which they are a part. For example, attempts are being made to develop instructional materials for communication planners and to create training opportunities. A group instructional package, originally developed by the University of Syracuse in the United States, was tested in a regional Asian workshop, held at the Universiti Sains Malaysia (Penang) in 1977, and refined during 1978 by the East-West Communication Institute. This package is now available in English for training purposes; meanwhile a Spanish adaptation for Latin American countries has been undertaken by the International Centre for Higher Studies in Communication for Latin America (CIESPAL), Quito, and was piloted during a regional workshop in late 1979. Later, the package will be extended to other regions.

Case studies are also being compiled of planning processes at various levels, including explorations of new concepts in participatory and open

planning. Similarly, the area of methodology is being extended: in the East-West Communication Institute, an interdisciplinary task force is working on a comprehensive study of planning methods and their applicability to communication problems. Above all, the practical business of planning continues; while requests to Unesco for assistance cannot be predicted in advance, at least it is possible to ensure that such exercises, when they take place, are adequately recorded and monitored, to add to our growing stockpile of experience.

The present collection of essays is intended for a variety of readers. Its main audience, evidently, is those who are responsible for the preparation of communication plans and for their implementation, at either sectoral or national levels. But as a contribution to a new field, it should also be of interest to communication researchers and practitioners, and to planners and researchers in other fields, especially those whose longer-established disciplines are now making a contribution to the communication sector.

The book has been edited by Dr John Middleton, Assistant Director for Planning and Evaluation at the East-West Communication Institute. Its authors have in all cases been asked to write in their personal capacity, and the opinions expressed are their own; they do not necessarily reflect the views of Unesco.

# Contents

Chapter 1	Approaches to communication planning: an overview John Middleton	9
	Ι	
	THE PROCESS APPROACH	
Chapter 2	Images and action: theories in and of communication planning	
	John Middleton	19
	II	
	THE SYSTEMS APPROACH	
Chapter 3	An overview of systems analysis James Lyons	65
Chapter 4	Modelling and simulation	
	Gus Root	113
	TTT	
	THE TECHNOLOGICAL APPROACH	
Chapter 5	The elements of communication technology	

Chapter 5 The elements of communication technology Bernard Webster 155

Chapter 6	Short- and long-term forecasting for communication technology planning	
	John Spence	183
	IV	
	THE ECONOMIC APPROACH	
Chapter 7	Economic analysis and communication decision-making Steven Klees	
	Stuart Wells	211
Chapter 8	Financing and control of capital-intensive broadcasting systems	
	Stuart Wells Station Klass	247
	Steven Rices	247
	v	
	THE EVALUATION APPROACH	
Chapter 9	Evaluation strategies for communication planning John Mayo	267

Notes on contributors 299

# Chapter 1

# Approaches to communication planning: an overview

### John Middleton

The importance of the flow of information and knowledge in society has long been accepted. Communication, the process through which information is shared, is often called the web of society, holding together the individuals, groups and institutions of which a society is comprised. As a broad social process, communication incorporates several essential functions: informing and being informed; persuading and being persuaded; teaching and learning; entertaining and being entertained.

Inevitably communication became of central concern when men and women turned their attention to planned action to change and improve the nature of society. Much of this attention, since the Second World War, has been devoted to the task of national social and economic development, particularly in what came to be known as 'less developed countries'. The purposive use of communication as part of development strategies has been manifest in various ways—through the development of extension services in agriculture and health, through the expansion and qualitative improvement of school systems, through the development and use of mass media, and through the expansion and improvement of telecommunication systems and services.

In industrialized countries, attention to communication systems has arisen from a somewhat different source. There the motivation for examination of the role and function of communication systems has stemmed primarily from rapid advances in communication technology. The potential of new technologies, particularly in the area of telecommunications, has promised significant advances in both the effectiveness and efficiency of communication systems, with the potential for important effects on society.

By the late 1960s, these two trends began to be drawn together. The continuing rapid development of communication technologies and their export to developing countries began to be taken into account in national development planning. At the same time, efforts for planned use of communication within development strategies came to be seen as fragmented, uncoordinated and occasionally in conflict. New technologies held the promise of significant improvement in communication systems, but at high cost and with considerable uncertainty about their effects.

At the same time, ideas about what 'development' ought to be, and how it ought to be achieved, were changing. The capital-intensive industrialization model of development that had dominated the first twenty years of development planning came into serious question. New goals emphasizing social equity over increased Gross National Product (GNP) began to take hold. Wider participation in development came to be seen as an alternative to centrally planned, top-down models.

It has been from these three trends—communication for development, technological advances, and changing images of development—that the new field of communication planning has emerged.

In a very real sense, this new field also emerges from a number of traditions, of levels and of action, with respect to communication and development. The traditions include the planning of communication in support of development projects in agriculture, health and community development. A major feature of this tradition has been the communication campaign, using mass and interpersonal communication. A second tradition has been the planned use of communication technology in support of education. Notable in this area have been educational radio and television experiments in a number of countries. A third tradition has been that of the role of mass media in national development, with the press and broadcast media taking responsibility for creating awareness of development programmes, providing information to stimulate the adoption of new practices (such as family planning), and providing information designed to serve a variety of purposes leading towards the creation of national unity and integration.

Four levels of action can be identified. At the lowest level we have the planning of communication projects, such as a family planning campaign. Above this we have the level of a sectoral programme, usually managed by a ministry or agency, which is comprised of a number of projects and activities, sometimes co-ordinated, sometimes not. At a third, national, level, communication planning is done in co-ordinated fashion across ministries and agencies, and occasionally across sectors. Sometimes this is done through a co-ordinating body, such as the Philippine Commission on Population. Also at the national level, certain aspects of communications—notably telecommunications and investments in new technologies—are part of national planning for development. The fourth level is international, where a policy debate over a number of critical aspects of world communication has emerged in recent years. Given the multiple trends, traditions and levels out of which communication planning is emerging, it is not surprising to find that the field is characterized by a variety of approaches. These grow in part from the fragmented record of experience that has characterized the recent past; and in part from different theoretical perspectives on planning, on development and on communication in society—what it is, how it can be planned.

The chapters in this volume reflect this variety. Five different approaches are presented: process, systems, technological, economic, and evaluation. There are elements in common across the approaches, notably a concern for how communication can be effectively planned as part of social and economic development. But the emphases and underlying assumptions differ in many respects, as do the aspects of communication in society that are addressed. However, taken as a whole, the chapters form a significant contribution to the development of the field of communication planning, providing a number of building blocks out of which more integrated approaches may be formed.

# The process approach

John Middleton's chapter on 'Images and Action: Theories in and of Communication Planning' deals directly with the communication planning process. The chapter deals with two central issues. The first relates to the use of theories within the planning process. Asserting that planning is, in essence, the application of theory, Middleton goes on to review four selected theory areas of relevance to planning communication: development theory, sociological theory, communication theory and organization theory. The emphasis is on the implications of these theories for planning—how and why they are used.

The second issue is that of the planning process itself. A number of theories of planning are reviewed. These theories provide alternative ways of organizing the planning function and process, given different purposes and planning contexts. The thrust of the argument is that there are alternatives to the widely known rational/comprehensive planning approach.

The chapter concludes with the identification of four 'design principles'. These are broad concepts that are thought to aid planners, early in the process of designing planning, to select and combine various theories.

This chapter provides a broad framework against which the following chapters may be viewed. It emphasizes, for example, that the broad view that planners hold of society has important implications, both for the nature of planned action and the way in which planning is carried out. The level of certainty in planning—in understanding of problems, in data,

~1

in theories—is seen as a major variable in deciding how to plan. The role of values in planning is emphasized, as is the need for planning to be a continual learning process.

# The systems approach

General systems theory, and its application through systems analysis and systems planning, is an important theoretical and applied tradition. Two chapters deal with this approach to planning. In 'An Overview of Systems Analysis' James Lyons begins with a broad review of the systems approach. He then goes on to describe a six-step model for systems development. Special attention to the systems development phase of 'systems analysis' follows, a process which Lyons sees as both an 'art and a science'. He provides an analysis of a number of techniques useful in systems analysis, including ways of presenting data, decision techniques, flow charting, interviewing, creativity techniques, and cost-benefit analysis.

While the chapter deals generally with systems approaches, there is emphasis on how to establish new systems within organizations. Thus this chapter is of particular value to planners faced with the task of setting up organizational systems to carry out communication functions.

Gus Root deals with two specific systems analysis techniques. In 'Modeling and Simulation' he stresses the analytical approach to understanding the problems which plans will address. Thus, like Lyons, he emphasizes the analysis phase of a full systems development approach. Root states the theme of his chapter as: 'Begin any effort to change a complex social system by constructing models and simulations of the problem and its social context; they will increase the likelihood of longterm success, decrease the chance of unexpected negative side effects, and help conserve scarce national resources.'

Three kinds of model are reviewed: prose, structural, and dynamic. Advantages and disadvantages of each are analysed, and techniques for the development of models of each type described and exemplified: brainstorming, Delphi, scenarios, cross-impact matrices, causal-loop models.

The discussion of simulations as a planning technique emphasizes human simulations as a dynamic way in which decision-makers can obtain personal experience with and insight into dynamic models.

The emphasis in Root's chapter is on systems analysis applied to problems in the environment, providing planners with an analytical perspective on problem analysis and a range of techniques to use in implementing this perspective. The variety of models provided gives planners a set of alternative procedures from which to choose, given resources available to them—in time, in money, in technology, and in data.

# The technological approach

For many communication planners, technology is mysterious in many of its technical components. Bernard Webster, in 'The Elements of Communication Technology', sheds considerable light on the basic technical elements of communication technologies. He provides a useful review of trends in the development of new technologies, including computers. The international policy aspects of communication, and the agencies concerned with these issues, are also reviewed.

Webster's emphasis is on telecommunications and broadcasting, with more attention to the former. He outlines a national communication structure, and provides an example of the kinds of technological alternative that can confront planners, showing how technological alternatives have different social implications.

John Spence deals with a particular aspect of communication technology in 'Short- and Long-term Forecasting for Communication Technology Planning'. Short-term forecasting applies to planning for the growth and expansion of existing communications systems. Long-term forecasting is concerned with alternative futures over long time spans, perhaps twentyfive years.

The use of alternative scenarios for long-term forecasting is of central concern. Given the rapid changes in technology and in society, long-term futures cannot be predicted from the past. Planners need to construct an array of alternative futures, then assess the impact that developing technologies may have on these futures, and vice-versa. The need for continual monitoring of events as they unfold, with constant revision of scenarios, is considered essential.

Techniques for scenario construction and use are provided in some detail, with examples.

Both technology papers deal directly with the issue of projection. Webster is somewhat doubtful about the value of projections for telecommunications services in developing-country situations, and gives reasons. Spence, writing from long-term planning experience in Australia, is more optimistic. Both see telecommunications planning as being done systematically, with evaluation of alternatives as a key element.

Both papers are also written from the perspective of the market economy, in which telecommunication services are 'sold'. This perspective is most clear in Spence's work.

# The economic approach

The economics of communication form an important area of inquiry, and economic perspectives on communication planning are essential. Stuart Wells and Steven Klees explore two aspects of economic approaches.

Ways in which economics concepts can be useful to communication decision-makers are the focus of 'Economic Analysis and Communication Decision-Making'. Two contrasting frameworks of economic thought are reviewed and compared: neoclassical economics and Marxist economics. Both frameworks are applied to communication decision-making in three contexts: communication as an economic good; communications as an industry; and the role of communication in national development.

The authors show how different economic frameworks lead to different kinds of communication decisions, and emphasize that communication planning must be guided by an overall framework.

The analysis is continued and made more specific in 'Financing and Control of Capital-intensive Broadcasting Systems'. Economic efficiency and social equity are contrasting goals in development. The authors analyse the effects on goals of both kinds of three types of ownership of capital-intensive broadcast media: private, non-regulated ownership; publicly regulated ownership; and public ownership. The strengths and weaknesses of each form of control are analysed, and the authors conclude that some form of public regulation may be most workable.

The analyses offered in the two chapters apply most directly to communication planning decisions at national levels, illustrating broad policy choices based on economic criteria and analysis.

# The evaluation approach

John Mayo approaches the functions of evaluation in communication planning systematically, and against an image of communication planning in less developed countries that separates policy decisions from planning and implementation. In 'Evaluation Strategies for Communication Planning' he outlines the functions of evaluation at four stages of planning: policy definition, systems design, systems implementation and maintenance, and assessment of impact. He thus shares an image of systems planning with Lyons and Root.

Evaluation is defined as the 'exercise of insight and informed judgement to assess programme effectiveness'. It is an ongoing process, closely linked with all phases of planning, with the primary purpose of strengthening decision-making.

The increasing realism of evaluators is noted, with a shift from heavily

quantified, social-science research-based approaches to a variety of methods for gathering reliable information. Mayo believes that stronger and more co-operative relationships between planners, administrators and evaluators are essential if evaluation is to become effective. He raises the important issue of the role of evaluation in increasing popular participation in planning.

The chapter is notable for its pragmatic thrust, its recognition of the importance of institutional relationships and values in planning, and its emphasis on flexibility in methodology. Planners and evaluators alike will find it a useful framework for a broad and thorough consideration of evaluation as an integral part of communication planning.

# Missing approaches

The five approaches to communication planning that are dealt with in this volume are representative of the larger number of approaches that deserve consideration. As the field develops, other approaches merit careful attention. A selected few of these are briefly noted below in order to provide a fuller sense of the range of issues that face communication planners—and those who would develop the field academically and professionally.

### THE POLITICAL APPROACH

Communication is intimately associated with political processes, regardless of ideological base. Communication systems are utilized for the flow of information around which political systems function.

What is communication planning from the political perspective? Does a political approach differ significantly from the approaches outlined in this volume? If so, how?

# THE ETHICAL/NORMATIVE APPROACH

The planned use of communication in society touches on deeply held ethical values. Should communication be planned at all? What are the ethical implications of the purposive use of communication to persuade people to adopt new practices, to change attitudes and behaviour? Who decides when such uses of communication are acceptable? How are these questions viewed in differing cultural/ethical contexts?

### THE CLIENT APPROACH

The chapters in this volume, and indeed much of the writing about communication planning that now exists, assume the perspective of a centralized planning unit. True, this unit may engage in a variety of strategies to obtain client input in planning. True also, measures may be taken to decentralize planning.

But beyond this, we need to see communication systems, and their planning, from the perspective of the clients of these systems—the individuals whose needs are to be served. Given different social and cultural conditions, how do clients view communication? What rights to communication do they feel that they possess? How aware are they of the potential of planned communication, and of their relationship to planned communication systems? What kinds of planning could they—or would they—do if given the opportunity?

# THE SECTORAL APPROACH

Are there any unique demands on communication planning within specific sectors of the national structure? Does agriculture have needs sufficiently different from education or health to warrant different ways of planning, or different allocations of communication resources? What are the special claims of telecommunications, of mass media, of educational systems, of extension systems on communication resources? If there are different or competing interests, how can they be reconciled? Should they be reconciled?

# THE INTERNATIONAL APPROACH

Levels of communication planning are increasingly perceived as interrelated. The chapters in this volume deal primarily with project, ministry and national levels of planning. What about the international level? How do international policies affect national systems and vice-versa? What about the complex question of international technology flows? What obligations should industrialized societies assume for communication development in the Third World?

# Towards the future

Despite the many missing approaches, we believe this volume offers a useful contribution to the evolution of communication-planning thinking. Those individuals with direct involvement in communication planning face an unknown terrain, one with many hidden pitfalls—yet, at the same time, one with vistas of the potential that communication has to bring about a better quality of life in a rapidly changing world. The value of this volume must be judged by the extent to which it aids practising communication planners, opening up alternative approaches that strengthen and illuminate their work.

Ι

The process approach

# Chapter 2

# Images and action: theories in and of communication planning

### John Middleton

What is planning? One way to approach the problem of definition is to first ask ourselves why we 'plan' at all. On reflection, it seems clear that we plan in order that certain things may be accomplished. Planners seek to organize resources (things, people, money) into actions which, if carried out, will lead to expected results. Thus planning is, at a most basic level, an expression of the elemental scientific statement: 'if . . . then . . . ' If these actions are carried out, then certain desirable consequences will occur.

This reasoning lies at the heart of most definitions of planning, some of which (Faludi, 1973, p. 90-1) make explicit the connection between planning and the development of scientific theory. Friedmann, for example, defines planning as the 'application of a scientific and technical intelligence to organized activities'. (1973, p. 19.) I have recently defined planning as 'the application of theory to reality in order to decide what to do, when and how'. (Middleton, 1978, p. 2.)

Thus the essence of planning is the application of theory. Many practising planners and administrators will find themselves uncomfortable with this definition. They see themselves as practical persons with a job to do: theory is something that scholars are concerned with. Moreover, theory is often seen as something that is used to explain natural phenomena, not as a guide to action designed to affect the future.

Yet all planners use theory in their work, consciously or unconsciously (Appleyard, 1977). Whenever a planner recommends a course of action leading towards some goal, theory is being applied. The planner will be able to explain (perhaps with considerable uncertainty) why the proposed actions are likely to have the desired effects. Theory is, in its most basic sense, an explanation; thus this explanation is a statement of theory. The explanation may be based on prior experience, on customary ways of doing things, or on explicit models or theories drawn from scientific knowledge; but it is theory in all cases.

In this chapter I will seek to bring to the surface the ways in which

theory is important in planning. The focus will be on communication planning, although, as we will see, many of the ideas to be discussed are relevant to planning of any kind. Moreover, the discussion will centre on communication planning as part of social and economic development, primarily (though not exclusively) in the Third World.

We will begin with brief discussions of various kinds of communication planning. We will then consider how theories are used in planning, defining two different categories of theory: theory in planning and theory of planning. Major theories within each category will be briefly reviewed in relation to communication planning, leading towards the development of principles to guide the selection and use of theory. These principles form an initial statement of a theory of communication design, linking together a range of theories on the one hand, and the practical context and functions of planning on the other.

# Communication planning

As other chapters in this volume show, there are many kinds of communication planning. A distinction can be made, for example, between planning of telecommunication systems and planning for mass-media development and utilization (Middleton, 1979). It is one thing to plan for a national telephone system and quite another to plan for educational radio. Another possible distinction relates not to the communication structure, but to the reasons why communication is being planned. This view would separate planning of communication as a means of achieving certain goals from communication as an end in itself (Holmstrom, 1978). 'Development communication', as the planned use of communication media to support development goals and programmes, would fall into the category of planned communication as a means to an end. Planning to expand a general system of telephone and telegraph services for general public use might fall into the category of planned communication as an end in itself. The distinction can also be drawn at an individual level-between communication to accomplish work and communication as play, done solely for selfsatisfaction and enjoyment (Stephenson, cited in Schramm, 1973, p. 25-6).

A third perspective on communication planning is holistic. Communication is defined broadly as a social/economic resource that can and should be allocated (planned for) just as any other resource is. In this view, both telecommunications and mass media (or 'communication with a technological base') can be 'planned' as part of a comprehensive development strategy (Hancock, 1978). This approach formed the basis, for example, of the development of a proposed national communication plan for Afghanistan in recent years (Unesco, 1978). Variants of this holistic approach would extend the reach of planning beyond mass media and telecommunications to include a wide range of kinds of interpersonal communication (Middleton and Lin, 1975).

The discussion of the use of theory in communication planning that follows applies, at a general level, to each of these perspectives on communication planning. To the extent that communication planning, regardless of type, embraces both a process of planning and the use of theories within that planning process to reach decisions, the broad structure of the following analysis should be relevant.

# How theory is used in communication planning

Communication planners seek to create action that will in some way affect the future nature of society. In some societies, planners will explain this task as projecting social and individual needs (or demands, in the economic sense) in order to develop communication systems and activities to meet these needs. Communication is, ostensibly at least, designed to meet felt or demonstrated needs of individuals or groups. Examples of this kind of planning can be found in market economies, such as the long-range telecommunications planning done in Australia and Canada (National Telecommunication Planning Branch, 1975). In other societies, planners will work towards goals established to reflect an image of the ideal man and society, using communication to bring about individual and group psychological and behavioural change so that these goals may be met (Chu, 1977). Much development-support communication is of this type, notably in health, family planning and agricultural programmes. These two types of planning are often found in the same society.

# THEORY IN PLANNING: A GUIDE TO STRATEGY, ACTION, EVALUATION

In all cases, actions are proposed in plans with some idea of their probable effects, in relation to some goal. The importance of theory in planning rests fundamentally on this fact. Consciously or unconsciously, planners make assumptions that form links in the chain of proposed actions leading to a desired goal. Take, for example, a radio-based teacher education project designed to raise pupil test scores by improving teacher knowledge of how to teach. This 'strategy' has several theoretical links. One is the assumption that if teachers know more about teaching they will then change the way they teach. Clearly, this link is essential if the end goal, improved student performance, is to be reached through this particular strategy. Theory of this type is often called substantive theory, or theory in planning. It is theory about the causes of the problems the planners are seeking to ameliorate. Alternatively, it can be thought of as theory about the complex relationships among the elements of a system that a planner is seeking to change. This kind of theory is used by planners as an aid in making decisions in the course of developing a plan. Theory in planning helps planners predict outcomes of planned actions, or to work backward from desired outcomes to identify actions that are likely to be effective. In short, it helps planners decide what to do and how to do it.

A few more examples of the use of theory in planning will help us see the range of uses more clearly. Evaluation also depends on theory. The connection is most clear for formative evaluation, through which planners seek to understand how their programmes are working. Once a programme has been designed with explicit use of theories as the links in the chain of actions, planners can identify the critical assumptions that bind these actions together. These in turn become the focus for evaluation. In our example of radio teacher education above, one would certainly wish to collect data on the theoretical assumption that links changes in teacher knowledge with changes in classroom behaviour, as well as on other, similar linkages in the programme design. Without such information, it would be difficult to know, if the programme does not succeed in raising pupil test scores, why it failed. Which assumption has proven to be the weak link in the chain? If the programme succeeds, is it because the theoretical linkages were correctly assumed-or in spite of the possibility that one or more linkages have not functioned as expected? Information of this kind is essential if the programme design is to be used again. Suchman (1967) discusses this concept in terms of a 'hierarchy of goals' in which elements of planned actions are explicitly linked together by assumptions.

Theory also has its place in summative evaluation, which planners use to find out the effects of planned action, both in terms of planned goals and unintended outcomes. At a most basic level, theory has helped planners decide what outcomes are likely or possible, and has thus been an element in deciding on goals. Moreover, specific goals for planned communication are usually justified in some way by reference to some high-level, general and abstract statement of a social goal. The specific goal for a telecommunications expansion project, for example, might be to increase the number of telephones available to rural people by a particular number. Communication planners will have had to argue in support of this goal to get the budget required to achieve it. They may have said that expanded telephone services in rural areas will have important effects on agricultural productivity (through better market information), on participation in

22

development programmes (by making it possible for more rural people to make their needs known), and so on. The relationships between a communication project and these larger goals are difficult to prove because of the many factors outside the communication effort itself that contribute to or work against their attainment. Planners often can only make theoretical arguments, showing the outputs of the communication system as inputs to the achievement of the larger goal.

Similarly, theoretical argument plays an important part in the evaluation of alternate plans and strategies that can be part of the planning process. In evaluating alternative strategies for action, planners must rely on information from past efforts (often carried out in other places and always carried out at other times) or on theory to decide what is likely to work. Pilot projects are the most operational and observable example of the use of theory for evaluation of this type, especially when whole alternative strategies are tested. In pilot projects an attempt is made to make theories operational through real actions.

Various theories can also be brought to bear on the problems of organization and administration, message design, systems design, and so on, as we will see below.

## THEORY OF PLANNING: A GUIDE TO HOW TO PLAN

Planners can also use theory to guide their own actions. These theories can be thought of as procedural theory, or theory of planning. There are many different theories on how planning is, or ought to be, done. Each supports a different approach to the planning task. These theories grow out of different views of society and the place of planning in society. They also rest in part on the nature of the problems that planners address.

# WHY USE OF THEORY IS IMPORTANT

For the forseeable future, at least, planners will make less than perfect plans. This is particularly true for communication planning, where many of the problems being addressed (especially social problems as opposed to engineering problems) are 'wicked' ones, for which no known solution is easily available, and whose shape is affected by whatever planned actions are taken (Rittel and Weber, 1973). Thus a constant and critically important aspect of communication planning is, and will continue to be, explicit attention to learning from the linked processes of planning and action. The explicit and conscious use of theory in and of planning helps planners to understand and manipulate both what they do and how they do it, increasing their ability to learn as they go along. If things work, planners will have a better idea of why; if they do not, why not.

Beyond this, the conscious use of theory can help planners justify and support their proposed actions. This, in turn, can have the effect of making both plans and planning a more open process, exposing the rationales underlying planned action, and enabling individuals and groups with an interest in the planning process to participate more effectively.

# Theories in planning: a brief review

It is commonplace now to say that communication planning draws on many disciplines. This, of course, is equally true of planning for other sectors. Economic planners use economic theory primarily, but increasingly look to theories from sociology and organizational science for guidance. Urban planners use theory drawn from sociology, architecture, geography, and so on.

Communication planners, if we include planners dealing with the spectrum of problems ranging from the engineering aspects of telecommunications to the use of mass media to teach, can draw on an exceptionally broad range of theory areas. This range is, in fact, so broad that few, if any, individual planners will be able to master it. This, among other things, points towards planning by interdisciplinary teams, with attendant problems (Hancock, 1977).

It also points towards a responsible degree of humility with respect to the review of theories that follows. Not all useful theory areas can be reviewed. First, because the size and scope of the paper is limited. Second, there are the inevitable limitations on my familiarity with various theories. The theories I have chosen to include are those with most immediate applicability to communication planning and which I know well enough to review. Thus, for example, many of the theories that engineers use in designing telecommunications systems are not here, even through they are important. Moreover, the theory reviews themselves are not complete. Each field has an extremely large body of literature of its own. The intent of the reviews is to identify the underlying problems around which the theory has been developed, to highlight some of the major components of the theory, and to show how the theory area is useful to communication planners.

Four theory areas are reviewed in turn: development theory, sociological theory, communication theory and theories of organizations. The order of the review is not random, but flows from my earlier discussion of communication planning. Development theory is seen as a sort of 'touchstone' of philosophy and approach that determines the nature of the communication system(s) to be planned for. Sociological theory, as we shall see, provides ways of looking at the societies to be 'developed'. Communication theory forms a kind of bridge between theories of societies and theories about how individuals and groups learn and change, dealing with social and individual functions as well as with processes of communication. Organization theory not only helps planners look at implementation and programme administration, but also provides a convenient stepping stone for the following section on theories of planning.

### DEVELOPMENT THEORY

Most of the world is poor; some of it is rich. This is the essential problem which development theory seeks to address. How can poor nations achieve a better life for their citizens?

Agreement on this statement of the problem probably exists among development theorists with very different views on the purposes and nature of development. Beyond that, however, disagreement on some very important issues breaks out. It is safe to say that we are in a period when the field of development theory is in the process of rejecting old models and searching for new ones.

The old model emerged after the Second World War. Development was defined in economic terms, with the Gross National Product (GNP) being seen as the leading indicator. A better life was, by and large, defined as having a higher income; this, after all, was the way one moved from being poor to being rich. Increased incomes were to come primarily from industrialization of developing societies, and this would be brought about by transfer of capital and technology from rich countries to poor ones. The model was based on the underlying assumption that economic growth was both good and possible in developing societies.

As many critics have pointed out, this theory of development was largely imitative of historical development in Western industrialized societies.<sup>1</sup> The most basic 'if . . . then . . .' explanatory statement of this theory was that 'if developing countries follow the path to development that Western societies followed, and accept the same definition of development as increased income, then the quality of life will be improved'. From this theory flowed much of the action that characterized development strategies through the 1960s. Capital and technology were imported; factories were built; educational systems were expanded to provide the manpower necessary for an industrialized economy. Centralized planning

<sup>1.</sup> Although Kenneth Arrow (1975) believes that the architects of these development models misread the then current economic theory.

of economies was justified as necessary to manage this complex process of capital formation and use.

And this model appears to have worked reasonably well in some isolated instances—in countries such as Brazil, Singapore and the Republic of Korea. But it did not lead to much progress in most of the Third World, and by the 1970s the theory was under heavy attack.

A first target was the kind of development that the theory implied. The GNP is a crude measure of the overall economic output of society: it says little about distribution of benefits within society. Moreover, it does not deal well with non-economic development goals. New indicators of development have come forward, stressing the fulfillment of 'basic human needs' and emphasizing the distribution of wealth and other benefits within society. Equality and social justice, autonomy and mass participation in development, and the importance of ecological balance in the face of limited and increasingly scarce resources became important development goals (Galtung, 1976).

Once the goals of the 'traditional' model were in question, its strategies also came in for critical review. Industrialization was one of the first targets of the debate. It was noted that this strategy accelerated urbanization, creating immense social problems. Somehow, unemployment and underemployment continued, even grew. The 'dual' economy was created, in which élite groups benefited from a modern industrial, consumeroriented economy, while the mass of society in rural and urban areas saw their standard of living steadily decline.

The nature and process of international assistance emerged as a factor hindering national development, creating systematic dependencies between developing and developed countries—with the advantage to the developed. Development problems came to be seen as the consequence of an oppressive international economic system, not as the result of 'backward' or 'traditionbound' cultures (Inayatullah, 1976).

New development theories are in the process of formulation. There has been a realization of the importance of the rural, agricultural sector. There is a movement towards 'self-reliance' in the Third World as a way to break dependency relationships (Pugwash, 1977). The Chinese model of social, ideological and economic development has received much attention as an alternative. Third World countries have begun efforts to achieve a 'New International Economic Order', through which Third World countries can become equal participants in the world economic system. Though such efforts have been largely frustrated to date, they remain a significant force in development thinking (Barraclough, 1978).

Centralized planning is increasingly questioned. If an end of development is increased mass participation, then centralized planning becomes difficult to justify. As economic development models move away from centrally planned industrialization strategies towards rural development, increased agricultural productivity and better distribution of the benefits of development, the value of decentralized planning, which seeks to place decision-making capability where needs are best known and information for planning most readily available (for some problems, at least), becomes more clear.

A consensus on development theory similar to that which held sway in the near past has yet to emerge. However, it is clear that alternatives to the industrialization strategy exist and are being followed—in Tanzania, in China, in Cuba, and elsewhere. The primary effect of the passing of the traditional theory is to increase the alternatives that development planners can explore as models to guide their country's path.

# Implications for communication planning

It seems obvious that different development theories are based on different images of future society. Communication of all kinds is instrumental in the achievement of any imagined society, and it follows that the ways in which communication is developed in society will follow from the dominant social image. The old development strategy had many implications for the function of communication and communication systems. Telecommunications were seen primarily as part of economic and bureaucratic for control) infrastructure. The mass media were to persuade the people to go along with centrally determined policies, priorities and plans. Extension systems brought new technologies to the people so that they could play their part in the grand scheme, whether by adopting an agricultural innovation or contraceptive practice. Campaigns were to change values or attitudes that posed 'barriers' to modernization. Education systems were to be tuned to the manpower needs of the industrial economy.

But things are changing. As new development theories emerge, communication is asked to perform new functions consistent with these new models. The implications have yet to be worked through in general, and perhaps should—or cannot—be, given the plural nature of new development theory. But we can easily imagine some of the alternatives for communication planners. Telecommunications, for example, can be seen to serve the need for mass participation in development. Why can't market information by telex be available in rural villages? Shouldn't a farmers' group be able to telephone problems in to the extension headquarters? Shouldn't the mass media serve as a further vehicle through which the aspirations of different groups are aired? What if campaigns were turned around, with disadvantaged groups using them to persuade government officials to 'do the right thing'?

For communication planners the implication seems clear, and critically important. Planning for communication in society must conform to some development theory. Which one? The old one? The currently popular one? The problem is made difficult by the fact that the theory is in transition, and even more difficult because implementation of the new theories is a slow and piecemeal process that can leave communication planners groping for guidance. Their search, however, will be made easier if they understand these alternative development theories well enough to enter into the development theory debate both within their own country and internationally.

### SOCIOLOGICAL THEORY

Development theory rests on images of what society ought to be, and how to get there. Sociological theory is a way of viewing society and the way in which it changes. The problem which sociology addresses is the explanation of society and social processes, including social change.

In a recent book, Kinloch (1977) has identified three major types of sociological theory. These differ in important ways—in degrees of formality, in underlying assumptions, in the level at which society is studied, in methodology, and so on. Briefly, these three types of theory are as follows.

# Society as integrated structure

This is the sociological theory with the longest tradition. Society is seen as an integrated system of institutions; emphasis is on the structure of society. In simplistic terms, this structure might consist of the social institutions of education, the government, commercial enterprise, political groups, and so forth. Society is seen as moving towards increased integration and efficiency. Institutions perform certain functions, and change as functions change. This theory deals with society at a high level of abstraction.

Social systems are seen as organic, and are sometimes explained by analogy to natural systems. Importantly, society is thought to operate according to a number of natural laws, rather like other natural systems. In the more recent work in this area of theory (structure functionalism), these natural laws are seen to operate as a system.

# Society in conflict

Society can be seen as a set of conflicting and competing forces rather than as moving towards integration. Marxist thought is a well-known example of this kind of theory. Theory in this area still views society in systemic terms, but the system is the result of struggle, not co-operation and integration.

# Society as individuals and groups

In this area of theory society is viewed at the microscopic level, and explained in terms of the behaviour of individuals and groups. This is the province of social psychology and psychology. Society is the collective result of the behaviour of individuals, but the individual is also the product of the society. The interaction of the individual with the social system is an important area of concern.

# Implications for communication planning

Sociological theory is much richer than the very brief summaries above indicate. However, even this simple overview helps us see how theory of this kind enters communication-planning thought. The comments below on implications are meant to be suggestive and exemplary; much work remains to be done on these ideas.

The relationship of sociological theory to development theory is evident: both help planners form images of society. Planners who think of society as a set of integrated social institutions, either consciously or unconsciously, will tend to create communication systems that conform to this image. Such systems are, for example, likely to be oriented towards establishing social consensus; towards minimization of conflict; towards incremental change and adjustment. Planners working from conflict-oriented theory would be likely to reverse many if not all of these characteristics of their systems. In both instances the emphasis would be on communication in relation to and as part of social institutions. The emphasis in planning might be broad and macroscopic. Communication might be linked directly to larger strategies of institutional change and development.

Planners operating principally from the view of society as the collective of individuals and groups might differ further. Their emphasis would tend to be on achieving social change through individual and group change. This approach could be linked with either of the first two theory areas. Individual and group change could be oriented towards integration or conflict. Much of what is known as development-support communication appears to be planned on the basis of a view of society as an integrated set of institutions combined with an orientation towards achievement of social change through individual and group change.

Sociological theory has as its purpose to explain what society is and how it operates. It arose from efforts to deal with problems created by social change (Kinloch, 1977, p. 35-6). Thus it provides an array of ideas on how change occurs and can be made to occur. Because of its broad sweep, it can, and often does, incorporate a variety of theories from other fields—politics, economics, philosophy, communication. In many ways then, the 'if . . . then' assumptions of sociological theory form the framework for the use of development theory as a guide to action. The first tends to be more empirical and rational; the second more normative and ideological. The first tends to say 'this is the way things are, how they got that way, and how they will change'. The second tends to say how things ought to be, and how they ought to be changed.

In the actual policy and planning process, most 'sociological' thinking appears to be buried in assumptions that are rarely stated. For example, recent work on the development of integrated development communication planning is apparently based on the assumption that society is a set of integrated institutions (Hancock, 1978; Middleton, 1979). But these assumptions are not stated. It may be a good thing for planners to examine their underlying sociological-theory assumptions as part of the planning process. Explicit attention to these underlying images may raise alternative actions for consideration; it may also increase (eventually) consensus on plans and their results.

# COMMUNICATION THEORY

It is commonplace now to speak of communication as the web of society, the essential link among individuals, groups and institutions. Thus communication theory tends to cover a wide range of levels and functions. At the risk of oversimplification (for there are many interconnections between levels), these theories will be discussed in two categories. The first is at the level of communication as a social function. This is the macro level, at which communication is viewed in terms of its broad functions in society. The second level can be thought of as the level of communication process and effects. This level includes what is often called human communication, which deals with the problems of how communication works with individuals and groups.

It is worth noting that communication is an exceptionally interdisciplinary field. There are communication theories and concepts. There are communication scholars and researchers. But much of communication theory has been built from other disciplines and the people who work and have worked in them: political science, economics, sociology, anthropology, psychology. As with the other theory areas reviewed here, only the broadest outlines of the theory can be covered, although this will be sufficient to draw out some implications for communication planning.<sup>1</sup>

# Communication as a social function

The political scientist Harold Laswell identified three major social functions of communication. These are broad roles that communication is held to play in society. These are surveillance, correlation and transmission. To these has since been added a fourth, entertainment.

- Surveillance is the scanning of the environment for information. This, of course, is something that all individuals do. But society has developed special roles and institutions to embody this process. Included are diplomats, investigative journalists, and, one supposes, researchers and planners—at least in some of their roles.
- Correlation can be thought of as the co-ordination of response of different parts of society to the environment. It includes the interpretation of information on the environment, policy-making, planning, operation of organizations, and—at the level of individuals and groups—the 'co-ordination of public understanding and will; [the] operation of social control' (Schramm, 1973, p. 31).
- Transmission of social heritage, customs and laws is the third function. We commonly think of education systems as performing this function, but in the communication sense, at least, this function encompasses much more: the communication that occurs between parents and children, and through other institutions—libraries, museums, etc.
- Entertainment is an important fourth social function, one that consumes many mass-media resources. This function can also lead to learning of various kinds.

These four functions tend to be independent of the type of society or political system. They are held to be four necessary categories of communication through which the 'web of society' is established. Overall, we can summarize, in the words of Melvin DeFluer, the communication act as 'the means by which a group's norms are expressed, by means of which social control is exerted, roles are allocated, co-ordination of effort is achieved, expectations are made manifest, and the entire social process is carried on. . . Without such exchanges of influence human society would simply collapse' (1970, quoted in Schramm, 1973, p. 30).

1. This review is drawn primarily from Wilbur Schramm (1973).

# Communication process and effects: the individual level

Communication serves these social functions as a process that leads to certain effects. This process involves, at the most basic level, a relationship in which information is shared through a communication act. The relationship that we most often think of is between people, in pairs and groups. But there are other relationships, such as between computers linked with telephone lines. Some of the relationships between people operate over such long periods of time that we tend to forget their essential nature—as when we read a book by an author long dead.

There is also a tendency to identify communication theory as dealing, essentially, with the media of communication; and there are theories on the mass media. But as our basic definition indicates, the media are one link in the chain of actions that taken together constitute the communication process.

Communication theory at this level is varied and complex, and difficult to summarize fairly and completely. However, we can take a look at some of the many questions that this theory deals with in order to get a feeling for the ways in which it can be useful in the planning process.

The theory of information: Information is a central component of the communication process. Much attention has been given to the definition of information and analysis of its functions. A major idea is that information is anything that affects the level of certainty a person has in a situation. This concept clearly relates information to decision-making.

The communication relationship: Persons enter into a communication relationship with some intention, or purpose. The purpose of the relationship affects the nature of the communication process. This idea has led communication scholars to differentiate communication processes according to goals. Kincaid and Schramm (1975) identify four such purposes: informing, persuading, instructing, and establishing dialogue. Different theories about the communication process, including ways in which information is structured, have been developed for these different purposes.

The psychology of communication: A number of theories have been developed to explain how the communication works. An important example is the idea that the meaning of a message—or its effect—lies as much with the receiver of the message as it does with the message itself. It is the cognitive make-up and communication skills of the receiver that determine if the message is perceived and accepted, and how it is interpreted. We often think of this as theory of perception and meaning. The sender of a message encodes it by expressing meaning in symbols. These may be verbal, written, pictorial, or non-verbal. Depending on the communication skill of the sender, these will represent to a greater or lesser degree the meaning that is to be conveyed. The symbols are transmitted through a channel. These channels may or may not involve media—a printed page, radio, etc. The receiver, if the message reaches him, then decodes the symbols and interprets their meaning. In decoding and interpreting, the receiver is guided by his/her existing knowledge, values, attitudes and communication skills. The receiver may then provide feedback to the sender by sharing a certain message.

Thinking like this has led to the development of a number of theories on how communication achieves different effects: for example, on knowledge and on attitudes. An important contribution to theories of communication effects is learning theory, which has developed in psychology and education, and has many facets and at least two different components—cognitive theory and behavioural theories. Both of these (but especially the latter) have been developed into instructional design models and techniques that guide the planning of teaching and learning.

Communication effects: Researchers have identified a number of components of communication effects. These include theories about the source of messages, the style of persuasive communication, the order in which information should be presented to achieve persuasive effects, the role of appeals, the effects of group membership, and so on. These theories help explain why and how communication can help achieve different kinds of goals.

The media of communication: There are two approaches to the study of the media of communication. In one approach, theory is developed about how different media function. In terms of communication effects a major distinction is made between mass and interpersonal media, although such instruments as the telephone tend to make the distinctions not totally clear. In studying the mass media, communication theorists have developed some ideas about how the mass media relate to interpersonal communication. Notable among these ideas are the 'two-step flow' theory, which explains how information flows through mass media to 'opinion leaders', who in turn pass the information on to others. The theory of the diffusion of innovations fits loosely here, explaining the relationship between mass-media information (as well as interpersonal communication and influence) to the process through which an individual adopts a new process or technique. The work of Marshall McLuhan has raised, though not demonstrated, the possibility of inherent effects of the media themselves, independent of their content.

In a second approach, theory is developed on the structure and organization of communication systems in society. This is clearly a complex field. It encompasses studies of the press, broadcasting and the telecommunications infrastructure. Studies showing the contributions of the information industry to the gross national product of a single nation have been completed (Porat, 1976).

Research of this nature is clearly oriented to policy. It deals with issues of media ownership and control, and of the financing of communication systems. What are the effects of private versus public ownership and control of media? What are the effects on communication in society of a media system financed through advertising?

The models of communication: Much work has been done to find ways to reflect the various components of communication theory as models that describe the entire process. There are engineering models, which are used to predict the performance of such things as telecommunication systems. There are models that focus on the process of human communication.

These models tend to have common elements. Almost all, for example, include a message source, a message, a channel, a receiver and feedback. Many contain more elaborate elements, including the processes of encoding, decoding and interpreting information. The general model shown in Figure 1 is typical of many.

Other models have been developed which emphasize different aspects of the communication process. Convergence models (Kincaid, 1979) portray communication as a dynamic, cyclical process through which individuals move closer together in terms of their common understanding



FIG. I. A general model of the communication process. Source: Kincaid and Schramm (1975).



FIG. 2. The convergence model of communication. Source: Kincaid and Schramm (1975).

of concepts, ideas, values (Fig. 2). This model grows out of and is closely linked with network models (Fig. 3) of communication. Social networks portray the information and influence relationships among individuals and groups. They are developed through empirical research.

There are other models of course. And it is probably safe to say that there is no 'correct' model of communication. Instead, there is a variety of models, each presenting a particular way of looking at the processes through which communication of different kinds takes place.

# Implications for communication planning

Theory on communication as a social function describes, in the broadest sense, roles that communication plays in society. These are thought to be necessary roles, essential if society is to function. Planners dealing at the most general or macro level of communication systems can use concepts such as these as overall guides to the development of communication systems. Alternatively, they can be used as rough criteria to examine the state of communication systems at any one time. What are the systems that aid correlation? Transmission? Surveillance? Are they functioning effectively? How much of available resources should go to entertainment functions of communication? How important is this function in a particular society?

In short, theories or images of the social functions of communication can help planners frame the broad questions that guide policy and resourceallocation decisions. These theories (or other ones that may exist or be developed to explain the same level of communication problems) help planners and policy-makers deal with some of the important 'why'





questions: Why should we invest in social research (surveillance)? Why should we develop teletype systems linking rural areas with large cities (correlation)? Why should we invest in a national curriculum for schools (transmission)? Why is it important to retain, or encourage, traditional forms of art, dance or drama (entertainment, transmission)?

Theory on communication process and effects at the individual and group level serves a different planning function. This theory helps planners design communication strategies. At this level planners are concerned with designing messages to share through selected channels or media to achieve certain effects. Different components of communication theory are brought to bear at this level on different kinds of problems.

The idea of purposes of communication seems especially important. Depending on the goals of planned communication, different theories can be most usefully applied (Middleton and Lin, 1975). For example, planners working to increase levels of information or awareness of certain ideas might apply aspects of information theory, drawing on concepts from the psychology of communication, especially as it relates to perception and meaning. Planners working to change attitudes, values or behaviours of individuals and groups (such as in a campaign to prevent cancer by persuading people to stop smoking) would use theories of persuasion. They might even draw on applied components of communication theory, such as marketing. Planners working to teach (instructional goals) through media, such as radio, might draw on learning theories and their applied versions, models of instructional design.

The models of communication help planners form overall images of how communication strategies may have desired effects. The concept of social networks, for example, linked with the idea of the 'two-step flow through opinion leaders', is often used as the justification for campaigns that combine mass media and forms of interpersonal communication. The diffusion-of-innovations theory appears to have been the basis for many health, family planning and agriculture campaigns.

The application of these theories at this level is as much an art as a science. In particular, we need to learn much more about how theories may be most usefully combined as the basis for action.

It is especially important to recognize that most of these theories have been developed in the United States. Their validity to other cultures has not been well established, although they have often been used. This has two important implications. The first applies most directly to planning in communication at this level. This is the importance of continuing testing and evaluation of programmes and materials based on these ideas to determine, at a practical level, if things work the way they are intended to. The second is a challenge to communication research: we need to know
much more about the cross-cultural transferability of these theories. Perhaps most importantly, we need to find out if there are other theories and models that provide more valid and useful 'if . . . then' statements for non-Western cultures.

#### ORGANIZATION THEORY

Planners work both in and through organizations. Organizations form the context in which planners act, and are the principal resources through which planned actions are carried out.

٠

Organization theory seeks to explain how organizations function and can be made to function more effectively. There are several traditions in organizational research and theory building. The first, and oldest, tends to be normative. Through rational analysis and argument theories are constructed that say how organizations ought to be structured and operated. Organizations are seen as structures to be manipulated. A second, more recent tradition relies more on empirical observation of how organizations actually do function; the emphasis is on organizations as patterns of interaction among people. A third and emerging area of theory seems to blend aspects of both with systems theory. All of these traditions have important things to say to communication planners.<sup>1</sup>

#### Organization as structures

Not surprisingly, organizational theory has common roots with sociological theory: organizations are significant social institutions. Max Weber's analysis of bureaucratic organizations is such a common starting point. In this perspective, organizations are seen as goal-oriented, created and structured to achieve efficiency in the use of resources. They are rational and information-based in the way decisions are made. Rules are established to guide actions of those who work in the organization (the 'members'). Members are organized hierarchically, with clearly defined responsibilities and authorities, and are chosen on the basis of their competence to perform required tasks. The image of such an organization often held is that of the great bureaucratic machine, with people functioning as efficient cogs. They are motivated to work primarily by economic rewards.

A variety of subsequent normative theories has been built on this foundation. One is the scientific management school, which began at the turn of the century with the work of Frederick Taylor. In this perspective the worker came truly to be perceived as a component of a machine designed to

<sup>1.</sup> This review is drawn principally from Haas and Drabek (1973) and from Scott and Mitchell (1972).

achieve highest efficiency and productivity. The organization was seen as exceedingly rational. Later theories emphasized the function of decisionmaking in rational organizations. Concepts associated with these perspectives included such ideas as the span of control (the number of employees one could supervise effectively) and rational models of the organization process such as POSDCORB (planning, organizing, staffing, directing, co-ordinating, reporting, budgeting).

## Organizations as people

A key assumption in the normative theories was the belief that if employees are properly paid, they will then perform as desired. This assumption began to be questioned as empirical observation of organizations in action expanded. The central importance of the human beings inside the rational structures of traditional theories became obvious. It was found that employees respond to many needs other than that for financial reward, that human problems created loss of efficiency, and that organizations are, in fact, not as rational as would appear. At the same time, organizations were growing larger and much more complex.

Research and thinking of this type led to a variety of new perspectives. Part of this has been known as the neoclassical school of thought, in which explicit recognition of the human factor in organizations has emerged. Personality needs and their effects on organizations are recognized in the work of Chris Argyris and others, with influence from the field of psychology. The human relations school emphasized the resolution of organizational problems through small group work, greater employee participation, bottom-up planning, and the like.

## Organizations as people in systems

Perhaps no other idea has had so much influence on the study and operation of society as general systems theory. This is a way of looking at reality which is in essence quite simple but which in its implications is quite profound. In systems thinking, the world is seen as a complex set of interdependent subsystems. Everything is related, in some way, to everything else. When applied to smaller (than the world) systems, this line of thought identifies several critical elements of a system: (a) its objectives; (b) its environment; (c) its components; (d) the relationships among the components; (e) the resources of the system (adapted from Churchman, 1968).

Organizations can easily be seen as systems. In earlier organizational theory, organizations were viewed as natural systems, organic in their

nature. Not surprisingly, much of this thinking was done by sociologists who viewed society as a natural system (cf. page 28). Current thinking, however, tends to see organizations as man-made systems, artificial phenomena which 'are as they are only because of a system's being molded, by goals or purposes, to the environment in which it occurs' (Simon, 1969, ix.)

Current theory of organizations draws on systems thinking and ideas about organizations as made up of the complex interactions of people in them. An important example of such theory is the 'stress-strain' theory (Haas and Drabek, 1973), in which the dynamism of organizations is seen to flow from stress and competition between members. Another important aspect of the systems view of organizations is the idea of organizations as open systems, drawing resources from the environment in return for a valued product.

This emerging synthesis of theories appears promising, if only because it provides a framework within which much earlier theory can be incorporated usefully.

## Implications for communication planning

We have noted that planners work both in and through organizations. The implications of organization theory for planners in organizations will be discussed in some detail in a later section of this chapter. At this point, it is probably sufficient to note that the way in which planners do their work will be affected by the nature of the organizations they work in, and that these in turn will have been designed, consciously or unconsciously, with some theory of organization in mind.

When we turn to implications of the theory for the work that planners do through organizations, implementation springs immediately to mind. Communication planners operate with a number of theories or images in their minds at one time. Some, as we have seen, have to do with the nature of development, of society and of communication. These theories, among others, help planners to design and create communication systems. At the same time, planners have theories or ideas about how to make their systems come to life—in short, how to implement plans. These ideas encompass resources of many kinds, and a key resource is the organization that will carry out all or part of a plan.

Failure in implementation is a common charge against plans and planners. Often this is attributed to faulty analysis and lack of attention to the 'administrative' component of action. It may be that poor analysis and lack of attention are surface symptoms of a more pervasive underlying problem, one which communication planners may not be able to deal with. This is the potential inconsistency between the way in which organizations are seen by planners, and the reality of the way they are. A simple example is easily found. Planners may expect organizations to carry out efficiently their plans, only to find that the human dynamics of organizations in fact forms barriers to 'obviously' good and rational plans. Here planners operate with a normative, bureaucratic image, only to find reality demonstrating 'organizations as people'.

Planners may thus face a critical 'trade-off' decision. Let us assume that they can, using theories of organization and common sense, reach a reasonably accurate diagnosis of how implementation agencies actually function. Then, do they design communications systems that can be implemented with the structures at hand? Or do they attempt to redesign organizational structures capable of implementing an alternative (presumably 'better') communication system?

Regardless of the answer, the key step is the analysis and understanding of organizations as they are. It is in this task that various theories of organization have their value.

It is worth noting the parallel between sociological theory and organizational theory. In practical terms, both help the planner understand more clearly the environment in which planning and action are to occur. In theoretical terms, both grow from a common tradition, and both share methods of investigation and concepts. For example, conflict theories of society have their counterparts in organizational theory.

The importance of the parallel is fundamental to many issues in communication planning. Many of the 'newer' ideas about planning—participation of key actors in organizations, to take one example—flow from underlying theories about what organizations are or ought to be. How to carry out such an idea if the normative ideal of the organization does not match reality? How many 'co-ordination' mechanisms fail to achieve desired results because they are based on ideas of organizations as open systems, exchanging resources, when in fact they may be faced with closed systems operating according to interpersonal dynamics with historical origins?

# Theories in planning: a pause to summarize

The foregoing review of theories has been brief and simple, perhaps too much so. Important theories in planning for communication planners, notably economic, have been left out, and others—such as systems theory and learning theory—hardly more than mentioned. However, there seems to be enough substance to step back from detail to look at the larger implications.

#### MULTIPLE THEORIES

As Table 1 illustrates, the range of theory with potential for application in communication planning is indeed great. In this chart are summarized the theories reviewed here. The implications of all theories are briefly noted.

Perhaps the range is why communication planning, and indeed social planning of all kinds, appears so difficult, challenging and likely to fail. How can these theories be effectively applied, particularly in combination?

The problems are formidable, but not insurmountable. Let us consider a few of them in turn, and see what might be done.

#### Knowing the theories

Writing this chapter has been a humbling reminder of the difficulty of achieving enough familiarity with theories to be able to use them. Clearly,

	Development theory	Sociological theory	Communication theory	Organization theory
Components	Capital- intensive, industri- alization model Social- development model	Society as integrated institutional structures Society in conflict Society as individuals and groups	Social functions Surveillance Correlation Transmission Entertainment Process and effects Information Communication relationship Psychology of communication effects Media of communication Models of communication	Organ- izations as structures Organ- izations as people Organ- izations as people in systems
Implications	Purposes, goals and nature of communi- cation systems within national development	How society is structured; how it changes; degree of change contemplated	Broad social functions of communication systems Strategies: how to make communication work	Patterns of implemen- tation

TABLE 1.	Theories	in	planning:	imp	lications
----------	----------	----	-----------	-----	-----------

planners must establish some mechanism to develop knowledge of theory within the planning setting. This must be done in addition to the more accepted, though equally complex, tasks of developing data bases, negotiating political agreements, and applying methods and techniques.

A familiar mechanism is the use of consultants. Another is planning through interdisciplinary teams. Both of these have inherent weaknesses, a chief one of which is the tendency for ideas to get lost in the routine press of planning. One way to deal with this problem might be the creation of a theory bank as a complement to more standard data bases. This might be nothing more elaborate than a notebook of reports done by consultants and planning staff on selected theories and their applicability to the kinds of communication problems planners are working on.

# Exposing the hidden theories

Much has been made in this chapter of the idea that planners often apply theories unconsciously. These are 'hidden' theories or assumptions that guide such things as the choice of problems, the establishment of goals, the collection of data, and so on. In effect, they permeate planning. It could be useful to conduct a series of exercises within the planning team through which individuals could be encouraged to expose these theories for discussion. This might well result in more consensus within the team, and it could also result in more careful establishment of the theoretical assumptions on which plans rest. Many of these assumptions are likely to be original, creative and not found in theory books. These are the theories of experience that planners have learned through their work. These theories, too, should go into the theory bank.

# Testing theories

Most theory, at least the kind found in books, has been developed in a very few cultures; yet we apply it in many. Many of the theories of experience that guide planners have, because of their hidden nature, not been overtly tested. It would seem essential that, once theoretical assumptions in plans and planning have been identified, they be explicitly tested in reality. This testing need not be done through formal research: many evaluation strategies will be adequate to give planners some degree of confidence in the effectiveness of some theory or assumption. Information of this sort should also go into the theory bank.

## ON TO THEORIES OF PLANNING

With these last comments we have begun to move into the planning process, speculating on certain aspects of how planners apply theory to reality in order to reach decisions. In the section that follows we will review a number of different theories or perspectives on the planning process itself.

# Theories of planning

What should planners do? How? How should they be organized? How should they relate to each other and to other individuals, groups and institutions? In short, what is the planner's role and how should it be played?

These are questions of central importance to planning of all kinds, including communication planning. Theories of planning (or planning theory) do not provide simple answers to these questions. They do, however, describe alternative models of and approaches to planning, relating these to different kinds of problems, conditions and purposes. Thus, taken as a whole body of thought, planning theory provides a range of concepts which can help communication planners develop an approach to the planning process which is appropriate to their situation.

This basic idea needs, perhaps, some emphasis. There are different ways to plan. This should not be too surprising, given the considerable complexity and uncertainty that comes with planning of social programmes. There is, however, a tendency for communication planners (and indeed planners in many fields) to assume that there is one approach to planning. This approach varies, but usually consists of some set of steps such as (a) problem analysis, (b) policy and goal setting, (c) evaluation of all alternatives, (d) choice of the best alternative, (e) establishment of objectives, (f) implementation planning and (g) evaluation, etc. And this is one useful way to plan.

But there are many alternatives. The literature on planning theory contains more than twenty different approaches (Bryson, n.d.). These differ in greater or lesser degree from the model described above, but the differences are important.

To make matters confusing, a clear structure for classifying and making sense of the differences has yet to emerge, save perhaps at the most abstract level (see Friedmann, 1973; Faludi, 1973; Bolan, 1975 for attempts in this direction). Planning theories tend to be distinguished from each other along several different dimensions; for example, the degree of stability in the social environment; the degree to which goals are fixed or set as part of planning. These differences are sometimes subtle, and each theory seems to share some characteristics with another, while differing sharply with others.

A full analysis of this complex problem is beyond the scope of this paper. Happily, there are some broadly contrasting models of planning. By reviewing these, we can highlight some major differences and draw out implications for communication planning.

The review will begin with a brief discussion of rational thought and the function of values as basic issues in planning theory. Two pairs of contrasting planning theories will then be reviewed in turn: (a) the rational/comprehensive mode and the disjointed-incremental mode; (b) allocative planning and innovative planning. The discussion will then turn to a brief review of selected additional planning theories—participatory planning, transactive planning, and radical planning.

## RATIONAL THOUGHT AND VALUES: THE HEART OF PLANNING

All planning is concerned with two central issues. One is rational action. At the simplest level, this means acting or planning to act with reason, showing by logic, theory, data—or all three—why proposed action will lead to desired results. Rational communication planners propose a new national radio network by arguing logically and empirically that it is necessary because without one, large segments of the population cannot be reached with development information; and theoretically that a certain kind of education programme on radio will stimulate new and desired behaviours. Rational arguments are central to planning because they demonstrate the assumptions and knowledge on which plans rest. This process in turn is essential for the analysis, criticism, and negotiation among interested individuals and groups that make 'participation' in planning possible.

Rational argument, as we will see in our review of theories below, does not mean certainty or the truth. It simply means that the reasons and reasoning that lie behind conclusions are laid bare.

A second central issue is values. Values have to do both with what ought to be done and with how things ought to be done. Thus values are important both for ends and means; both goals and courses of *action*. Development communication planners will immediately recognize the crucial role of values in trade-off decisions between communication investments. For example, should telephone services be extended to urban industrial centres or to pockets of rural poverty? Rational arguments can be made for both courses of action; a final decision will inevitably involve values (often expressed as social goals). Similarly, values will play a part in determining how the decision is reached: who should participate in making such a decision? Government officials only? Industrialists? Rural villagers? Values are involved in choices of how to achieve extended telephone services: should foreign telecommunication firms be retained? If so, should indigenous participation in a joint venture be sought, even at the risk of a possible loss of efficiency?

It is possible to view the evolution of different theories of planning as the result of the interplay of three factors. One is the rapid growth of science, technology and knowledge. A second is the concept of rationality. The third is the existence of different sets of social and individual values. The questions may come down to: What do we know? What can be done? What should be done?

SELECTED PLANNING THEORIES: A BRIEF COMPARATIVE REVIEW

Planning theories are not clearly-cut, agreed-upon models. The labels used tend instead to describe rather broad sets of ideas and approaches with underlying unity and similarity. In the discussions that follow, each theory will be described rather simply, with a tendency to describe rather 'pure' versions of the theory in order to highlight important differences between theories.

## Rational/comprehensive v. disjointed-incremental planning

We begin our review of planning theories with this pair for two reasons. First, the rational/comprehensive mode of planning is widely thought to be an 'ideal' way to plan. Secondly, the contrasting disjointed-incrementalist approach to planning—as its name implies—stands in stark contrast to the rational/comprehensive mode in many ways, so much so that it is often accused of being 'non-planning'.

## Rational/comprehensive planning

Rational/comprehensive (R/C) planning is characterized by a centralized linear planning process, emphasizing the flow of planning thought from needs analysis through evaluation. Much emphasis is given to evaluation of alternative plans as part of the planning process. In the ideal state, all possible alternatives are evaluated, with the optimal alternative, usually in terms of cost-effectiveness, being selected. The process repeats itself to the extent that feedback from evaluation is used to modify plans, usually over some relatively long time frame. This is a basically simple approach to planning; consequently, it is very widely imitated, though it is rarely completely followed.

#### Images and action

R/C planning emphasizes the formal use of knowledge and theory in planning. There tends to be significant emphasis on the use of planning techniques, including mathematical and other kinds of models. Planning is seen as a staff function. The ends—or goals—of planning are seen as falling outside the sphere of the planner's competence. Goals and policies are set through political processes, separated from planning. Thus planning tends to be a technical staff function, focused on means to attain given ends. It follows naturally that there is considerable emphasis on the plan. Plans are seen as the guiding blueprint for the actions of implementors, who typically are not part of the planning group. This separation of plans and implementation leads to the use of centralized authority to bring about compliance with plans. Society is seen as a set of co-ordinated institutions, moving towards integration and efficiency.

Clearly, a great many assumptions underlie this approach to planning. Perhaps the most important one has to do with the degree of certainty in the planning environment. The rational/comprehensive approach rests heavily on the assumption of a high degree of certainty: in the reliability of the data used; in the relevance and strength of theories in planning; in the stability of the problems being addressed and the social/political environment in which problems—and planning—exist. Where such certainty exists, the rational/ comprehensive approach can bring the power of analysis and systematic, rational thinking to the long-term solution of important problems. The development of telecommunication infrastructures, as a planning problem, represents the kind of situation in which this level of certainty may be approached.

The problem of dealing with uncertainty and with large amounts of information is central to rational/comprehensive planning. A number of strategies have been developed within this approach to handle the problem (Faludi, 1973). One is routinization, in which programmes and responses are developed for routine application to certain kinds of problem. A plan can consist of a number of such routines. In communication planning, such a routine may be seen in procedures for preparing educational software for radio or television. A second strategy is sequential decision-making. Large problems are broken down into smaller problems, rules are set for decisionmaking on these smaller problems, which are arranged in a certain order and solved in sequence. Clearly this strategy rests on the quality and relevance of these rules. This procedure is often used in evaluating alternatives. For example, in communication planning, one might evaluate alternative media combinations by establishing criteria and point scores for such things as reach, accessibility to programming, suitability for a particular message, and so on. Various combinations would be rated, and the one with the highest point score selected. A third strategy is mixed scanning,

in which broad, fundamental decisions are made based on available information; these are then used to make smaller decisions. Basic decisions can be modified as necessary as planning proceeds. Suppose planners have as a goal to improve communication linkages within rural areas. They may consider several broad alternatives: telephone service, community radio, mobile vans with video-tape recorders, and so on. Rather than work out complete plans/strategies for each of these alternatives before choosing, this initial choice can be made on broad criteria, often including values. A choice for community radio, for example, is then followed by detailed planning and analysis. If this planning were to reveal difficulties—such as lack of personnel to man community radio stations—the original decision might be modified. The advantage of this strategy is concentration of planning resources; the disadvantage, clearly, lies with the initial broad decision being made without full information.

## Disjointed incrementalism

There has been a great deal of criticism of the rational/comprehensive mode of planning. As a form of development planning it has been charged as a costly failure (Wildavsky, 1973). The degree of fit between the real world problems and the data, models and theories that form the basis for R/C plans has been sharply questioned (Strauch, 1976). We may know less about the world than we think, especially with respect to 'wicked' social problems, as opposed to 'tame' engineering problems (Rittel and Weber, 1973). Other assumptions of this approach appear doubtful for various reasons. Centralized authority conflicts with emerging values and concepts about the nature of the development process. The separation of ends and means, of goals and planned actions, may be artificial: it certainly contributes to the 'failure of implementation'.

Perhaps the most interesting criticism of the rational/comprehensive approach has been described by Friedmann (1973). Noting that effective rational/comprehensive planning requires high degrees of certainty of data and theory, and high levels of certainty in the problem environment, he points out that in planning situations where rational/comprehensive approaches are most useful, they are least needed; conversely, where the rational/comprehensive is most needed (for significant and difficult problems, with associated high degrees of uncertainty), they are least useful.

Disjointed incrementalism stands in stark contrast to rational/comprehensive planning. As developed by Lindblom (1968) and colleagues, this theory of planning outlines an approach that describes much day-to-day planning. The disjointed-incrementalist approach is founded on the lack of certainty in the problem environment, the unreliability of data and theory, and the limited capabilities of planners to know and control the future.

This approach emphasizes decentralized bargaining among established institutions for limited resources, a process called 'partisan mutual adjustment'. Planning proceeds in small steps, with limited consideration of alternatives. Decisions are made 'on the margin'; planners seek to do the next best thing that appears workable. Strategies are evolved through 'successive limited comparisons', not through full rational analysis, although each choice, within its limits, may be rational.

Policy, goals, planning and implementation tend to become intertwined. Society is seen as consisting of many competing institutions, not capable of achieving integration. Control and authority are seen as relatively decentralized. The emphasis on small steps, and on 'marginal analysis', tends to lead towards emphasis on planning as a constantly evolving learning process, although evaluation may be relatively informal and focused on a limited number of outcomes. Autonomy of action for any planning unit tends to be relatively great.

Critics of the disjointed-incrementalist approach often call it 'nonplanning'. Indeed, the approach is commonly known as 'muddling through'. Clearly, it offers an extreme alternative to rational/comprehensive planning. Significantly, it grows primarily out of criticisms of the failures of the rational/comprehensive approach. As a planning theory, its primary strength lies in its correspondence with reality, and its emphasis on the interrelationships between choice of ends (goals) and means (strategies of action).

# Allocative v. innovative planning

Rational/comprehensive planning and disjointed incrementalism differ sharply along several dimensions—the degree of certainty possible, the nature of control and authority, the separation of ends and means. Allocative and innovative planning differ primarily according to the nature of institutions and the way in which resources are employed. As we will see, these theories cut across the dichotomies between rational/comprehensive and incremental planning.

## Allocative planning

Allocative planning is concerned essentially with the distribution of limited resources among competing users (Friedmann, 1973). It is thus oriented towards maintenance of existing patterns of distribution of power and influence. This form of planning is characteristic of most bureaucracies. Rational/comprehensive approaches are often associated with allocative planning. A well-known example is the Program Planning, Budgeting System (PPBS).

The image of society is again of integrated institutions, operating rationally. Authority may be more or less decentralized, but is formally allocated throughout the system. Change is incremental, with a strong tendency to project and maintain dominant interests in society. The emphasis is on short-term benefits and plans.

## Innovative planning

This mode of planning emphasizes action through institutional change. In its most pragmatic form, innovative planning appears through the setting up of new institutions to carry out change programmes. The emphasis is on the mobilization of resources, as opposed to the allocation of existing resources.

The establishment of new institutional structures to achieve significant change leads to a relatively uncoordinated, competitive system. New institutions may form networks of change. After a time, these change-oriented institutions may stabilize and become part of an allocative planning process, in which case the cycle may be begun again with new institutions. Friedmann finds evidence of innovative planning as part of development strategies in developing countries (1973, p. 59).

Innovative planning employs incremental approaches, if only because of the newness of the institution and the problems it seeks to address. 'Planning' is sometimes difficult to locate, much of it being done in fragmented, informal ways.

### SOME ADDITIONAL PLANNING THEORIES

Some of the major dimensions of difference among planning theories have emerged in the brief analysis above. There are other important differences, many having to do with the style of planning. Three additional planning theories are reviewed briefly below to gain further perspective on these differences.

## Participatory planning

This is a style of allocative planning that has emerged in response to the desire of people affected by a plan to take part in its formulation. In this kind of planning individuals and groups become involved in a variety of ways, sometimes with encouragement from central planning authorities, sometimes spontaneously. The challenge seems to be to develop effective participation mechanisms. There is some evidence, for example in the United States, that programmes to increase participation lead to stabilized control of interest groups over the planning process; to increased disagreement; to delayed implementation. Other examples, notably in industry, point to more positive results.

# Transactive planning

Here again the emphasis shifts. Transactive planning is centrally concerned with human growth and learning. This approach emphasizes direct personal involvement of people affected by planning in the planning process. Transactive planning changes knowledge into action 'through an unbroken sequence of interpersonal relations' (Friedmann, 1973, p. 171). The emphasis is on the establishment of effective communication links between technical planners, who rely on processed, abstract knowledge, and clients, who rely on personal knowledge.

This approach necessarily emphasizes decentralization of planning and authority to permit a great deal of interpersonal dialogue. Plans are evaluated not only in terms of what they accomplish for people, in terms of services, but also in terms of their effects on people—their values, their competence, their sense of self-worth (Hudson, 1979).

# Radical planning

Barclay Hudson (1979) perceives two streams of radical planning that 'occasionally flow together' (1979, p. 8). One stream approximates 'grass-roots' collective action, with emphasis on self-reliance. The government plays a minimal role in planning, with control resting with the people. The work of Schumacher (1973) and Ivan Illich (1973) is representative of this stream of radical planning thought.

The second stream of radical planning thought encompasses large-scale processes, including conflict between economic classes, social movements, the effects of colonialism and dependency relationships. The emphasis in this stream of thought is on political theory and institutional structures. Planning as customarily defined is not the answer; radical re-structuring of society is.

# Implications for communication planning

## DIVERSITY OF THEORIES: A PLANNER'S CHOICE

The primary implication of theories of planning for communication planning lies in their diversity. There are many approaches to the planning process, and these differ along important dimensions. Thus communication planners face important choices not only about the communication problems they address and the solutions they propose in their plans, but also about the way in which planning will be organized and carried out.

The existence of a choice implies the need for criteria, or ways for choosing. How can communication planners decide which planning theory is most relevant and useful for their particular situation?

This is a complex problem indeed, made even more difficult by the fact that, as Hudson (1979) points out, it is not merely a question of choosing a single planning theory or approach. Different kinds of problem and situation may call for either a combination of planning approaches, or the creation of a hybrid approach drawing on several theories.

Our brief review of selected planning theories provides some initial ideas about how planners might approach the task of selecting or modifying planning theories. The dimensions along which the theories differ seem to fall into two categories. The first category contains five factors that describe characteristics of the environment in which planning takes place. One is the degree of certainty of knowledge—of the problem, of resources, of theories in planning. The second is the degree of change contemplated—is change to be massive and widespread, or smaller and more narrowly focused? The third has to do with resources; are they primarily to be allocated? Bargained for? Mobilized? The fourth is the values that are associated with planning in society—are they oriented towards control, action, learning or human growth? The fifth category is that of the predominant social image—how do planners view the society in which they live and plan? Is it stable and integrated? Fragmented and competitive? Based on institutions or on individuals and small groups?

The second category contains factors that typically describe key process characteristics of different planning theories or approaches. The first is the degree to which ends and means are separated in planning: do planners set goals or only define programmes? Second is the way knowledge is organized and applied in the planning process: Is this a formal or informal process? Is it comprehensive or limited? Third is the nature of power and authority in planning: Is it centralized? Decentralized? Shared among competing groups?

Each of the planning theories reviewed above can be described in terms of these two categories of factor. As shown in Table 2, each theory is associated with a different pattern of environmental and process characteristics. The table thus represents an initial guide for the selection of a particular planning theory or theories. Given certain environmental characteristics, different planning theories seem more appropriate. These theories in turn have certain process characteristics.

It must be emphasized that the conceptual map represented by Table 2

Characteristics of the planning environment				Planning process characteristics				
Certainty	Degree of change	Resources	Values	Social image	Planning theory	Ends/ means	Knowledge use	Authority
High	$High \to Low$	Allocated (?)	Control	Integration	Rational/ Comprehensive	Empha- sizes means	Formal, highly organized	Centralized
Low	Low	Bargaining	Consensus Learning	Conflict, competition	Disjointed- incremental	Both	Informal, limited	Decentralized
$High \to Low$	Low	Allocated	Stability	Integration	Allocative	Means	Formal	Centralized → Decentralized; Formal
Low	High	Mobilized	Action Institutional change	Competition	Innovative	Both	Informal, fragmented	Decentralized
$High \to Low$	$\operatorname{High} \rightarrow \operatorname{Low}$	5	Participation	Integration Competition; Individuals/ groups	Participatory	Both	Formal → Informal	Shared
$\mathrm{High} \to \mathrm{Low}$	$High \to Low$	Applies to all conditions	Human growth Learning Dialogue	Individuals/ groups; Integration	Transactive	Both	Linking processed/ personal knowledge	Shared
\$	High	?	Self-reliance Social trans- formation	Conflict	Radical	Both, empha- sized ends	Fragmented	Decentralized Shared

is very tentative. Much more work needs to be done in this area (note for example the question marks in the resources column), and a full analysis is beyond the scope of this chapter. However, the range of choice available to planners in the task of organizing for planning, as well as some of the possible determinants of such choices, becomes relatively clear.

We can test the matrix with some examples. First, let us look at a relatively straightforward telecommunications planning problem, such as the establishment of a telephone system. The decision for such a system has already been taken. There is likely to be a relatively high level of certainty about the technical and construction problems involved, as well as on their solutions. The degree of change may be large, as in a situation where a system will be constructed for the first time. Planners and implementors value control over the technical system. And let us further suppose that planners hold an integrated image of society. Clearly, the rational/ comprehensive planning theory, with associated characteristics, would fit well here.

Let us suppose, however, that the planning team is aware of several potential problems with the apparently straightforward task of developing a telephone system. Moreover, let us further assume that there is considerable competition among institutions in society, that new resources will have to be developed, and that planners recognize this. Under these circumstances, they might want to consider an innovative planning approach, emphasizing the development of a new institution—such as a rural telecommunications authority.

Suppose the general goal were widened considerably, to become something like 'improve the quality of rural life through communication'. Planners who saw society as in conflict, with a view of social change operating through individual and group change, might wish to build a planning approach drawing both on transactive and radical planning.

Again, the intent here is not to present an exhaustive discussion of the alternative planning theories, but simply to highlight the primary implication of planning theory for communication planning—that there are alternative planning approaches, and that one or more of these may be more or less appropriate, given the environment of planning.

### THE PROBLEM OF CERTAINTY

It has been a central theme of this chapter that planning, in its most basic sense, is a statement of the theory or explanation: 'if . . . then . . .' The level of confidence that planners have in their understanding of problems, in the data and information they have, in the theories they apply in planning, and the resources (including organizations) that are available has a great

deal to do with how strongly these 'if ... then ...' statements can be made.

Dealing with uncertainty is central to planning theories. It seems clear that the greater the level of uncertainty, and the lower the corresponding level of confidence that planners have, the more important learning becomes as part of the planning process. Planning under uncertainty becomes much more of an experimental process in which feedback and evaluation assume central importance.

Under such circumstances, more incremental planning approaches may be called for. Certainly, in communication planning, where the record of experience is thin and where uncertainty is correspondingly high, formal rational/comprehensive planning may not be the best approach. To take one example, communication planners are often uncertain about what might be reasonably achievable goals. Consequently, there is a need to link goal-setting, planning, implementation and evaluation closely together. Such close linkages increase the probability of learning and of constant adjustment of goals and strategies. The rational/comprehensive separation of ends and means poses barriers to learning of this sort. Similar levels of uncertainty exist with theories in planning, such as that with regard to the cross-cultural use of Western communication theory.

## THE PROBLEM OF VALUES: PLANNING AS MEANS TO AN END OR AN END IN ITSELF

It has been customary in development planning to see planning as a means to important ends. Yet a number of planning theories emphasize the importance of planning itself in human growth and learning. Participatory planning theory grows out of the idea that people who are affected by planned action should have input into the planning process. Transactive and radical planning theories would shift the opportunity of planning closer and closer to the grassroots, in part because of the need for putting planning where problems and information are located, but also because participating in planning has value in itself.

This is not a trivial problem. For communication planners, the implications are immediate and serious. The choice may well be between an efficient planning process, done centrally, to allocate scarce resources to fixed goals, and an extremely decentralized form of planning, in which a diversity of goals is encouraged, where resources are to be mobilized (as through community action groups), and where efficiency becomes a secondary consideration. Planner values and social image will be important elements in this choice.

#### ON TO COMMUNICATION DESIGN

Pity the poor communication planner. The range of choice in theories of planning is large; so too is the range of theories in planning from which planned action can be constructed. Is it not possible for someone to state in a simple and straightforward way how communication planning ought to be done?

The answer is 'not yet'. However, we can begin to identify some ideas that may help planners make their own choices and to develop their own way of planning, one appropriate to their planning environment and challenges. We will call these 'communication design principles'.

# Conclusion: towards principles of design

The reviews of theory in planning and theory of planning, though brief, have mapped a complicated terrain. How can planners move through this complex territory, making effective decisions on how—and what—to plan? There are no easy answers, but a number of practical design principles can be drawn out that may function like signposts along the way, helping planners chart their course.

These principles may form an initial statement of a theory of communication design. They are broad concepts that planners can use in making an initial assessment of the environment of planning—social images, values, problems—to reach decisions on how to proceed, and on the limits of their plans.

These principles are tentative. Moreover, they do not form a complete and integrated map. The terrain of communication planning is too new, and our experience still too thin to offer stronger conclusions. Hopefully, however, these ideas can serve as a beginning point for further thought and, most importantly, testing in the real world of planners.

### THE PRINCIPLE OF CONGRUITY

The review of theories in planning, especially development and sociological theory, pointed out the power of underlying assumptions in the communication planning process. Clearly, the planner's ideas of what development is and how it should be attained, and of how society is structured and how it changes, set important conditions for the planning of communication systems.

The degree to which planned communication systems fit with—or are congruent with—these broad views of development and society is important. The fit between the planning process used (or the planning theory or theories followed) and these underlying images is also significant.

In most circumstances planners will seek to achieve a relatively close fit between images of development and of society on the one hand, and communications planning processes and planned communication systems on the other. Planners with an integrated view of society, and an industrial model of development, will plan one way—possibly following rational/ comprehensive approaches. Planners with an image of society in conflict, with a social-equity view of development, might plan differently, perhaps under a combination of such approaches as transactive and radical planning. It goes without saying that the kinds of communication system that would emerge under these two different sets of images and planning approaches would also differ.

This argument, of course, is itself based on an underlying assumption that congruence—or unity within the system—it a 'good thing'. It is also possible that planners might wish to deliberately violate congruence in order to create change of a certain sort. Rational/comprehensive planning is often imposed as a way of forcing order into a chaotic situation, even when the results may be 'not real' and costly. (Rondinelli, 1976, has pointed out that much planning for development projects supported with foreign assistance falls into this category.) The coin can be turned, of course: radical or incremental planning might be adopted to 'open up' a system that is perceived as 'too' integrated.

The point of the principle—and of the argument—is that the issue of congruity is one that planners should deal with explicitly.

### THE PRINCIPLE OF VALUES

The role of values in planning is obvious in the discussion of congruity above. We have seen in the review of theories that values play central roles—in determining what development ought to be and how it ought to be achieved; in determining the way in which planning should be carried out. Again, planners may benefit by an explicit assessment of their values and the values of key 'stakeholders' in the planned communication system (Middleton and Root, 1979). Knowledge of values will not make decisions easier, but it may make them more effective.

#### THE PRINCIPLE OF CERTAINTY

Planners are concerned with the future, with events that have yet to occur, with patterns of change in society that establish both the opportunities for and constraints on action. Communication planners deal with essential individual, group and social processes—such as learning, attitude and value change, and change in the structures of society. They deal with an increasingly rapidly evolving technology, which opens new opportunities to strengthen communication systems and to create new patterns. The communication process and communication resources are closely related to the political systems of societies, and are thus bound up with change and development in the way in which power and authority are distributed in society. Increasingly, national communication systems affect and are affected by international policy issues, such as the New World Information Order. As we have seen in our earlier review of theories, the basic philosophies and strategies of national development are changing.

These different patterns of change and development create a complex environment for communication planning, an environment that is increasingly uncertain, one in which prediction becomes an increasingly difficult task. Fifty years ago, change over ten years in most countries was relatively easy to predict. Today, a society may be transformed in ten years, great technological advances may be made, and new theories about how communication works may appear.

Uncertainty is a key feature of planning. The degree of uncertainty in any planning situation has significant implications, both for what is planned and for how planning is carried out. Planners must take a careful look at their level of confidence in knowledge of the problem, their data, and their theories in planning.

## Problems and models

Communication tends to operate on social problems through effects on individuals and groups. Thus planners must use data and theories of various kinds to build models of the cause-and-effect relationships that surround complex social problems and that link them with human behaviour. These models then become guides to planned action.<sup>1</sup>

These models, once built, have a tendency to assume a reality that they often do not deserve. The origins of a model in untested theory and poor data can easily be forgotten. The example of the cause-and-effect model through which radio-based teacher education was to affect student learning, given in the first part of this chapter, is a good example.

Thus planners must proceed with caution in their definition of a problem and its causes, and in how communication may be brought to bear on the problem. Ralph Strauch has pointed out that models of this kind can serve planners in two ways (1976). One is when models are used as

<sup>1.</sup> Various kinds of model are discussed later in this book in chapters by Root, Spence and Lyons.

surrogates for the problem they are intended to portray. In this approach, the model is assumed to have a direct, real relationship with the problem. As a consequence, the model tends to be seen as the problem itself, and planners proceed with relatively high levels of confidence in their designs. This use of models is typical of rational/comprehensive planning—such as in the long-range forecasting of communication technology.

The second use of models is as perspectives. Here planners are much less confident in the correspondence between their models and reality. Models of problems are seen as suggesting, but not firmly defining, causeand-effect relationships. The level of certainty is much lower.

The trick is to be clear which kind of model is being used, and to plan accordingly. The greater the certainty, the more formal planning processes and techniques can become. The less certainty, the more experimental planning should be.

#### Data

Planners use both processed and personal knowledge. The former is the kind of data represented by statistics, survey results, and other forms of empirical observation of social processes. The latter is the knowledge that comes from personal experience and insight.

Data of both kinds can be accurate, valid and reliable—or they can be inaccurate, not relevant to the problem or process they seek to represent, and not to be trusted over time. Data are used both to build and operate the models, formal and informal, that planners use. Consequently, planners need to make an assessment of their confidence in data of all kinds. High levels of confidence in data lead to stronger models, and planning with more certainty. Low levels of confidence lead to a much different situation. Bad data used with good models yield misleading results. As they say in the computer world, garbage in, garbage out. Good data and bad models yield the same result.

## Theories in planning

The same considerations of certainty apply to the theories in planning that are used, with data, to create images of problems and solutions. Much of the theory available to communication planners has, for example, been developed largely in Western cultures. How confident can planners be in such theories?

### Certainty and planning

As with the principles of congruity and values, the principle of certainty points towards careful assessment early in the planning process. The degree to which planners are confident in their models, data and theories has direct bearing on the way in which planning is carried out—the degree of formality, the length of time for which plans are developed, the techniques that may be used. Much of the criticism that one hears of planning in developing countries—that it is weak, too informal—ignores the possibility that planners may be acting, perhaps intuitively, with the understanding that their levels of confidence are necessarily low.

#### THE PRINCIPLE OF LEARNING

We began this chapter with the assertion that planning is, essentially, the application of theory. We have come to the end of the chapter with some strong statements about the need for explicit assessment of the level of certainty that planners have in theories, models and data—the stuff out of which plans are made. Along the way we have discussed a number of theories in planning, as well as some contrasting theories of planning.

Planners can rarely be fully confident in their understanding of problems and how to solve them. This is particularly true at the present time for communication planning of most kinds. In conditions of uncertainty, planners should follow planning processes and planned action that emphasize learning from experience. Careful attention to the design of effective evaluation strategies is a necessary component of planning for learning. In addition, planners need to be explicit about the theoretical assumptions in their plans as the basis for evaluation design. They need to assess the quality of their data, and create mechanisms to get new and better information as part of the processes of planning and, importantly, implementation. In an ideal (but uncertain) world, resource allocations—budgets, manpower, time—should reflect this emphasis on learning.

Finally, planners need to create an experimental attitude, one that encourages innovation, rewards those who learn from mistakes, and sees communication planning not as a routine bureaucratic exercise but as an exciting and responsible effort to improve the quality of life in society.

## References

APPLEYARD, D. 1977. A Planner's Guide to Environmental Psychology: Review Essay. Journal of the American Institute of Planners, Vol. 43, No. 2, April.

ARROW, K. 1975. Economic Development: The Present State of the Art. Honolulu, East-West Communication Institute. (Paper No. 14.)

- BARRACLOUGH, G. 1978. Waiting for the New Order. New York Times Review of Books, 26 October.
- BOLAN, R. S. 1975. Mapping the Planning Theory Terrain. In: David R. Godschalk (ed.), *Planning in America: Learning From Turbulence*. Washington, D.C., American Institute of Planners.
- BRYSON, J. M. n. d. Political Models, The Public Interest, and Planning Methods, Strategies and Roles: A Bibliographic Essay. Madison, Department of Urban and Regional Planning, University of Wisconsin. Available from Council of Planning Libraries, P.O. Box 229, Monticello, Ill. 61856. (Mimeo.)
- CHU, G. C. 1977. Radical Change Through Communication in Mao's China. Honolulu, University Press of Hawaii.
- CHURCHMAN, C. W. 1968. The Systems Approach. New York, Delacorte Press.
- FALUDI, A. 1973. Planning Theory. Oxford, Pergamon Press.
- FRIEDMANN, J. 1973. Retracking America. New York, Anchor Press/Doubleday. GALTUNG, J. 1976. Toward New Indicators of Development. Futures, June.
- HAAS, J. E.; DRABEK, T. E. 1973. Complex Organizations: A Sociological Perspective. New York, Macmillan Co.
- HANCOCK, A. 1977. Constraints on the Communication Planning Process. In: S. A. Rahim and John Middleton (eds.), *Perspectives on Communication Policy* and Planning. Honolulu, East-West Communication Institute.
- ——. 1978. Communication Planning for Development: An Operational Framework. Honolulu, East-West Communication Institute. (Paper No. 17.)
- HOLMSTROM, J. 1978. Urban and Regional Planning: Lessons for Communication Planning. In: Rahim et al., op. cit.
- HUDSON, B. 1979. Comparison of Current Planning Theories: Counterparts and Contradictions. *Journal of the American Planning Association*, Vol. 45, No. 4.
- ILLICH, I. 1973. Tools for Conviviality. New York, Harper & Row.
- INAYATULLAH. 1976. Western, Asian, or Global Models of Development—the Effect of the Transference of Models on the Development of Asian Societies. In: Wilbur Schramm and Daniel Lerner (eds.), Communication and Change: The Last Ten Years and the Next. Honolulu, University Press of Hawaii.
- KINCAID, D. L. 1979. The Convergence Model of Communication. Honolulu, East-West Communication Institute. (Paper No. 18.)
- KINCAID, D. L.; SCHRAMM, W. 1975. Fundamental Human Communication. Honolulu, East-West Communication Institute.
- KINLOCH, G. C. 1977. Sociological Theory: Its Development and Major Paradigms. New York, McGraw-Hill.
- LINDBLOM, C. 1968. The Policy-Making Process. New Jersey, Prentice-Hall.
- MIDDLETON, J. 1978. Using Theory in Communication Planning. Development Communication Report, No. 23, July. Washington, D.C., Academy for Educational Development.
- ——. 1979. An Introduction to Communication Planning. In: The Communication Planning Workshop. Honolulu, East-West Communication Institute.
- MIDDLETON, J.; HUDSON, B. 1978. Educational Planning: Lessons for Communication Planners. In: S. A. Rahim et al., *Planning Methods, Models, and Organizations. A Review Study for Communication Policy Making and Planning.* Honolulu, East-West Communication Institute.
- MIDDLETON, J.; LIN, Y. H. 1975. Planning Communication for Family Planning. Honolulu, East-West Communication Institute.
- MIDDLETON, J.; ROOT, G. 1979. Human Systems in Communication Planning. In: The Communication Planning Workshop. Honolulu, East-West Communication Institute.
- NATIONAL TELECOMMUNICATIONS PLANNING BRANCH. 1975. Telecomm 2000. Melbourne, Australian Government Printing Office.

- PORAT, M. U. 1976. *The Information Economy*. Program in Information Technology and Telecommunications. Stanford, Calif., Stanford University.
- PUGWASH SYMPOSIUM. 1977. The Role of Self-Reliance in Alternative Strategies for Development. World Development, Vol. 5, No. 3.
- RITTEL, H. W. J.; WEBER, M. M. 1973. Dilemmas in a General Theory of Planning. Policy Sciences, Vol. 4, No. 2, June.
  RONDINELLI, D. A. 1976. International Requirements for Project Preparation: Aids
- RONDINELLI, D. A. 1976. International Requirements for Project Preparation: Aids or Obstacles to Development Planning. *Journal of the American Institute of Planners*, Vol. 42, No. 3, July.

SCHRAMM, W. 1973. Men, Messages and Media. New York, Harper & Row.

SCHUMACHER, E. F. 1973. Small Is Beautiful. New York, Harper & Row.

- SCOTT, W. G.; MITCHELL, T. R. 1972. Organization Theory: A Structural and Behavioral Analysis. Homewood, Ill., Richard D. Irwin and the Dorsey Press.
- SIMON, H. A. 1969. The Science of the Artificial. Cambridge, Mass., M.I.T. Press. STRAUCH, R. E. 1976. A Critical Look at Quantitative Methodology. Policy Analysis,

Vol. 2, No. 1, Winter.

SUCHMAN, D. 1967. Evaluation Research. New York, Russell Sage.

- UNESCO. 1978. Communication Planning for Afghanistan: An Eight-Year Projection of Communication Development. 3 Vols. Paris, Unesco.
- WILDAVSKY, A. 1973. If Planning is Everything, Maybe it is Nothing. *Policy Sciences*, Vol. 4, No. 2, June.

Π

The systems approach

.

# Chapter 3

# An overview of systems analysis

James Lyons

Systems analysis has evolved over the last several decades in response to the increasing complexity of organizational processes and technologies. Communication planning is typical of the areas in which technological, organizational, cultural, social, political and financial factors come together, making the identification and implementation of any feasible solution very difficult. In such an area, the discipline and structure brought to bear through the application of systems analysis can be of substantial assistance to planners.

This chapter explains the concepts and methods of systems analysis at a level which should permit planners to assess its value in their context. It will not obviate the need to have an expert in the field as a member of the team in order to apply the approach successfully. The chapter moves from the general to the specific through the following sequence of topics: the systems approach, the systems development process, systems analysis, objectives of systems analysis, a systems analysis methodology, and tools and techniques.

The systems approach

The traditional approach employed by scientists to tackle complex problems is to subdivide them into subproblems that are analysed independently. The solution to the original problem is then supposed to be the sum of the independent solutions to the subproblems. This approach has served mankind well and no doubt will continue to do so. However, in recent decades, the severe limitations of this approach have become clear. There are many situations in which the best overall solution to a problem is not the sum of the solutions of a number of subproblems. The overall solution must take into account the interactions between the various subproblems. It is for this reason that a 'systems perspective' is essential for the resolution of problems arising within complex systems, particularly organization systems. What is a system? It is deceptively easy to define the concept of a system and surprisingly difficult to employ the concept in a practical sense. All that can be attempted here is to state the basic concepts, emphasize those aspects particularly relevant to the theme of communications planning and refer the reader to the relevant literature for future study.

A system is comprised of a number of elements (two or more), each of which is in some form of interaction with at least one other element of the system. A subsystem is the system formed by a specific subset of the elements of the original system. No subsystem can be defined that does not interact with at least one other element of the system. The relevant environment of the system consists of those elements external to the system whose interactions with elements of the system influence its functioning.

The boundary between a system and its environment is a function of the problem at hand. The exact choice of elements to include in the system is the result of the reason for defining the particular system in the first place. Another person dealing with the same situation with a different objective might incorporate additional elements of the environment in his definition of the system under study.

For example, suppose a particular health problem is not being solved. One might define the problem as ineffective messages in a health education campaign. The 'system' to be analysed would consist of those persons, resources and activities related to message design. The environment of this 'system' would consist of the other organizational units that interact with the message design system (such as a health research office), the specific audience for the messages, and perhaps the political leaders who influence the content of public information. However, if the problem is defined as an ineffective health education and delivery structure, the 'system' grows to incorporate many additional persons, resources and activities, including some (such as the health research office) that were outside the boundaries of the first, smaller system. The environment for this second, large system (of which 'message design' is a subsystem) expands and changes.

A system that has no interaction whatsoever with its environment is a closed system. The concept of a closed system is a mathematical abstraction. As such systems do not exist in nature or society, this discussion will focus on open systems. An important aspect of any communications system may well be its interactions with its environment.

A system is referred to as an organizational system if it is specifically conceived and structured by people within a societal context. All systems designed and implemented by business and government are therefore organizational systems. All organizational systems are purposeful. They are intended to achieve certain (explicit or implicit) goals of the organization of which they form a part. Why make these self-evident statements? One of the most essential reasons for the need of a systems approach derives directly from these definitions. This is often referred to as systems synergy and is the result of the fact that the relations between the elements are just as important as the elements themselves in determining system behaviour. In general, there is no way correctly to predict or understand system behaviour at a level lower than that of the entire system. In order not to miss potentially important aspects it is necessary to grasp the system in its entirety.

# The systems development process

A systems approach involves viewing the organization as a complex interplay between systems, subsystems and elements (internal and external) whose functioning requires identification and understanding of its structure, objectives and constraints (technological, financial, etc.), as well as the conception and implementation of new systems as they arise.

## SIX PHASES

The systems development process is usually broken down into a number of phases. The precise breakdown and the specific nature of each phase vary with each situation, but there is a generally applicable methodology. A representative breakdown consists of six phases: I, problem identification; II, systems analysis; III, systems design; IV, implementation; V, operations; and VI, maintenance and modification.

Problem identification is the first phase. It generally begins with the realization that a problem exists. There is no saying how this realization arises. It could come from some obvious factors or from an intuitive feeling that a thorough analysis would reveal either improved methods of handling current activities or areas in which new operating systems and methods could effectively be developed. There is no all-purpose structure to Phase I. The catalyst can be a person anywhere within a system. The availability of new or improved technology is a common stimulus. For example:

- A new five-year plan is adopted containing specific objectives to be met for which we have no current system in place.
- A country-wide television system is to be designed and installed—how to go about it?
- How to decide which system of colour television to use in our network?
- The plan calls for 500 telephones to be installed each month, and we seem unable to manage more than 350—what is wrong with our system and how can we improve it?

Efforts to improve crop productivity always seem blocked before they get

started; how can we mount a programme to communicate new methods and technology to the farmers?

Systems analysis (Phase II) begins when an explicit decision has been made that a problem exists and that something must be done about it. Systems analysis is the process of understanding the problem, describing in detail how any existing system or systems deal with it, and developing the specifications of a new system to deal with it. This process is the subject of the rest of the chapter, so further discussion is left for later.

Phase III, systems design, takes the specification of systems analysis as a starting point and works out in detail the functioning of the new system: all logical steps and paths of the system are worked through; job descriptions of personnel are written; instruction manuals for system activities are written; forms and/or reports required or produced by the system are designed; for a computer-based system, the actual computer programs are specified. In short, all details of the design of the new system are resolved. For technologically-based systems (such as computer systems), this is extremely tedious work requiring the consideration of minute details.

Implementation (Phase IV) entails training of personnel, acquiring equipment, printing of forms and procedure manuals, etc. For a computerbased system, there will be programs to write and test. This phase involves all activities necessary to make the system operational. One aspect frequently overlooked or underemphasized is the hiring or training of the necessary personnel. This phase must be managed by a skilled project manager. Large, complex systems require a project manager who is experienced in bringing together the huge number of divergent elements of which systems are inevitably composed. This person must understand the underlying technology of the system sufficiently well to communicate with and supervise the technical specialists. In a typical case, his skills in dealing with the people involved, particularly with the different types of technically-trained people, will be his most important contribution to the success of his mission.

Operation (Phase V) begins when the newly designed and developed system is handed over to its ultimate users and operators. This is the point when, in principle, the organization begins to reap the benefits of its investment in the systematic process described above. This is also the point at which the organization discovers whether the assumptions made during the planning process were correct and complete. The usual result is that not all of the interactions between the system and the organization have been foreseen or described. However, if the job of systems analysis has been professionally done, the positive and negative aspects of these unforeseen consequences balance in such a way that, within acceptable tolerances, system objectives are achieved.

The fit between system and organization is always temporary. Therefore Phase VI, maintenance and modification, is essential over time. All manmade organizational systems have a finite useful lifetime. Therefore, soundly managed systems development provides for explicit monitoring of all operating systems which have been implemented by the process. This monitoring procedure should permit comparison of system performance with original system design objectives as well as with conventional measures of performance. These might be standards used generally by the organization to monitor performance such as return-on-investment, adherence to budget, achievement of plan targets, etc. Perhaps even more important in the monitoring process is a continuing effort to evaluate system performance in the face of inevitable change. There are numerous aspects of the communication planning environment that evolve and that influence the efficiency and relevance of current systems. These include changing organizational goals, technological change, change in international agreements and new or changed planning targets among others. To these concerns must be added the possibility of errors in the original system conception or the realization that another approach would better serve the original system objectives.

Whatever the source of the observed discrepancy between current system performance and that desired, the maintenance and modification phase is the vehicle for making the small changes needed to match the system's performance with the organization's wishes. This type of change must be kept small or the integrity of the whole process is threatened. If a change, or cumulation of changes, exceeds a certain threshold, the rationale of using systems analysis is brought into question. At some point in the evolution of the organization, it is inevitable that the process should be repeated from the beginning in order to bring the ageing system into accord with current objectives and practice. But short of drastic changes, the process should continually assure the correction of errors discovered subsequent to implementation as well as the addition of subsequent minor modifications.

It should be emphasized that this process will never be carried out in the linear fashion described above. In practice, there will be a substantial amount of iteration among the phases:

Feedback from the work of a given phase will trigger re-examination and modification of actions and decisions taken in a previous phase.

Those involved in a system design will find that errors of logic during the analysis phase render their current task impossible.

Implementation will inevitably bring to light points overlooked during all previous phases.

Sometimes the activities of different phases may occur in parallel.

Above all, it will be found that the lines dividing the phases are fuzzy and depend on circumstances.

This list is endless but the result is clear: the entire process is iterative, with the various phases serving more to define the activities to be completed than their sequence.

TOP-DOWN OR BOTTOM-UP?

This process can be viewed either from the top (top-down) or from the bottom (bottom-up). These refer respectively to a view that is global in nature, covering the entire organization and its mission, and to a local view restricted to the implementation and operation of a system for a single activity. As is implied in any such dichotomy of extremes, neither is to be adhered to exclusively. The optimal approach is generally found at some point between. The subject is raised in the current context, because it is an important dimension of the process that must always be recognized and dealt with explicitly. It is of general interest to note that a compromise that has proven increasingly popular is to use a top-down viewpoint to the extent possible in the initial phases of problem identification and systems analysis in order to optimize the organizational fit of the resulting systems. However, the implementation phase is organized from the bottom-up point of view in order better to control the expenses involved in system development before initial results are obtained or any aspects of the system are required to face the rigours of the real world.

#### THE SYSTEMS DEVELOPMENT TEAM

Responsibility for systems development in any given instance should be confided to a team. This team should include several persons representing ultimate users of the resulting system, a number of specialists knowledgeable in techniques appropriate to the particular study, and one or two persons to represent management. The overall responsibility for the study must be in the hands of one of the user members of the group. The users are the only ones who know how a particular activity is carried out today, understand deeply all aspects of the needs of the organization (written and unwritten), and who must live in the future with the system resulting from the study. Without their active participation, the study is doomed to end up in the form of a final report, whose only contribution is to gather dust on a shelf. The specialists bring skills not normally found or even needed among the members of this organizational unit. The specific skills depend upon the subject of the study; however, in any case there must be one or more systems analysts, trained in the techniques described in this chapter. They will often have to be aided by others trained in technological skills (for example, telecommunications or computer hardware and software), economic or financial skills, or behavioural skills. If, as is almost always the case, one result of the study will be to implement changes in people's work routines, someone experienced in assisting organizations to make such changes can be invaluable. The management members of the study team assure the continuing adherence of the project to original objectives and those of the organization, provide continuing feedback to top management, and often provide expertise in such areas as organizational finances and accounting conventions and procedures. The organization of the study around such a multidisciplinary team is one of the essential factors in the success of the systems development process.

#### PLANNING FOR SYSTEMS DEVELOPMENT

Another aspect of systems development that should be mentioned is the planning of the process itself. No organization should embark on such a process without first estimating the cost, time and manpower, etc., requirements of the process. The systems development process itself must be able to withstand cost/benefit tests, as well as any system which it seeks to develop. As with any research and development activity, these cost factors are difficult to forecast. An approach that is useful is to make an initial systems development plan serve as the basis for justifying the application of the process. This plan should include estimates of the various costs associated with each of the development phases, including steady-state operation of the resulting system itself. The insights that such an exercise gives into the components and functioning of the system will permit substantial refining of the original cost estimate. After systems analysis, it will be possible to estimate far more accurately the costs for the system design phase. At the beginning of the implementation phase it should be quite clear what is involved and the associated costs for both that phase and also for operations. Thus the planning of the development process itself should be carried out as a series of refinements to an initial estimate. If the refinements indicate a major departure from the initial estimate, appropriate decisions may be taken.

# Systems analysis

Having situated systems analysis within the context of systems development, it remains to examine the topic directly. What is systems analysis? And of what activities does it consist?

#### DEFINITIONS

One approach to understanding systems analysis is to examine some of the definitions that have been proposed.

Systems analysis is 'a reasoned approach to problems of decision. Some have defined it as quantitative common sense.... It is a systematic attempt to provide decision-makers with a full, accurate and meaningful summary of the information relevant to clearly defined issues and alternatives'. (Rudwick, 1969.)

'... systems analysis has two objectives: (1) to reach a thorough understanding of the underlying details surrounding a situation, and (2) to develop generalized solutions appropriate to the details of the specific situations.' (McFarlan and Nolan, 1975.)

'System analysis consists of collecting, organizing and evaluating facts about a system and the environment in which it operates. The objective of system analysis is to examine all aspects of the system—equipment, personnel, operating conditions, and its internal and external demands—to establish a basis for designing and implementing a better system.' (Couger and Knapp, 1974.)

'Systems analysis utilizes both quantitative methods, where they are applicable, as well as qualitative factors such as judgement, creativeness, heuristics, common sense and experience. Within the framework of a problem solving approach, there are six basic steps to applying systems analysis: (1) definition and formulation of the problem, (2) development of alternative solutions, (3) construction of models which formalize the alternatives, (4) determination of the cost/effectiveness of the alternatives, (5) presentation of recommendations, (6) implementation of the chosen alternative.' (Burch and Strater, 1974.)

### ART AND SCIENCE

There are about as many definitions as there are authors writing about systems analysis. However, five important points can be drawn from this diversity. First of all, systems analysis is more of an art than a science. It is a relatively young concept, having been developed since the Second World War, along with a number of other approaches to rationalizing the decision-

#### Systems analysis

making process. As the field matures, it will certainly become more uniformly understood and more consistently practised. Meanwhile, systems analysis is widely and effectively practised in all types of organization. But what is actually being called systems analysis varies greatly from one instance to the next. The approach employed here is to give a glimpse of the range of activities carried out under the name of systems analysis, and then to enter in detail into one specific approach that is (a) applicable to the field of communication planning, and (b) representative of the various methods in use without necessarily reproducing any of them.

#### BOTH MACRO AND MICRO PERSPECTIVES

The second observation is that some view systems analysis as necessarily being applied at a level in the organization where a global view exists. In other words, systems analysis must be applied to the entire scope of the organization's activities. The other position taken by practitioners is to apply systems analysis to those aspects of an organization's activities particularly in need of improvement or alteration. Clearly, as indicated earlier, the best response is to recognize that there are occasions on which either the macro or micro point of view is valid and of value to the organization.

### THE TIME DIMENSION

A third observation is similar, but adds the dimension of time. There are those who propose the application of systems analysis (under various names, of course) as an analysis of the entire organization over the entire range of its activities with the objective of identifying and rationalizing all operations, structures, systems, etc., of the organization. After an initial, thorough study of the entire organization, they propose that many smaller projects be set up according to some set of priorities, to eventually achieve an 'optimum' functioning of the entire organization. This lofty ambition is easier to propose than to achieve, and these definitions of systems analysis should not be allowed to obscure the value and success of many applications of the approach at the level of one department at one point in time and considering one activity of immediate interest. Many organizations have systems analyst staff who are continually using this approach to deal with a variety of large and small problems at varying levels within the organization.

#### THE COMPUTER

A fourth observation, which is inescapable in the literature of systems analysis, is that there is a major emphasis on the computer. There were (and are) two major causes for the development of the idea of systems analysis. These are the increasing complexity of organizations and the increasing use of the computer. More rational, logical approaches have become necessary to understand these complex systems and their design and control. Use of the computer has required entering into a level of detail probably never before attempted and certainly beyond the patience of man to deal with in the absence of carefully structured methodologies. These two tendencies are linked furthermore in that many of society's recently invented complex systems rely to at least some extent on the computer. This increasing complexity has led to the development of systems analysis techniques. Many of these techniques having been developed by those oriented towards computer-based systems, much of the literature and approaches in use is strongly associated with the design of computer-based systems.

Probably the most frequent use of systems analysis is in the conception and design of computer-based systems. In this area, systems analysis is absolutely essential—the level of detailed understanding necessary to employ a computer is such that less rigorous approaches have a very low probability of yielding useful systems.

There has been an effort of increasing strength over the last ten years to find ways to use the computer itself as a tool in carrying out systems analysis. This effort has yielded some results and will certainly yield many more. However, these efforts are even more strongly linked to the development of computer-based systems than the subject in general. An eventual outcome should be the availability of some computer-based tools to aid the systems analyst in general.

A final point to be made concerning the multitude of definitions of systems analysis is that one can identify common threads running through many of them. It is on these common elements of definition that the soundest understanding of systems analysis can be built. In this way, a working definition of the approach will be proposed for use in this chapter.

# Objectives of systems analysis

The fundamental objective of systems analysis is to provide the foundation for the development of a system. This system may be the modification of one that exists already or something entirely new. The analysis may be applied to the activities of the organization in a global sense, or it may be
confined to one small project involving only a few people in a small part of the total organization. But in all cases, it determines the ultimate form of the system to be implemented.

Two activities that are a part of systems analysis are understanding and conception: understanding of the existing system, the environment of the proposed system, the objectives of the new system and those of the organization, and the constraints relevant to the particular situation; conception of possible ways to satisfy the objectives within the constraints identified, possible technology bases for the new system, other ways of doing things than those used in the past, and enough of the details of a new system that the design phase can be started.

In a general sense, systems analysis can be viewed as considering three essential elements:

What are the desired outputs or results of the new system?

What are the inputs or sources necessary to produce those outputs?

What is the logic of the process that will convert the inputs into the outputs? An important result of using systems analysis should be to generate thorough documentation throughout the process. It should always be possible to go back and find the justification for each of the decisions taken along the way.

It should be re-emphasized that there is no one correct approach to systems analysis. There are some fundamental aspects that must be adhered to, but the specific methodology should be at the discretion of the organization concerned. However, for the purpose of exposition, the following section takes a specific approach to systems analysis.

# A systems analysis methodology

The process of applying systems analysis to a specific case can reasonably be broken down into six kinds of activity. These might be called initiation, documentation, interrogation, innovation, selection and specification. The purpose of this breakdown is to provide a framework for defining the process and for better understanding it. Each of these steps will now be considered at length.

It should be stressed at the outset that each of the terms above implies a type of activity carried out by the analyst team. The order of the list gives a rough idea of the sequence in which the activities are carried out, though they will never occur in a strictly sequential fashion from top to bottom. In general there will be a progression through these activities leading to a gradual development of the result that is sought.

An important concept underlying the whole methodology is that an interdisciplinary team will have responsibility for carrying out the process.

This team will normally consist of representatives of users of the new systems, together with a number of systems analysts. Each activity in the list represents a very different type of work for the members of the team: this is, in fact, the main aspect of the analysis of the process that follows.

### ACTIVITY I: INITIATION

Before beginning the discussion of the activities that make up the actual systems analysis process, it is important to note that the first essential step consists in launching the systems analysis project. The degree to which it is done precisely and thoughtfully is one of the most important factors in determining the ultimate success or failure of the project.

The initiation step consists of establishing unambiguously the terms of reference for the systems analysis study. Ideally, there should be a written contract between the systems analysis team and the requesting organization. This should be done whether or not the exchange of money is involved or external persons or organizations are concerned. A written contract, even between two internal groups, has the strong advantage of being an explicit and permanent statement of what the two parties have agreed to do. See Figure 1 for an example of the sort of document that is useful in this context. It must contain a clear, concise statement of what the systems analysis team is going to do. It must state who is to do it-a list of the members of the team. It must specify the objectives which the resulting system is to achieve as well as the constraints under which the study is to be carried out. These constraints will normally include a time horizon for the conclusion of the project and a budget for the study. It may well specify exact dates for the start, finish and perhaps intermediate points of the project. The greater the clarity of this document and of the resulting relationship between the team and requesting organization, the greater the chances of achieving the best solution to the problem that led to the study.

# ACTIVITY 2: DOCUMENTATION

The next activity to be undertaken by the team will usually be to study the existing system. (This does not necessarily imply that the objective of the systems analysis is always to replace an existing system.) The aim of this part of the project is to understand how things are done at present. There may be a system or systems to be replaced. There may simply be similar activities currently being done. There may be other similar organizations doing similar things that could be a source of experience from which the team could benefit.

PROPOSAL TO INITIATE SYSTEMS DEVELOPMENT PROJECT
DATE: <i>1 January 1976</i>
PROJECT TITLE: Establish external telephone link
RESPONSIBLE USER DEPARTMENT: Ministry of Communications
USER LIAISON: Section chief, Telecommunications
CHIEF ANALYST: Jones
PROJECT OBJECTIVE: Determine the most cost/effective type of international telephone link to meet the country's needs for the next 25 years.
REASON FOR PROPOSING SYSTEMS ANALYSIS: Current links are inadequate. No reliable way for private or business users to link into the international telephone system. Impeding development of country.
PROJECT SCOPE: Consider cable, microwave and satellite. Must handle up to 200 calls simultaneously.
SCHEDULE OF MAJOR ACTIVITIES: Initiate study: 1 March 1976 Complete analysis: 1 June 1976 Begin implementation: 1 October 1976 Begin operation: 1 October 1977

FIG. 1. Project initiation document.

In fact there will always be something to study. In every case there will be the environment of the current organization, into which the new system must be inserted. This environment must be thoroughly understood by the study team. It is never reasonable to propose to change an organization without being familiar with its history. The context of any proposed system will be the source of interactions with the different elements of the system. These interactions must be taken into account in order to assess the impact of the new system.

Moreover, there are always constraints operating in the organization, limits within which any proposed system must be capable of functioning. These often concern: money (the budget within which the system must operate); time (organizations generally live under the influence of certain time frames and cycles that any system must respect); numbers and categories of staff; administrative or legislative policies; technical constraints (the range of broadcast frequencies that exist is limited as is the amount of information that can be carried at a given frequency; equipment available has limitations of power, range, reliability).

The objectives of the organization as they relate to the project must be clearly understood. These will often impose limits on the types of system that might be proposed. These may be found in general documents or plans or long-term strategies of the organization, or they may be precisely laid out in the project initiation document. The scope of the project will likely result from organizational objectives and strategies.

If there is an existing system to be replaced by a new one, then documentation of the old one is fundamental. It must be understood in complete detail. The team members must immerse themselves in the current system at this stage as a prerequisite to their later activities—any proposed system must withstand a harsh examination in the context of the functioning of the existing system. This activity can be summed up in terms of searching for the precise and exhaustive answers to the questions: What is being done? Who is doing it? When is it done? Where is it done? How is it done?

Those team members who have no previous experience with past approaches must use this phase to familiarize themselves thoroughly. This process and the resulting written documentation provide the foundation for all subsequent activities.

Techniques: Various techniques are employed to record and represent the data gathered in the documentation activity. In general, these can be categorized as textual, tabular or graphical. Each type has advantages and disadvantages, and all are important in one case or another. It is frequently helpful to develop visual presentations of the data on the existing system. It has often been demonstrated that many people find graphic presentations easier to comprehend. The study team can perform a valuable service for the organization by writing out details on the operation of the system and by finding ways of presenting these data in graphic forms. This opportunity does not often occur in the life of most organizations, and the written documentation rarely exists in the absence of systems analysis activities.

Data-collection techniques are also important in the documentation activity. These include interviewing, observation, and questionnaire surveys. An essential element of the process is to interview those involved in the current system and those in its immediate environment within the organization. Much of the basic data on current operations is gathered in these interviews. As data are collected, and the team members develop

78

an understanding of how the organization works, these must be tested and verified through further interviews with those concerned.

As with the production of written documentation, the conducting of interviews requires certain skills. These are primarily behavioural and organizational and will be discussed later.

During the documentation activity, the systems analysis project team members function as investigative reporters. By the end of this activity, they should have become experts in how the current organization operates.

## ACTIVITY 3: INTERROGATION

The object of the interrogation activity is to examine the results of documentation (Activity 2). Having obtained and documented the answers to the questions concerning what, who, when, where and how, it is now time to ask the question: Why is it being done like that? Every aspect of the existing approach and how it is carried out should be explored:

What are the strengths and weaknesses of the current system?

What are the implications of carrying out the operations as they currently are?

What are the opinions of current systems users as to why the system is designed and used as it is?

What has been the evolution of the reasons for the current design of the system?

Techniques: During this activity the team will work mostly from the files of documents it has created during Activity 2. The bulk of the step is a critical examination of the contents of those files. Some further interviews may be desirable to clarify vague points or seek additional opinions. It is often necessary to go back to the activity of documentation to fill in gaps discovered during the interrogation phase. A typical instance of systems analysis involves frequent passage back and forth between the various activities as the process moves along. During interrogation team members function somewhat as detectives. By the end of this activity they should have become experts on the strengths and weaknesses of the existing system.

# ACTIVITY 4: INNOVATION

The fourth activity of the project team is to search for alternative ways in which operations could be carried out. The files of information developed and recorded in Activities 2 and 3 serve as the background for this phase. Much of this information is now firmly grasped by the team members.

Now they need skills of creativity and imagination. The relevant question now is: how could it be done?

The objective of the innovation phase must be to imagine all possible approaches: to search a large number of alternatives. In order to have maximum freedom for unrestricted thought, there should be minimal concern for the constraints. The environment or context must be treated as secondary. All efforts must be concentrated on the possible means of achieving the required objectives. The activity can and should be carried out both individually and as a team. Any available source of ideas should be explored. Ideas of those interviewed during Activities 2 and 3 as to alternative approaches to the system under study should certainly be introduced here. Effort is concentrated on idea generation, not yet on critical examination of the ideas generated.

*Techniques*: Many techniques have been developed and used in recent years to expand and structure individual and group creativity. Any approach that augments creativity or encourages imagination is valued in the innovation stage. Further discussion of this subject will be left for later.

During the innovation activity the team members function as innovators, as generators of ideas. By the end of the step, they should have expanded their horizons and become aware of the full range of possible alternatives.

#### ACTIVITY 5: SELECTION

The purpose of the fifth activity is to select, from among the alternatives generated during the innovation stage, the solution to be developed as the new system. The question to be addressed is: Which alternative is the best one to be implemented? This step can be compared to throwing a bucket of cold water on the uncontrolled enthusiasm that reigned during the previous step. The team must go back and take into account the hard facts of the reality of the environment and the constraints recognized in Activity 2. A detailed, rigorous analysis of the alternatives must be conducted in the light of the files of data that the team has developed previously.

The first step will of course be to discard any alternatives that are patently impossible. That done, the remaining candidates for new approaches must be examined. In the event that a substantial number of potentially viable alternatives remain, it is necessary to utilize an approach of successive elimination: coarse filters screening out the obviously inappropriate alternatives, finer meshes being used to make discriminating judgements later. The analysis must become progressively more precise

and exacting as the number of remaining alternatives is reduced. When the two or three most promising potential solutions have been identified, the team must carry out a careful feasibility study of each in order to determine which is to be retained. This must include an assessment of proposed system performance compared with required objectives, compatibility with applicable constraints, expected benefits (monetary and otherwise) compared with predicted costs, and so forth. The team will inevitably be forced to make trade-offs between mutually exclusive characteristics evidenced by competing solutions. For example, solutions often differ as to their relative efficiency in purely technological terms and at the same time they vary in attractiveness, in a human sense, in the eyes of the people affected by the system. When these two dimensions vary inversely among the alternatives under consideration, a compromise has to be made to sacrifice one in favour of some improvement in the other. This type of choice, which is always present in a technology-laden field such as communication planning, requires that team members be sceptical in their reliance upon technology and cautious but firm vis-à-vis the importance of human values.

An example of this might be seen in a community training programme in an area such as family planning. At one extreme one could design a technology-intensive programme based on audio-visual materials, or perhaps community television requiring no trained teaching personnel on site. The other extreme is represented by a classic programme of instruction with teachers going from one site to the next disseminating the information to groups of people in a classroom format. Chances are that neither of these extremes is optimal. One has the advantages of being capital-intensive, constant in message and delivery, and attractive to those who favour new machines and gadgets. The other is more flexible, more alive, more variable in content, and highly labour-intensive with all the associated advantages and problems. The choice facing the decision-maker in this case is difficult, involving trade-offs between options that are inherently not comparable.

*Techniques*: The precise techniques used by team members during Activity 5 are the typical tools of management in evaluating any investment proposal. They must use the same methods always used in cost/benefit and feasibility studies. Recently developed techniques for making choices when there are a number of criteria and those criteria are not quantifiable are also quite helpful. Some specific discussion of these tools is given later.

During the selection activity, the project team functions as a judge. By the end of this step, the solution is taking shape: the team has identified the approach to be used for the new system. All that remains is to flesh it out, to fill in the vital details.

## ACTIVITY 6: SPECIFICATION

The specification activity involves designing the new or modified system that was selected in the previous activity. This phase may appear more structured and more clearly defined than some of the others, but that does not make it any easier or less demanding. Successful implementation of the new system depends as much on the careful execution of this activity as it does on that of the others.

The specification activity answers the questions, with respect to the system: What is to be done? How is it to be done? Who is to be it? Where is it to be done? When is it to be done? All new or modified outputs of the system are defined. All required input data, forms and procedures are worked out. All system flows, procedures and logic necessary to produce system outputs are determined and described precisely. Any new or modified organizational structures, task definitions, equipment, etc., are explained. The project team must produce documents that clearly explain the new system to all concerned.

This process requires meticulous follow-through of the implications of all preceding activities. It must be assured that the solution satisfies the objectives or that a satisfactory modification of original objectives has been agreed upon. All relevant constraints must have been respected, or a way of overcoming them devised. Any changes in the context of the system necessary for its implementation must have been accepted. Any changes which might follow its implementation must be identified and preparation made. In other words, the team must understand the proposed system in Activities 2 and 3, documentation and interrogation.

Techniques: Many different techniques might be used by a system analysis team in a typical specification activity, depending on the type of system which they propose to define and the context into which they propose to install it. Tools that facilitate the representation of the logic of a system are often used. These include decision tables and various types of flowchart. All of the same documentation techniques used in Activity 2 are equally applicable in describing the proposed system. Many forms of tables and charts are useful to describe forms, structures and equipment needed.

A final class of techniques, which has been found valuable in working out the detailed operation of a new system, is the walkthrough or talkthrough. This technique is particularly useful to find bottlenecks or points overlooked in relatively complex systems involving interactions between many people and a large number of operations. Some of these tools are discussed further below. During Activity 6, the team members function as designers or architects. By the end of it they have specified a complete system in accordance with the terms of reference of Activity 1, initiation, perhaps subject to subsequent modifications.

#### THE FINAL ACTIVITY

The tangible product of the procedure described above should be a report that provides all the information necessary to build and implement the new system. The value of this report, and indeed that of the entire project, now rests in its clarity, its completeness, and its fidelity to the original terms of reference and to the final system conception as it exists in the minds of the members of the project team. With the presentation of this report to the requesting body, supported by all the documentation developed by the team throughout its study, the systems analysis team has completed the project.

# Tools and techniques

Systems analysis has proven far more efficient and successful when special tools are utilized. Many such tools exist: there are those that are simply the organized, systematic application of common sense; some are more sophisticated and require more substantial training before their use. Examples of the latter are the network models like PERT and CPM or detailed flowcharting techniques. Many of these techniques were developed around the computer.

The analyst should have at his disposal a wide variety of these tools from which he can select those that are of value in a given situation and utilize them to advance the project. A full analysis of all relevant tools is beyond the scope of this chapter. An overview of some representative examples will serve to illustrate the various roles played by these techniques and introduce some of the basic ideas to the decision-maker. With this objective in mind, examples of tools and techniques useful in various of the six areas of activity encountered in systems analysis are presented below.

#### DATA PRESENTATION

There are three broad categories of data presentation approach: textual, tabular and graphical. The choice of approach depends entirely upon the situation at hand and the objectives sought through the presentation. It is important to consider the impact of various alternatives when preparing any report. The amount of information conveyed, the clarity with which it is conveyed, the degree to which the message is received with a positive or negative connotation, can often be influenced by the way in which the data are presented. A graphical approach often proves to be far more efficient and impressive but requires a somewhat greater investment in preparation.

# Textual presentation

The textual presentation is, of course, the most common and perhaps the easiest approach. It is more common simply to write in textual form the information which one has to convey. However, in a technical area like communication planning, straight text is rarely the optimal way to present information. Text will probably always be of value as the explanation of various tables and graphs as well as to explain certain concepts and ideas. The type of text which is generally needed in the reports generated in systems analysis should be in a technical or report-writing style; that is, it must be clear, direct, with a minimum of complex language structures. Simple, declarative sentences should predominate. It is a good rule that when there is a choice between graph and tables or text, the former are to be preferred.

# Tabular presentation

Tabular presentation is the listing of data, usually in one or two dimensions. The one-dimensional case is usually called a list and is commonly used. It is worth emphasizing that even in this simplest of cases, a list separated out from the text has greater impact than the same data embedded within it. The success of the tabular presentation of data depends upon the clarity and simplicity of the organization, headings, physical layout as well as such details as, for example, the choice of the number of digits to use in listing numbers. Attention to such details has a large payoff in terms of rapid understanding of the resulting report.

Two examples will serve to illustrate the use of tables in presenting data. The first of these, Table I, contains a summary of data on the development of the use of cable television. Much of the content of this table is of a descriptive nature, but the presentation in table format allows quick comprehension and easy comparison of facts. The density of information conveyed in a document can be somewhat greater if that information is presented in such a format, since the presentation itself contains a certain amount of the information and facilitates understanding it. Table 2 is a simple tabular presentation of a series of numbers. Here it is obvious that the information could not be presented in the text. In fact, the presentation chosen is ideal in that it is oriented towards a specific type of use known as 'table look-up'. The two dimensions represent the two characteristics of

.

Country	First cable system started	No. of homes connected (1972)	% of homes	Notes
United Kingdom	1951	2.2 M	13	Growth rate 12% p.a.
United States	1950	6.5 M	II	CATV systems may not be owned by TV stations 20% of cable stations originate local programmes (1973)
Canada	1951	1.4 M	23	114 of 361 licensed cable systems are originating local programmes
Ireland	1960 for Irish programmes only 1973 for British and foreign material	65 000	8	
France	1972 experimental services in a few towns			
Italy	Independent cable companies ended officially 1973			Experiments in local broadcasting and community television, including some two way, continue
Netherlands		1.8 M	40	6 stations provide local programming
Belgium		450 000	14	II programme channels include several from neighbouring countries
Japan	Two large-scale exp New Town has 2,00	eriments cu o subscriber	rrently rs.	under way-one at Tama

# TABLE I. Summary of domestic cable developments

TABLE 2. Calculation of income tax payment

	Number of dependants							
income	I	2	3	4	5			
Under \$3 000	348	129	o	o	o			
\$3 000-\$3 999	463	326	204	68	0			
\$4 000-\$4 999	625	481	378	264	148			
\$5 000-\$5 999	811	636	528	422	306			
\$6 000-\$6 999	982	815	688	575	467			
\$7 000-\$7 999	1 257	978	859	740	621			
\$8 000-\$8 999	1 482	I I49	1 030	911	792			
\$9 000-\$9 999	1 729	1 320	1 201	1 032	963			

interest—taxable income and number of dependants—and the value given on the basis of these two parameters is the corresponding tax payment. The two characteristics define a row and a column of the table and the appropriate value is found at the intersection of this row and column.

#### Graphical presentation

The graphical approach is superior to either tables or text. Almost without exception, a graphical means is preferable when it can be employed without substantial increase in cost or effort. The quantity of information and the speed with which it is communicated are greater than with text. Selected examples of forms of presentation frequently used by systems analysts are described in the following paragraphs.

Everyone is familiar with bar charts, pie charts and line graphs. Though we see them every day in all types of context, this familiarity should not be overlooked as an advantage in their use in systems analysis reports. Figure 2 is a typical example of a bar chart.



FIG. 2. Bar-graph comparison of systems.

#### Systems analysis



FIG. 4. Total of 1,701,522 deaths in the United States (1960), distributed by age of the deceased.



In this particular case, the chart illustrates the evolution over the coming five years of the costs of three possible systems; the current one and two alternative proposals. It is clear that the current system is the least expensive during the current year, but becomes progressively more expensive, in a relative sense, with the passage of time.

Pie charts (see Fig. 3) are useful in displaying the distribution of some quantity, such as a budget. Line graphs efficiently display the relative frequency of distribution of one factor (variable) along a particular dimension (see Figs. 4 and 5).

Other commonly used graphic display techniques are organization charts in various levels of detail, charts showing the physical layout of facilities, and flow diagrams showing the flow of material or personnel in a work space.

An important kind of data display shows how work and responsibility in a given project are scheduled over time. An old but frequently used approach to this is the Gantt chart. A Gantt chart has three primary purposes:

First, it tells the manager what jobs are assigned and who has been assigned to them.

Second, it indicates the estimated dates on which jobs are assumed to start and end, and it represents graphically the estimated duration of the job.

Third, it indicates the actual dates on which jobs were started and completed and pictures this information.

An example of a Gantt chart is given in Figure 6. Across the top of a Gantt chart a time scale is represented, broken down into units relevant to the activities under consideration. Down the left side of the chart are listed the tasks to be completed, together with the person or department responsible for their completion. Within the chart, each task is represented by a bar whose length represents the duration of the task and whose position along the time axis indicates at which point in the period of the total project this task is programmed to take place. Various means are used to indicate progress against the initial schedule. In Figure 6, the bars labelled 'S' give the task lengths and timing according to the estimated schedule, and the bars labelled 'A' indicate actual progress. At any time in the realization of the project, the status of each task, and thus of the entire project, can be seen from the current state of the Gantt chart. In the example given, the project is currently in the week of 14 July, and Job 2 is way ahead of schedule. Job 1, however, seems to have fallen a few days behind. Gantt charts are useful analysis and management tools. In the case of a systems analysis project, they portray clearly the timing and roles involved in the achievement of certain project tasks. What they do not show at all is the dependencies between the various tasks: other techniques are required for this.

There is a group of techniques that present information in a grid format. The basic grid chart lists all departments concerned with a specified activity. Then, in the columns defined by these headings, the various tasks are described beneath the department responsible. These tasks may be numbered consecutively in the order in which they must be carried out and arrows may be used to emphasize this sequence. Figure 7 illustrates this approach.

Many other examples of grid presentations can be cited. Table 3 is a



FIG. 6. Gantt chart.

case illustrating the relationships between positions, functions and types of responsibility.

There are many ways to chart the flow of tasks necessary to achieve a specified activity; another approach is to employ easily recognized symbols to identify functions. This has the advantage of making a flow chart more concrete and more readily understood to those engaged in the processes under review (it is often utilized in commerce, to illustrate purchasing or marketing procedures), since it makes use of the tools of the trade.

While it may appear that charting is a relatively simple task, the novice analyst will find that early attempts at charting can be frustrating, timeconsuming and complicated. Even with experienced analysts, first attempts

#### Name of system: Fulfilling customer's orders Date: 20 Dec. 1975 Warehouse Shipping Billing Accounts Broker department department receivable Warehouseman Packer (1)Takes order from customer on form 87 ORDER (2) ··· (3) Sends order Pulls to warehouse materials (4) (5) Gives to Prepares materials packer with order for shipment (6) Stencils address on boxes (7) • (8) Sends to Loads shipping on truck (9) (10) Notifies Prepares billing by form 42 sending 'filled' INVOICE, from form order 87 ORDER (11)Mails original and duplicate to customer (12) .... + (13) Sends parts 3 and 4 to Sends (13)A commission copy to Broker Receives Accounts commission receivable copy of invoice (14) Upon receipt of invoice sets up ageing file

#### Approaches to communication planning

FIG. 7. Grid chart.

	New products	New markets	Training	Salary and commission
President	Α			Α
Vice-president, marketing	A	Α		Α
General manager, sales	С	Α	Α	AP
Regional manager, sales	С	Р	AP	С
Manager, product planning	PC	Р		
Vice-president, personnel				Α
Manager, supply and				
transportation	0	0		
Manager, product engineering	APC			
Manager, manufacturing	0			
Salesmen	Ō	0	РО	
Manager, research and				
development	Р			
Controller				PC
Manager, information				
systems	0	0	0	PO

TABLE 3. An illustration of a linear organization chart of some of the marketing functions. A: approval; C: control; P: planning; O: operation.

to chart an organization or activity are considered drafts and are quickly redrawn. The usefulness of charts, however, makes them a necessary tool in the analyst's repertoire.

The following set of guidelines may assist the analyst to gain maximum usefulness from charting attempts:

The first attempt at constructing a chart should be free hand.

If a chart is not readily understandable to the viewer, the chart should either be simplified or redrawn as two or more charts.

- The symbols used in a chart should aid the viewer and not hamper his understanding.
- Charts that are to be used as permanent documentation should be keyed to brief narratives as appropriate.

Charts used in presentations to non-technicians should minimize symbolism.

Charts used in visual presentations should contain text large enough so that all members of the audience can read it.

#### LOGIC REPRESENTATION

Closely related to the techniques for data presentation are the systems analysis tools for the representation of the logic involved in the processes and procedures of the system under study. Two such tools will be presented here, each of them in the form of a summary description.

### Approaches to communication planning



FIG. 8. Major sections of decision table.

# Decision tables

A decision table is a tabular representation of (a) conditions: factors to consider in making a decision; (b) actions: steps to be taken when a certain combination of conditions exists; and (c) rules: specific combinations of conditions and the actions to be taken under these conditions. These items are contained within the body of a decision table, which is divided into four major sections by double or heavy vertical and horizontal lines.

The various parts normally making up a table are.

- I. Condition statements: statements that introduce one or more conditions: 'If . . . '
- 2. Condition entries: entries that complete the condition statements.

		R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>
Condition 1	Credit limit OK	Y	N	Ν
Condition 2	Pay experience favourable	_	Υ	N
Action 1	Approve credit	х	х	_
Action 2	Return order to sales		—	x

Credit order approval procedure

Note: Be certain that the latest credit information is used.

X = Action taken

FIG. 9. Decision table example.

- 3. Action statements: statements which introduce one or more actions: 'Then . . .'
- 4. Action entries: entries that complete the action statements.

As an example of a decision table, consider the following situations:

Rule One  $(R_1)$ : If a customer's credit is okay, then the credit request may be approved. Rule Two  $(R_2)$ : If credit is not okay but the customer's pay experience is favourable, then the credit request may also be approved. Rule Three  $(R_3)$ : If the customer's credit limit and pay experience are unfavourable, credit should not be approved and the order should be returned to the sales department.

Figure 9 is a decision table embodying this logic. It is easy to see immediately from the table which action should be taken under which conditions. The three possible situations are arrayed clearly by this approach.

A second example of a decision table, given in Figure 11, illustrates a somewhat more complex decision situation. In this case there are three conditions used to decide which of eight rules to apply. These rules lead either to placing the student on the dean's list, or to applying one of three subsequent decision tables.

Here are some examples of the use of the decision table in Figure 10. What would be the decision on a full-time student, with a 3.0 grade average, who has not received any 'F' grade? We find out by finding the rule under which there is a 'Y' (yes) for each of these three condition statements.

	Dean's List decision table									
-	Condition/action statement	Rules								/
_		R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	R <sub>6</sub>	R <sub>7</sub>	R <sub>8</sub>	Ĩ
	If the student is a full-time student (12 credit hours)	N	N	N	Y	N	Y	Y	Y	
-	If the student has received a grade point average 3.0	N	N	Y	N	Y	N	Y	Y	
	If the student has not received an 'F' grade	N	Y	Y	N	N	Y	N	Y	
_	Then prepare a Dean's List statement	•	•	•	•	•	•	•	×	
	Then GO TO probation/ dismissal table	•	•	•	×	•	•	•	•	
	Then GO TO part-time student table	x	x	x	•	×	•	•	•	
	Then GO TO regular student table		•	•	•	•	×	×	•	
		X = /	Action	taker	ı					
	Probation/ dismissal decision table									
· · · · · · · · · · · · · · · · · · ·	Part-time student decision table									
	Regular student decision table									

FIG. 10. Multiple decision tables using a GO TO statement.

, , , , This is Rule 8 (R8). We then read down the column of Rule 8 until we find an 'X'. We then read left across the row to find the appropriate action. In this case the action is to 'prepare a dean's list statement'. How about a student who is not enrolled full-time, but has a 3.0 grade average and has not had any 'F' grades? The appropriate rule is  $R_{3}$ ; the action leads the decision-maker to the part-time student decision table.

When an analyst identifies an opportunity for constructing a decision table, the following approach is suggested:

- 1. Limit the decision process or objective of the decision table with firm boundaries.
- 2. List all the conditions that must be addressed before a decision can be executed.
- 3. List all the activities that must be accomplished, based on the exact nature of the decision.
- 4. Identify and define the values of all conditions and actions.
- 5. Classify and consolidate like conditions and actions.
- 6. A decision table can only have one entry point.
- 7. A decision table can have many exits.
- 8. Each rule in the final table should have at least one condition entry different from any other rule.
- 9. Test each rule in the table as well as the table as a whole for completeness, accuracy and proper format.

Some advantages of decision tables are that tables:

- 1. Display situations concisely.
- 2. Promote completeness and accuracy.
- 3. Can have completeness and accuracy techniques applied to them.
- 4. Permit a complex situation to be more easily grasped.
- 5. Have allowed the study and design of systems previously too complex to be handled by other methods.
- 6. Facilitate modularity.
- 7. Are relatively easy to construct, modify, and read.
- 8. Are useful in computer applications.
- 9. Improve client/analyst communication.

As a systems tool, decision tables work best in those instances that are both specific and detailed. For example, using a decision table that clearly outlines the credit action to be taken under specific conditions, a clerical worker with little or no training can accurately process or refer numerous credit applications.

Decision tables are also a beneficial addition to systems documentation. By including decision tables in the formal documentation, readers can quickly review some of the complex decision-making processes without wading through lengthy, written statements and procedures. More detailed discussion of decision tables is found in the references to the chapter (IBM, 1962; McDaniel, 1968, 1970*a* and *b*; Pollack et al., 1971).

#### Flowcharts

Flowcharts are another way of visually depicting logical relationships. They have had their greatest development in connection with computers, but are widely used in broader contexts. Various conventions exist as to the use of specific symbols to represent different types of operation, but such use is generally restricted to a specific field of activity. The important benefit of flowcharts, however, does not result from adherence to a specific convention as to format. The main objective is to represent all aspects of the system and their precedence relationships, the occurrence of decisions, and the actions that correspond to various alternatives. In many cases, a flowchart describes the entire process of transforming all system inputs into system outputs.

Flowcharting is widely used as a means of better understanding existing or proposed systems. It is a diagrammatic representation or map of the flow of events.

Flowcharts are constructed to conform to our natural tendency to read from left to right and top to bottom, though variations are sometimes desirable in order to achieve symmetry and to emphasize certain points. Within this framework, flow lines can be drawn horizontally, vertically and diagonally. The primary considerations are neatness, uniformity and clarity. Explanatory footnotes are encouraged and are frequently used.

Some symbols are used almost universally in identifying particular events in the flow of the system. The flow of data itself is usually indicated by a line and the direction of the flow by an arrowhead. In flowcharting,



#### Systems analysis

#### Basic symbols



FIG. 12. Basic flowchart symbols.

solid lines are normally used to indicate the direction of flow, but occasionally it is desirable to differentiate between the physical movement of work and the transfer of information. This can be done by using solid lines in the first case and dotted lines in the second. Also, for one symbol to have a meaningful relationship to other symbols in a chart, it is necessary that it be connected to one or more other symbols in such a way as to indicate the sequence in which the operations occurs.

A standard set of flowchart symbols is given in Figure 12. These were established specifically for use in computer-based systems development. However, today they are often found in applications far removed from any computer implication.



FIG. 13. System flowchart.

The first step in flowcharting is to define the problem itself. In doing this, answers must be found to the following questions:

- 1. What information is required from the system in the form of reports?
- 2. What is required from this portion of the system to integrate it into the overall system of which it will be a part?
- 3. Where does the input information arise?
- 4. What exceptions can arise and under what circumstances?
- 5. How many transactions and how many exceptions of each type are involved?
- 6. How should each transaction and each exception be handled?
- 7. What files are involved and what information is involved in them?
- 8. Are results required on schedule or on demand?

Answers to these questions will supply much of the necessary information for the preparation of the initial flowcharts of the system.

An example of a simple system flowchart is illustrated in the pictorial presentation of the major considerations in processing a payroll in Figure 14. This figure also illustrates the use of the four basic symbols. The general symbol for any type of input or output is used in this flowchart. More specialized symbols may of course be employed if desired to indicate more precisely in a computer context the type of input or output to be used.

The process symbol is used to represent each of the series of calculations involved in calculating the payroll. First is the determination of the gross pay amount, then the various payroll deductions are developed, and last,



Input from the organization

Input from the telephone company

Users

FIG. 14. Telephone information system.

Users

the calculation is made of the net pay amount for which the cheque is to be written. Each symbol in the flowchart provides only general information and even the notations are in summary form, but the flowchart does provide an overall view of the procedure necessary to prepare the cheques for a payroll.

In the figure, the arrow leading from the final output symbol back up to the initial input symbol represents a looping procedure. Without this loop notation, the chart would indicate that only one cheque is to be processed. The loop indicates that a number of cases are to be processed in sequence.

A more complex flowchart illustrating a telephone information system is presented in Figure 14. In this system, all calls are logged by the switchboard operator. These data are then recorded on punched cards and entered into the phone charges. Another set of inputs comes from the telephone company in the form of detailed bills. All inputs go to a computerbased system which provides a series of reports for various users. These include departmental usage data for department heads, total institutional cost data for the accounts department, overall statistics on usage for those responsible for the internal telephone system, and directory information for the switchboard operators.

Another example of this type of presentation comes from an information system developed by the BBC. Figure 15 shows one aspect of that system, the 'programme file'. This file contains data on some 6,000 programmes produced by the BBC during one year; other files contain data on resources available and scheduling and cost data. Together, these files permit the generation of many reports for management. In particular, detailed cost data and comparisons with budget are provided as well as reports to aid in scheduling the various programme material for broadcast.

Whatever the context and whatever sets of symbols are used, flowcharts enable better communication and more explicit statements of system objectives and functioning than would otherwise be possible. The principal concern of the analyst should be to prepare complete and clear flowcharts in order to portray with accuracy and clarity the logic of these aspects of the system covered by the chart. Further information on flowcharts may be found in chapter references (Burch and Hod, 1975; Burch and Strater, 1974; Chapin, 1971; Couger and Knapp, 1974).

#### INTERVIEWING

Interviews are critical in the data collection process; they are also critical in that the essential relationships of the system analysis team with the organization being studied are established by them. An arrogant or flippant

#### Systems analysis



FIG. 15. Programme file.

approach to the questioning can kill a project no matter how fine the design of a new system. Even if a system is installed, it will not succeed in the face of antagonistic users, or even more to the point, of the antagonism or resistance of those who must provide input. During the interviews, an image of sincerely wanting to solve the problems of the people being interviewed should be established. This will enhance the success of any system ultimately installed.

An interview must be planned. The systems analyst must know why he is there and what information he is after. Prior planning will not only ensure that the needed information is obtained, it will also enhance the respect paid the interviewer by the interviewee.

Before beginning the interview, the systems analyst should confer with

and obtain co-operation from all department managers to be included in the systems project. In addition to obtaining co-operation from the department managers, the analyst also learns more about the department structure and its relationship to other departments, the employees and their function within the department, and facilities available for conducting the interview.

As in any planning, the first step is to determine the objective of the interview. This will depend not only on the analyst's requirements and the project objectives, but on the person being interviewed. He may be principally a user of information, or he may be significant as a source of information to be input to the system, or he may be an operator or administrator.

The other part of planning an interview is to determine what information is needed. In this regard, it is well to review the tasks that are ahead during the rest of the analysis phase of the project. These tasks determine the requirements that the analyst will have for particular information.

In conducting the interview, the system analyst should ask questions that will get the required study facts in as little time as possible. The analyst should not present himself as a 'know-it-all' or as an interrogator. Before going into the interview, he should have a fair understanding of the duties and responsibilities of the interviewee, along with his working and personal relationships with others in the organization. He should also already have some awareness of the kinds of answers he is looking for and will probably get.

Following are some points that are helpful in conducting interviews:

Explain who you are, what the purpose of the interview is, what the systems project is about, and what contribution the interviewee will make in the development of the new system. A typical follow-up question subsequent to this introduction is: 'At this point, is there anything more you would like to know concerning the systems project?'

Make sure that you have a correct understanding of the interviewee's job responsibilities and duties. A typical question: 'It is my understanding that your job is (a brief job description). Is this correct?'

If the interviewee is responsible for one or more decisions, it is extremely important to ascertain his decision-making model, i.e. what decisions he makes, and how he makes them. The reason that this aspect is so important is because the information produced by the proposed system will be used to help the interviewee make more effective decisions. In this way, both the analyst and the interviewee can come to grips with what information is needed and when. Typical questions: 'As a cost accountant, it is my understanding that to prepare a monthly cost-analysis summary, you need to decide how telephone costs are allocated between departments. Can you decide how this is to be done with the information you are now receiving? If not, what precisely is the information you need and how many days before closing do you need it?'

As much as possible, try to ask specific questions that allow for quantitative responses. A typical question: 'How many telephones do you now have in this department?'

Avoid buzz words, meaningless jargon, and broad generalizations. The type of statement to avoid is: 'The preponderance of favourable responses manifests the apparent feasibility of this undertaking at this point in time and we further believe that the operative status of your function will be significantly enhanced in a most immeasurable manner.'

Develop an awareness of the feelings of the person being interviewed. Learn to listen well. Guard against anticipating answers before the interviewee has had sufficient time to respond.

Maintain control of the interview by using tact and discrimination in ending ramblings and extraneous comments. A typical response: 'Now back to that problem of cost allocation we were talking about earlier. Do you propose that we use toll-call usage as an allocation basis?'

Vague answers to questions should be pursued for full clarification. A typical statement: 'Please bear with me, but I do not quite understand how you propose to handle this.'

Determine if the interviewee has any additional ideas or suggestions or recommendations. It is very important for the systems analyst to give credit where it is due. A typical question: 'Do you want your supervisor or others to know of your suggestions?'

At the end of the interview, summarize the main points of the session, thank the interviewee, and indicate that if there are any further questions, you will return.

Taking good notes of an interview is very important; it is also awkward. People react in different ways to having their words taken down. Some freeze, others start dictating. If the interviewee is too uncomfortable with your note-taking, it is better to put the notebook away and confine your questions to those that are most important. If necessary, come back later for details.

Interview notes should be transcribed in final form as soon after the interview as possible. It is surprising how much that is heard and understood during the interview becomes a mystery or a blank the next day.

Beginning at the introduction and continuing throughout the interview, the analyst should be honest about his lack of knowledge of the interviewee's field of endeavour. If the analyst is not an expert in this field, or has never operated in this type of organization before, he should let the respondent know. The analyst should not try to bluff.

Because the interview process is 'give and take', it is imperative that

a good rapport be established at the outset between the systems analyst and the interviewee. Otherwise, not only the interview, but the entire systems project is doomed to failure.

There is a large literature available on techniques of interviewing, both in the context of systems analysis and in general (see Backstrom and Hursh, 1963; Burch and Hod, 1975; Burch and Strater, 1974; Carlson and Lewis, 1973; Cox et al., 1976; DeMasi, 1969; Matthies, 1976; Survey Research Center, 1969).

#### CREATIVITY TECHNIQUES

In recent years, we have come to the realization that (a) without assistance, creativity varies substantially among different individuals, and (b) techniques can be employed to assist individuals to exercise a much broader and deeper capacity to innovate and produce creative ideas.

A number of specific techniques have been developed and become well known. Examples are known as the Delphi technique, brainstorming, scenario building, and lateral thinking. Each of these techniques can be learned, and they have all been shown to be of substantial value in enabling people to enlarge their horizons, relax mental constraints and inhibitions, and generate ideas and concepts otherwise missed. Some of the basic elements of these approaches are illustrated by the following quotations:

Ideas on better ways will begin to come-if you have done a thorough job of analysis.

No one knows quite how the subconscious mind works, but we know that it does work.

Ideas on what one facet of the new system should be (the output of your subconscious mind) may come to you at any time. You may be in church, travelling in your car, at the office or at a party. You may be at the theatre. Carry a pencil and paper with you at all times.

After a period of gestation, ideas about what the new system should be will flutter out of your mind. They won't necessarily flutter out during working hours. They may come to you at any time. Be prepared to catch them. When the idea comes, write it down. You may be tempted to say, 'Oh, I'll write that down later. I'll remember it.'

You won't remember it. At least you won't remember it with all its potency. You might lose the idea completely. And you can't afford to lose a good idea. You worked hard to get it. Capture it now by writing it down.

These methods will not be discussed further here, since a number of them are covered in some detail in the chapter by Gus Root in this volume. For references see Barron, 1963; Campbell, 1977; de Bono, 1969, 1971, 1972; Linstone and Turuff, 1975; Matthies, 1976; Parnes and Harding, 1962; Prince, 1970.

# COST/BENEFIT ANALYSIS

Justification of a proposed system requiring a capital investment should always incorporate cost/benefit analysis. This analysis weighs the direct and indirect benefits of a proposed system against resource constraints or costs. In short, it determines if the proposed system produces benefits that outweigh its costs. Normally, this analysis is performed on a number of desirable alternative systems and by simple comparison shows which one is best.

Cost/benefit analysis of systems faces some extremely complex issues. We are a long way from developing fully satisfactory approaches. Nevertheless, an organization that sinks vast sums of money into the development and operation of a system cannot ignore the complexities; it must deal with them in the best way it can.

The various costs involved in a system project can be categorized by type, behaviour, function and time.

# Definition of costs by type

- Direct costs: These costs represent expenditures that can be identified as resulting from the proposed system.
- Indirect costs: These are overhead costs that cannot be easily identified with a specific system and are thus portioned out to each area in the organization. Examples are rent, insurance, taxes, management salaries and employee benefits.

# Definition of costs by behaviour

- Variable costs: These costs fluctuate with volume changes in a direct manner. Examples are electrical power and supplies (i.e. if the volume of work increases, the use of electrical power and supplies will also increase).
- Non-variable costs: These costs might vary from period to period, but not in response to volume changes in a particular period. Examples are depreciation, rent, taxes and management salaries.

# Definition of costs by function

- Development costs: These are costs incurred to bring something into being or making it better, more useful, etc. Examples are those costs expended to develop a system, such as training and programming.
- Operational costs: These are costs that must be expended to make

something work or perform. The employment of a computer operator involves operational costs.

Maintenance costs: These costs are expended towards the support, upkeep and repair of the system. Examples are replacement parts and components, and the wages paid to maintenance technicians.

#### Definition of costs by time

- Recurring costs: These costs are repeated at regular intervals, such as payroll costs or equipment rental payments.
- Non-recurring costs: These costs will be incurred at some specific point in time. The cost of computer program development is a non-recurring cost (the cost of maintaining the computer programs is recurring).

Tangible benefits are those that are easily known in monetary terms. This clearly is possible in the case of projects aimed at clear-cut reductions. Elimination of clerical operations or lowering the cost of equipment rental (through greater efficiency or use of a later-generation computer, for example), are common examples of such projects. They are often accompanied by little or no change in information quality; essentially the same information is provided (hopefully) at lower cost. Under these circumstances, the analyst need not concern himself with determining benefits, since they remain the same as before.

A cost-reduction project thus can be viewed as a straightforward



FIG. 16. Analysis of tangible savings from an investment in a system project. Tangible costs and benefits can be handled using traditional methods of investment analysis.

investment. As such, traditional methods of analysis can be applied. Certain expenditures must be made to implement the modified system. These are then followed (if all goes well) by reduced future operating expenditures. This is illustrated in Figure 16. One can analyse the investment in terms of net present value, internal rate of return, or payback period (Helfert, 1967; Weston and Brigham, 1972), as with any investment proposal. Project selection should be dealt with in the same way as that used to set priorities on other forms of investment.

There lurks a hazard in this; it is too easy. Because cost-reduction projects can be analysed in traditional ways, the organization is often tempted to concentrate on them. This avoids the problem of having to assess benefits of improved applications but runs the risk of misdirecting efforts away from projects that can make more fundamental improvements in organizational performance. This is not to say that worthwhile cost reductions should not be pursued, but the bias is often too heavily weighed in their favour.

Really significant contributions normally come through enhanced information quality. Many of the benefits from improved quality may be perfectly tangible. The benefits are certainly tangible, for example, when faster customer billing reduces cash requirements. The translation into monetary savings requires an estimate of both an annual rate of return and the total expected cash reduction. If the rate of return is estimated to be 20 per cent (based perhaps on the opportunity cost of internal investments), a \$100,000 reduction in cash is worth \$20,000 per year.

The principal difficulty in assessing benefits of this sort is the estimation of the effects of the improvements. In the above example the analyst must estimate the relation between cash requirements and billing time. If it can be assumed that faster billing will not change the distribution of time lags between the receipt of an invoice by a customer and the receipt of his payment, then the relation can be estimated quite easily. For instance, a two-day reduction in billing, with an average daily sales of \$50,000, will result in a \$100,000 reduction in cash requirements.

Sometimes a simple model can be used to provide boundary estimates of benefits—i.e. an upper or lower limit on possible benefits. An upper limit can be used to reject unworthy projects, while a lower limit can identify worthy projects. For example, in considering the implementation of an improved forecasting system, one can establish an upper limit by assuming that the system will give perfect forecasts. In a similar fashion, an upper limit on the improvement in production scheduling can be obtained by assuming 100-per-cent-capacity utilization and no interference among jobs. A proposed system can be rejected out of hand if the estimated cost of implementation exceeds the upper limit on benefits. If this is not the



FIG. 17. Use of upper and lower limits in analysing benefits. Suppose that the upper and lower limits shown in the diagram have been estimated for a given project. The project must provide sufficient benefits to justify its estimated cost. If point 'A' represents the cost, the project is worthwhile by virtue of the fact that the lower limit on benefits exceeds the expected cost. Similarly, the project can be rejected if point 'D' is the minimum expected cost, since this value exceeds the upper limit on benefits. The acceptance of point 'B' and the rejection of point 'C' require further analysis to refine the estimates.

case, more refined analysis can often lower the upper limit in order to bring it closer to the actual value.

Lower limits can be estimated in a variety of ways. One means is to put a value on only the most easily determined benefits. In analysing a customer billing project estimated savings from a reduction in cash requirements give a lower limit on total benefits. Actual benefits might exceed this limit by the (unknown) value of fewer errors in billing, better by-product information for market analysis, and the like. Another technique is to calculate benefits by using worst-case estimates for probabilistic variables. If a project is acceptable using the most pessimistic estimates, actual benefits are highly likely to prove even more attractive (see Fig. 17).

A similar approach is to estimate the break-even improvement necessary to balance the cost of implementing a project. Suppose, for example, that a firm spends \$1 million per year on advertising in magazines and newspapers. The cost of obtaining and analysing readership data is estimated to be \$10,000. Thus, a 1-per-cent reduction in advertising expenditures (while holding exposure constant) would justify the cost of the analysis. Similarly, a supermarket firm might determine that the cost of maintaining detailed sales statistics could be balanced by an increase in sales of 0.3 per cent. Given such an analysis, an experienced manager can often judge whether or not the likely improvement exceeds the break-even point.

The benefits discussed so far are tangible enough, even though they may be difficult to estimate. Some benefits, however, are especially difficult

#### Systems analysis

to translate into a monetary value. For example, it would be very difficult indeed for General Motors to put a monetary value on the customer goodwill brought about by reduction from 10 to 5 per cent in the probability that a dealer will not have a needed repair part. It would be equally difficult for the Southern Railway System to attach a monetary benefit to the improved service achieved by a one-day reduction in average delivery time of freight shipments.

Almost any benefit can be assigned a tangible value if sufficient effort is devoted to the task; the difference between a 'tangible' and an 'intangible' benefit thus lies in the difficulty of estimating monetary value. Even a benefit such as goodwill could be translated into a reasonable estimate of monetary value if the effort were justified. If this is not the case, however, we must deal with the problem in other ways.

It should be pointed out that difficulty in expressing a benefit in monetary terms does not imply that the benefit cannot be quantified. Failure to appreciate this fact has often resulted in an unnecessary lack of specificity in describing intangible benefits. Thus, in the above example, it was possible to quantify benefits (e.g. out-of-stock probability reduced from 10 to 5 per cent), even if no monetary value was attributed to them. In some cases, very little quantification may be possible (and so a narrative description must suffice), but this is the exception.

Under some circumstances, the analysis of intangible benefits becomes fairly straightforward. This is the case when a proposed system can be justified on tangible grounds alone (i.e. through some combination of cost reductions and tangible benefits), while also contributing significant intangible benefits. The proposed system is said to dominate the existing one, since it is superior in terms of both its tangible and intangible characteristics.

Opportunities of this sort are not as uncommon as one might suppose. Existing systems are rarely as efficient as they could be, and therefore offer considerable potential for cost reductions. Such savings are made all the greater when technical advances permit new economies. Rather than concentrating solely on tangible effects, however, an organization usually finds it worthwhile to take some of its gains in intangible form.

For example, a comprehensive airlines reservation system may be justified on such tangible grounds as reduction in the salaries of reservation and ticketing clerks, higher seat bookings and more efficient routing and scheduling of flights. Improved customer service may be as valuable as any of these, but it is probably not possible to assign a monetary estimate to it.

As systems become more comprehensive and integrated, formal methods of analysis become increasingly difficult to apply. Joint costs and joint benefits often make it impossible to determine the payoff from any one subsystem. As the system begins to pervade day-to-day operations and higher level decision-making, benefits become increasingly difficult to evaluate in monetary terms. Inevitably, then, we must rely on experienced judgement, as well as technical analysis, to make cost/benefit decisions.

There are a number of treatises dealing in substantial detail with the techniques of cost/benefit analysis (Frost, 1971), some of them oriented more specifically to its application to computer systems. In addition, there are texts of more general interest concerning the criteria commonly used to choose among competing financial investment alternatives.

# Conclusion

The beginning of this discussion was at a very broad level, while recent sections have entered into considerable detail. Before concluding, it is important to return to a global level and reflect upon the links. What is the relationship of system analysis to systems approach?

The systems approach recognizes that we must not lose sight of a global view—one from which we see the entire forest. In order to understand the entire system adequately and to be able to generate new or improved subsystems within it, systems analysis requires us to descend to a level of detail to examine the various aspects of each tree. In making this breakdown for analysis into trees and coming to know the individual tree well, we must never lose sight of the entire forest and the interaction between the one part and the whole. Whatever modification is made to one tree will affect the entire forest. Any changes made within one department, whether changes of job definition, work flow, procedure or equipment, will have an impact on the entire organization. The analyst must always retain the ability to step back and re-insert the parts into the whole, to reflect on the impact on the entire organization of proposed changes in one small unit.

This whole chapter has been devoted to organized common sense; systems analysis is little more than that. But many years and countless applications of the approach have proven its value beyond doubt.
## References

- BACKSTROM, C. H.; HURSH, G. D. 1965. Survey Research. Evanston, Ill., Northwestern University Press.
- BARRON, F. 1963. Creativity and Psychological Health. New York, Van Nostrand.
- BURCH, J. G. Jr; HOD, N. 1975. Information Systems: A Case-Workbook. New York, John Wiley & Sons, Inc.
- BURCH, J. G. Jr; STRATER, F. R. Jr. 1974. Information Systems: Theory and Practice. Santa Barbara, Calif., Hamilton Publishing Company.
- CAMPBELL, D. 1977. Take the Road to Creativity and Get Off Your Dead End. Niles, Ill., Argus Communications.
- CARLSON, R. D.; LEWIS, J. A. 1973. The Systems Analysis Workbook: A Complete Guide to Project Implementation and Control. Englewood Cliffs, N.J., Prentice-Hall Inc.
- CHAPIN, N. 1971. Flowcharts. Princeton, N.J., Auerbach.
- COUGER, J. D.; KNAPP, R. W. 1974. System Analysis Techniques. New York, John Wiley & Sons, Inc.
- COX, K. K.; HIGGINBOTHAM, J. B.; BURTON, J. 1976. Applications of Focus Group Interviews in Marketing. *Journal of Marketing*. (Chicago), January, p. 77-80.
   DE BONO, E. 1969. *The Mechanism of Mind*. New York, Simon and Shuster.
- -----. 1971. Lateral Thinking for Management. New York, American Management Association.
- -----. 1972. PO: Beyond Yes and No. New York, Simon and Shuster.
- DEMASI, R. 1969. An Introduction to Business Systems Analysis. Reading, Mass., Addison-Wesley Publishing Co.
- FROST, M. J. 1971. Values for Money: The Techniques of Cost Benefit Analysis. London, Gower Press Ltd.
- HELFERT, E. A. 1967. Techniques of Financial Analysis. Homewood, Ill., Richard D. Irwin, Inc.
- IBM. 1962. Decision Tables, A Systems Analysis and Documentation Technique. IBM Data Processing Techniques. (White Plains, N.Y., IBM Technical Publications Department.)
- LINSTONE, H. A.; TURUFF, M. 1975. The Delphi Method. Reading, Mass., Addison-Wesley.
- MATTHIES, L. H. 1976. The Management System: Systems are for People. New York, John Wiley & Sons, Inc.
- MCDANIEL, H. 1968. An Introduction to Decision Logic Tables. New York, Wiley.
  —. 1970a. Applications of Decision Tables—A Reader. Princeton, N.J., Brandon/ Systems Press.
- -----. 1970b. Decision Table Software. Princeton, N.J., Auerbach.
- McFARLAN, F. W.; NOLAN, R. L. 1975. The Information Systems Handbook. Homewood, Ill., Dow-Jones-Irwin, Inc.
- PARNES, S. J.; HARDING, H. F. (eds.). 1962. A Source Book for Creative Thinking. New York, Charles Scribner's Sons.
- POLLACK, S.; HICKS, H.; HARRISON, W. J. 1971. Decision Tables Theory and Practice. New York, Wiley.
- PRINCE, G. M. 1970. The Practice of Creativity: A Manual for Dynamic Group Problem Solving. New York, Harper & Row.
- RUDWICK, B. H. 1969. Systems Analysis for Effective Planning: Principles and Cases. New York, John Wiley & Sons, Inc.
- SURVEY RESEARCH CENTER. 1969. Interviewer's Manual. Ann Arbor, Mich., University of Michigan.
- WESTON, J. W.; BRIGHAM, E. F. 1972. Managerial Finance. New York, Holt, Rinehart and Winston.

# Chapter 4

# Modelling and simulation

Gus Root

And the Prime Minister said: 'Only 76 per cent of our men and 57 per cent of our women are literate. Literacy rates are the lowest in the rural, agricultural areas, and among the poor, the undereducated and the underemployed. Your task is to plan and conduct a campaign to improve literacy rates among our poorest and most needy people.'

How should we approach such a task? Where should we begin? There are at least two different types of approach to this kind of social problem, and it is important to look at and understand the strengths and limitations of each:

- First, the 'immediate-action' approach: We could study and learn from the experiences of those who have worked on literacy problems in many different countries around the world, and act upon the very best recommendations that fit our situation.
- Second, the 'analytical' approach: We could perform a detailed analysis of the society in which literacy appears to be a primary problem, identify the main interacting factors that enhance as well as hinder literacy in that environment, and find those places in the social system where small changes would be most likely to produce long-term and lasting improvements, with the smallest number of undesired and unanticipated negative effects.

For those who choose to follow the first (immediate-action) approach, there are many reports available and many persons with a wealth of practical and useful recommendations. This is the approach usually taken by those who face an urgent problem and feel compelled to take quick and effective action. This approach may lead to enough success to encourage increased effort.

For those who choose to follow the second (analytical) approach, there are only a limited number of reports on systematic planning in a particular project area, and a small number of persons with experience who can make specific recommendations on techniques for analysis-based planning. This is the approach usually taken by those who are deeply concerned about the use of scarce resources (money, talent and time) in problem areas where many efforts have had limited success and occasional unexpected difficulty. Often this approach leads to cautious progress in many related sectors of the social environment.

Regardless of the basic orientation a communication planner brings to the task, most would agree that a certain amount of analysis should precede action—it is only a question of the amount and type of analysis. At the same time that analysis is started, it is also advisable to gather the available information in a project area (literacy) and get to know those who have worked in that area.

So, the beginning of any project is a busy time. The main theme of this paper is: Begin any effort to change a complex social system by constructing models<sup>1</sup> and simulations<sup>2</sup> of the problem and its social context; they will increase the likelihood of long-term success, decrease the chance of unexpected negative side effects, and help conserve scarce national resources. From this point of view, the wisest approach may be a combination of the two approaches suggested above (analysis and action). It would be well to wait until the completion of the first (modelling and simulation) before acting upon any recommendations from the second.

The following sections of this chapter will introduce and illustrate a variety of modelling and simulation techniques, and will try to illustrate the strengths and limitations of each—as well as to suggest how several techniques, used in combination, may provide the best insights and approaches to a practical social problem such as the reduction of inequalities in literacy rates.

# Notes on changing social systems

From the 'systems' point of view every social event is the result of the interaction of a large number of cause-effect relationships, in which each 'cause' is itself influenced by the 'effect' it produces. Thus, to introduce a new 'cause' into a social system (such as an effort to improve literacy) with confidence that it will produce the 'effect' we want (an improved distribution of literacy and social development) with the fewest negative side-effects (migration of the illiterate to cities and social erosion), we must

<sup>1.</sup> *Model*: A representation of selected characteristics of some situation or problem of interest, e.g. a description, diagram or scaled-down replica.

<sup>2.</sup> Simulation: An activity in which a model is operated (run, played), with the expectation that its performance will duplicate (simulate) certain critical features of the situation that is represented by the model.

understand the dynamic structure of the many interactions that are interdependent in their influence on literacy and social development.

But the human mind is severely limited in its ability to perceive and interrelate all of the complexities of a social system and, from this understanding, predict the system's response to a change effort. A developing society includes many interconnected sectors (economics, education, agriculture, communications, etc.) with many feedback loops and time delays between an action in one sector and its reactions to other sectors (and, of course, the subsequent effect back upon the original action). The human brain cannot take in, process and act upon an integrated understanding of all of these complexities without the use of some outside-the-brain assistance. This assistance usually comes in the form of theories that describe these complex relationships, models that represent these networks of cause-effect loops, and simulations that imitate the most important behaviours of these systems.

Figure I suggests that these kinds of 'aids to thinking' progress from verbal or 'prose' models (such as descriptions and/or theoretical statements), to representational or 'structural' models (such as diagrams and charts), and finally to simultaneous or 'dynamic' models (such as models that operate within wind-tunnels or computers, and human simulation-games).

The sequence of lines in Figure 1 suggests both (a) the successive stages of model building and the enriched understanding that can come from each successively more complex stage, and (b) the organization of the following



FIG. I. A representation of the manner in which successively more complex models can enrich understanding (our mental models), which determines the policies, decisions and actions we undertake.

sections of this chapter. A brief review of these basic ideas may help to interpret the lines in Figure 1 and also give an introduction to the remainder of this chapter.

- Line 1: Initial efforts to understand a problem and the situation within which the problem is embedded (the problematique) usually lead to prose models, such as descriptions, special vocabularies, scenarios, operational definitions and theoretical statements. Prose models are discussed in Section I of this chapter.
- Line 2: New information and understanding can come from the process of defining, describing, talking about and arguing the characteristics of the problematique. A fruitful outcome of such prose models is an enrichment of the mental models that provide the basis for action.
- Line 3: There are special techniques for manipulating prose models, which in turn can lead to
- Line 4: further enrichment of our mental models.
- Line 5: Beyond prose models, additional techniques have been developed for the construction of structural models, which portray graphically the relationships among the critical components of the problematique (as identified and defined in prose models). Structural models are discussed in Section II.
- Line 6: The process of developing structural models often leads to new insights (enriched mental models) in those who participate. Experience suggests that these new understandings lead to enriched prose models, which in turn aid in the refinement of the structural models.
- Line 7: Techniques for manipulating structural models can lead to
- Line 4: even further understandings.
- Line 8: Beyond structural models, special techniques can lead to the development of dynamic models, which imitate or simulate the critical performances of the problematique, as have been expressed in both prose and structural models. Dynamic models are discussed in Section III.
- Line 9: The development of dynamic models leads to a greater understanding of the problematique—and this usually requires frequent cycling through the earlier stages of prose and structural models.
- Line 10: Manipulating dynamic models can lead to
- Line 11: still further enrichment of our mental models, and this is the continuing goal of all modelling effort.

The following sections of the chapter will describe each of these three types of model, suggest some simple techniques for constructing models of each type, and will give illustrations of both the procedures and results of several modelling projects.

# Some preliminary definitions

Simple definitions and illustrations of concepts that are central to the ideas of this chapter are relatively easy to provide. Slightly more complex descriptions and illustrations will follow. A full understanding of these concepts can be developed only by those who read widely in the literature of this field and participate in a variety of modelling projects. (That is to say, it seems reasonable to predict that a reader's mental models of these concepts will also develop through the processes suggested in Figure 1.)

The purpose of this chapter is to introduce these ideas and provide sufficient information to enable practical, non-expert kinds of people to build simple but useful models (of all three types), and use them as aids to decision-making in important, practical situations.

Models are simplified representations of selected characteristics of some situation of concern (some problematique). Social problems exist within a complicated network of factors, which produce the problem symptoms. This network is usually sufficiently complicated to obscure the basic structure of the cause-effect relationships within which the problem is embedded. A model attempts to simplify this situation, and to abstract the most important characteristics of this network. Common examples of models include verbal descriptions; drawings, maps and charts; toys, replicas, and scaled-down copies; and sets of equations and computer programs.

Simulations are activities in which models are 'operated' or 'played' in the hope that the model will imitate (simulate) the critical performances of the situation (problematique). The purposes of simulations may vary to include education, research, design improvement and/or the exploration of the probable effect of different policy decisions. Common examples of simulations include model airplanes that 'fly' within wind tunnels in order to let designers observe and improve the performance characteristics; business games that are based upon models of the market environment and are conducted to train executives; computer models that are 'run' under a variety of initial conditions to indicate the probable effects of different allocations of resources among the sectors of the model.

Decisions are the action-commitments or resource-assignments of individuals or groups when faced with a situation for which there is no existing agreement, routine or habit. Decisions are choices made after an analysis of the implications of a set of alternative possible commitments or assignments. The purpose of a decision is to respond to the changing demands of a situation in a planned or deliberate manner---rather than reacting on the basis of habit, tradition or whim. Examples of decisions are the selection of a particular type of telephone system after the analysis of several proposals and a set of goals for the region to be served; the allocation of resources among educational, agricultural and literacy programmes according to a formula, rather than according to the rhetoric or social power of the persons requesting support; the establishment of a national information-and-news bureau, rather than leaving the news services in the hands of private-profit corporations or international agencies.

## I: Beginning with prose models

In building any model, it is as though we try to express all we know and believe about the system of concern (the problematique) and represent that current knowledge in the model, thus leaving our minds free to think about other information or our next activity. If we can bring back this model whenever we need it, we will not have to remember and reconstruct all of that information every time we need to think about that problematique. In this way, models act as temporary storage for large amounts of organized knowledge and are very useful aids to our thinking about social problems.

The process of model building usually begins with an awareness that something is not quite as hoped, anticipated or expected—it begins with a wonder, a worry or a problem. This 'awareness' stage (Fig. 2) is usually followed by an attempt to define the problem and its context (which we are calling the 'problematique').



FIG. 2. The awareness stage.



FIG. 3. An analysis of the problematique.

The analysis phase of model building finds the individual (or group) examining the characteristics of the problematique over time. The analyst seeks to determine what has happened, what is happening now, and what is likely to happen if the current trends continue (Fig. 3). It is not enough for a single individual to work this out alone: usually he asks, listens and reads.

The analyst's concerns, observations and expectations are often expressed in words (his mental model is transformed into a prose model) as he talks with others and writes about his preliminary conclusions. There are many advantages to be gained from the preparation of prose models:

- 1. Public prose models can be examined by others whose perceptions may be different, thus leading to refinement, correction and confirmation.
- 2. Prose models can include a larger number of elements and relationships that can be handled within a single person's mental model.
- 3. Written prose models can be examined in a non-sequential manner, without the danger of losing those elements of the model that are not currently being examined.
- 4. Preparation of prose models demands that concepts be expressed with a degree of precision and organization that is not required by mental models, whose 'fuzziness' may not be apparent until the precision of language and the structure of formal presentation are imposed upon those concepts.

A number of techniques have been developed to facilitate the preparation of prose models. Because the prose model appears to be the first step in any modelling effort, it is appropriate to examine a few of these techniques in some detail.

# BRAINSTORMING

The essential characteristics of this technique centre around the group production of a full range of ideas (from the immediately practical to delightful flights of fantasy) under conditions of acceptance and the absence of criticism.

The technique of brainstorming is usually to:

- 1. Bring together a relatively small group of persons (four to fourteen) with knowledge and interest in the problem and its context.
- 2. Introduce the topic in its most gross and simple form.
- 3. Present the rules for working together on the problem, such as: (a) each person is encouraged to contribute any idea that may be relevant to the topic, regardless of how strange or absurd, and the contribution will be accepted and recorded; (b) negative or critical judgements of others' contributions are to be withheld (and small fines may be extracted from anyone who criticizes another's suggestion).
- 4. Have a recorder write each contribution publicly on a chalkboard or easel pad, to keep all information before the group.
- 5. Introduce the next level of problem complexity and ask individuals (working alone) to simplify, cluster and organize the items, to select the most promising items according to criteria provided, and to expand or elaborate upon the previous items to account for the new information provided.
- 6. Repeat steps 3, 4 and 5, above, as many times as are fruitful, with additional information supplied at each iteration. This process may be continued until the group has come to some acceptable degree of completion of the tasks of analysis, model definition, examination of alternatives, and selection of a preferred approach to problem resolution. How far the process is carried may depend upon the skill of the leader and the patience of the participants.

# AN EXAMPLE OF BRAINSTORMING

A group of educators was concerned about ways to improve the effectiveness of community education (locally initiated education activities designed to enhance citizens' ability to work together for the common good). Because of strong national interest in stimulating these local activities, agencies had been established at federal, state and regional levels to stimulate and support community education with money and programmes. Professional organizations had been established to encourage the professional development of the persons involved in community education. However, there was little agreement on the exact elements of the networks that interrelated all of these people, organizations and activities. There was less agreement on the characteristics of the networks that interrelated all of these people, organizations and activities. There was even less agreement on which of these elements and/or relationships would be most important for the further development of the network's effectiveness.

Four community educators came together for a two-hour brainstorming session in an office with a chalkboard. A consultant (not a community educator) led the group to create a long list of the factors that affected and were affected by community education (CE), or the causes that influenced CE, and also the effects produced by CE. In the second round of the brainstorming activity, individuals clustered and categorized these factors in order to simplify the list that had been generated. A set of factor-clusters was developed that all agreed was adequate to account for the major elements and relationships within the CE-network. In a third 'round', the participants were led to identify the magnitudes and directions of the major influences among these factor-clusters within the network, with a resulting prose-model that expressed the agreements among their individual perceptions (and de-emphasized the disagreements). Spontaneous expressions of surprise and pleasure followed this exercise, as a result of being able to see actions that could be taken that all felt would be constructive and each felt had a higher probability of being effective than the individual ideas that did not get group confirmation.

Advantages of the technique include:

- A high level of emotional involvement in the problem and the group effort (if the setting has been adequately planned and the activities are directed with skill and good humour). This can lead to wide acceptance of the outcomes of the process.
- Possibilities that might initially seem absurd can be given serious consideration, and may be developed and adopted when they might not have been considered under other conditions.
- The contributions of persons with widely differing backgrounds can be focused and integrated within a short period of time.

Disadvantages of this approach may arise from:

The personal characteristics of participants, among whom there may be persons with charisma or political power who may attempt to push for their own ideas, or others who are shy or socially unskilled and whose ideas may not be heard or recognized. These difficulties may be overcome (or minimized) by skillful group leadership, or by the use of external consultants with sufficient prestige to moderate the social interaction and permit the group to focus on the problem.

The technique, which may appear to encourage trivial and ridiculous thinking about serious problems. This criticism is most likely to be raised by those who feel strong pressures for immediate action, based upon solutions that have been 'proven' in many similar situations. It is very difficult to convince such 'practical' people of the advantages of exploration and analysis, and it is sometimes constructive to assign such persons to the task of getting to know professionals who have been effective in other situations and the approaches they have used. That information could also help to enrich our verbal and mental models.

## DELPHI

The essential characteristics of the Delphi approach to the construction of prose models centre on the response of a large number of persons to printed materials through successive 'rounds' of questioning and on giving feedback to each participant about his/her response compared with those of all other participants.

The technique usually involves:

- 1. Identifying a large number of persons with various types of expertise relative to the problem under consideration, and securing their agreement to participate in an extended series of questions-feedback-questions.
- 2. Asking each participant to identify/describe the most important characteristics of the problem, influences that affect it, and possible future conditions relative to the problematique. Each person is asked to respond individually.
- 3. All responses are analysed and organized to identify similarities and differences among subgroups of the participants.
- 4. Each participant is given information on all other responses including the means and deviations of various subgroups, and shown his/her response relative to the others'. The individual is asked to consider modifying the initial responses or justify any differences observed. Additional information may then be given to the individual and new questions presented.
- 5. Repeating the above steps until the various positions have been clarified and defended, and the purposes of the analyst have been achieved (or the patience of the participants exhausted).

#### AN EXAMPLE OF A DELPHI APPROACH

A school of social work wanted to design an undergraduate curriculum derived from the best thoughts of a wide variety of administrators, scholars and practitioners within the field of social work. The first rounds of the Delphi process were used to identify the categories of persons who should be involved in the study, and the specific persons within each of these categories. In the successive rounds of the study, broad subject areas were identified, specific details were developed within each broad area, and the appropriate balance of emphasis among the topics was established. At each step of the process, the major differences of opinion were identified, specific details were developed within each broad area, and the appropriate balance of emphasis among the topics was established. At each step of the process, the major differences of opinion were identified and defended or criticized. The school's responsible faculty was immediately involved in each step of the process, and was able to observe differences between its perceptions and those of respected colleagues, as well as significant differences within its own body. The outcome of the study was a high degree of confidence that the resulting curriculum was the best that could be devised at that time, an increase in the reputation of the school within the profession of social work, and a new appreciation of the faculty for the differences and agreements among its own members.

Advantages of the technique may include:

- The involvement of a large number of persons from many different disciplines and perspectives.
- The participation of persons throughout a broad geographical area without the costs and time required to come together at a meeting.
- The avoidance of interpersonal forces that might inhibit some contributions and give undue weight to others because of the command or charisma associated with the person supporting a particular position.

Disadvantages of this approach to building a prose model may include:

- The difficulty of preparing and presenting questions and issues in a way that is interpreted appropriately by a wide variety of persons at some distance from the analyst.
- Difficulties associated with the interpretation and analysis of the responses, including the question of the reliability of procedures for clustering and simplifying responses received.
- The difficulty of maintaining the serious attention of the participants throughout the long (and sometimes boring) process of responding to many detailed items, repeatedly, over a long period of time, with minimum social interaction with other participants and the analyst.

These difficulties can best be handled by the employment of research-trained analysts, selecting participants on the basis of their inherent interest in the problem to be studied, minimizing the time required for the study, establishing at least minimal personal contact with the respondents, giving public recognition to their participation, and building confidence that results will influence important decisions and actions—which will, in turn, influence them.

#### SCENARIO WRITING

The essential characteristics of this approach to the construction of prose models centre on the description of alternative possible future conditions, each attributable to a particular policy or decision on the allocation of available resources (money, time and talent).

The technique usually involves:

- 1. The identification of groups of persons who support particular positions with respect to a problem or decision that will affect a large number of persons or groups.
- 2. The preparation of a written description of each group's anticipated outcomes resulting from the decision it supports, including a logical analysis of the steps anticipated from the decision to that anticipated future.
- 3. The analysis of all 'scenarios' by a select group of participants and consultants, to identify the critical descriptive-dimensions of each scenario.
- 4. An agreement among all groups to rewrite their scenarios in a common format, organized to include consideration of the significant descriptivedimensions, and edited by external consultants to assure a common level of rhetoric and language usage (to remove advantages or disadvantages for any one group due to the different verbal skills of its members).
- 5. An agreement among all groups of the value-dimensions on which a decision should be based when choosing from among the alternatives. A variation of the 'brainstorming' approach described above can be used to identify these value-dimensions and establish a relative weighting among them.
- 6. All scenarios are distributed and studied by each group.
- 7. The groups then come together to evaluate all scenarios according to the weighted value-dimensions. Depending upon the amount of continuing disagreement and emotional rhetoric elicited by the need to select one scenario as the basis for action, the technique at this time may range from simple discussion to a closely organized procedure for differentially weighting the alternatives (see the example for a description of a closely organized approach).
- 8. Following a detailed consideration of all scenarios, the responsible decision-making group (or individual) reaches and announces a decision.

#### AN EXAMPLE OF SCENARIO WRITING

A moderately large organization had suffered a major disaster that had all but destroyed its physical facilities. The problem facing the organization concerned deciding among the alternatives of (a) rebuilding in the same location, (b) merging with one or more other similar organizations, or (c) building new facilities in a totally different location. Groups had formed to support each of the alternatives, and each had enough power (financial, political and personal) to prevent any clear decision from being reached without major fragmentation of the organization.

Planner-consultants were asked to assist all groups in reaching a high degree of consensus. In a series of meetings, five scenarios were prepared—one for each of the positions strongly defended. In the process of writing and editing the scenarios, all groups came to agree upon the characteristics of the future organization that should be dealt with in each scenario. Following this analysis, a set of value-dimensions was developed by the entire group of decision-makers, as well as the relative weighting that should be given to each dimension. The process was approximately as follows:

- 1. In a brainstorming session, a set of seven value-dimensions was established and given enough agreement that everyone was willing to proceed.
- 2. In a small number of 'rounds', the value-dimensions were placed in a list by 'rank order', from the most important at the top to the least important at the bottom of the list.
- 3. The bottom item was given an arbitrary weight of ten.
- 4. The next to the last item was then assigned a number relative to the bottom item.
- 5. The third to the last item was assigned a weight that placed it relative to the two lower items. This led to reassignment of numbers to the lower items.
- 6. This process was continued until all items were assigned numbers indicating their relative weightings.
- 7. All of these 'importance weights' were added; each value-dimension's weight was divided by the sum of all weights and multiplied by 1,000. This computation merely converted the weights into numbers something like percentages (purely arbitrary).

A matrix was then prepared that displayed these seven weighted valuedimensions down the left side, and the five alternatives across the top as suggested in Figure 4.

Each individual decision-maker was then asked to consider each row of this matrix, and distribute the number of points allocated to that valuedimension among the alternative scenarios according to his/her belief

#### Approaches to communication planning

Value dimensions	Weight	Alternative scenarios				
		A	В	С	D	E
1	210					
2	190					
3	170					
4	150					
5	130					
6	90					
7	60					
Total	1,000					

FIG. 4. A decision-oriented matrix of value-dimensions on alternative scenarios.

about the relative potentiality of each scenario for achieving that value. Each decision-maker then publicly displayed his/her point assignments (the number in each cell of his/her matrix). When the columns for the scenarios were summed up, alternative 'C' amassed more than half of the total available points, and its total was more than twice that of any other alternative. After a long moment of silence as the group considered these outcomes, a quiet and nearly unanimous vote was taken to adopt scenario 'C' and its implied decisions. One group of participants that had been highly vocal in earlier discussions announced that it would separate from the others, and that separation was accepted by all parties with significant understanding and goodwill. Action was immediately taken to implement proposal 'C'. Advantages of this approach usually include:

- Each participant and group has a full and equal chance to have its arguments and convictions received and carefully considered.
- Differences between groups in terms of power, charisma and verbal skills are equalized.
- The critical value-dimensions on which all alternatives will be judged are openly examined and given weightings that are physically determined and displayed.
- Decision-making power is permitted to rest in normally constituted authorities, but every viewpoint has the same opportunity to contribute information to influence that decision.
- Disadvantages of this approach to prose models include:
- The unequal abilities of groups to formulate and express their opinions, perceptions of cause-effect relationships, and decision criteria—despite the efforts of consultants to assist them.
- The special skills required by the consultants to perform in this highly emotional, deadline-oriented situation.
- The difficulties which may follow this procedure, after the decision has been made, when persons with power may take action relative to the others based upon actions they may have taken during scenario processes—and participants are usually well aware of that possibility during the experience, and may adjust their behaviour accordingly.

Although these difficulties may be insurmountable, the value of scenario writing—often in a truncated form of the procedure noted above—is very high, because of the clarification of issues that results.

#### MANIPULATING PROSE MODELS

As suggested in Figure 1, prose models can enrich our mental models in two ways: through the procedures for developing the prose models, and through manipulating prose models after they have been developed. The discipline required to clarify and formulate a verbal description of one's perceptions can be a cleansing and purifying experience. But the benefits of prose models are not limited to the values of this developmental process.

Prose models can be manipulated in ways that provide further insight. One example of such manipulation was included in the discussion of scenario writing. Multiple scenarios can be prepared and compared. Even beyond the description of the views of groups who have strong differences, another group may be assigned the task of preparing one or more brainstorming scenarios, or inventing some radically different alternatives, or possible futures. These procedures can force the analyst/designer to (a) identify the most important descriptive-dimensions of a good scenario, or the characteristics essential to a good prose model, and (b) define the critical value-dimension to be used in choosing one from among several scenarios.

# II: Structural models

A major step in modelling occurs when the models used are capable of including and representing a large number of elements and relationships in some compact form. Prose or verbal models have certain inherent limitations that can be overcome by using another system of symbols, other than words. Verbal models are limited in that they can express concepts only through a set of symbols that must be processed in a particular linear sequence, one word or word cluster at a time; the characteristics of an object, event, relationship or rule are not in the words or word-clusters themselves, but are developed by the observer as a result of processing the word-clusters (verbal symbols) in a particular sequence. Other symbol systems have been developed that are also capable of representing and displaying factors that have been found to be critical in prose models-and these other symbol systems may have a number of advantages over verbal symbols (words and word-clusters). Special symbols can represent directions of influence, relative magnitudes and forms of relationships, e.g. arrows, numbers, mathematical operations, which would require many word-clusters to describe them. In addition, other symbols may be considerably less ambiguous than words, forcing greater precision on the symbol user. Models with this additional precision are also more open to analysis, empirical testing and confirmation/criticism/change; verbal or prose models are usually difficult to validate or refute because of their ambiguity.

The above discussion brings us back to a reconsideration of the purpose of models; in this paper, the primary purposes of models are assumed to be (a) the representation of what are thought to be critical factors (elements and relationships) that account for a significant portion of the performance of the problematique, (b) the development of increased understanding of that object system and ways to represent that understanding, and (c) the guidance of efforts to work with and change the performance of that object system in ways thought or felt desired. In simpler terms, the three purposes of models are for improved conversations, research and application. As the symbol systems used in our modelling efforts become more complex, concise and precise, the possibilities of what we can do with our models increase. Thus, this second stage of model development, moving from prose to structural models, opens up new and powerful techniques for understanding and dealing with systems.

#### CROSS-IMPACT MATRIX MODELS

The essential characteristics of this technique include the production of a diagram that shows (a) the significant factors that affect and are affected by the problem of concern, (b) the directions in which some factors influence others, and (c) the relative magnitude of those influences. The intent of this diagram (matrix) is to show visually the relationships among the most important factors related to the problem of concern. When this structure of relationships is arrived at through group participation, it can organize and simplify the group's thinking about the most effective places to initiate design changes to improve the problem-system's performance.

The technique can be illustrated by the following procedures:

- 1. Bring together a small group of knowledgeable and concerned people. Special leadership skills may be required if the group is larger than about eight people and/or there are members with special expertise in the problem area who have social or political power relative to the others, or possess personal characteristics that affect group functions (talkative, shy, strong social needs).
- 2. Follow the brainstorming procedures described above and identify a set of five to twelve factors that most agree are critical to an understanding of the system and its possible influence by the group.
- 3. Prepare a 'matrix' as illustrated below, with these factors arrayed both down the left side and along the top edge of the matrix. Those on the left are labelled 'causes', and those along the top 'effects'.

FIG. 5. A simplified illustration of a blank cross-impact matrix.



- 4. Each person, working individually, is asked to consider each factor as a cause (starting at the left and working across one row of the matrix) and to estimate the relative strength of its effect on each of the other factors. They are to ask themselves: 'If this cause were to increase substantially, how would that effect react?' Any of several symbols could be marked in the 'cell' of the matrix to indicate the relative strength of this cause-effect relationship; a simple procedure would be to ask the participant to insert a '+2' if the influence is perceived to be strongly positive (an increase in the cause, produces a large increase in the effect), '+1' if the influence is moderately positive, 'O' if there is no influence, '-1' if there is a moderate negative influence (an increase in the cause produces a decrease in the effect), and '-2' if the influence is strongly negative.
- 5. A large blank matrix is mounted on the wall, and the sum of the individual estimates in each cell is determined and displayed. Any large differences among group members may be discussed if there is both time and a constructive atmosphere for exploring such differences (there may be times when anonymity must be maintained for personal or political reasons).
- 6. The final matrix is closely examined to identify the places where the group members might initiate actions with (a) a high probability of achieving desired effects, and (b) little likelihood of causing undesired effects, even through an interconnected series of influences or second-and-third-order effects.

#### AN EXAMPLE OF A CROSS-IMPACT MATRIX

A planning group in a late-developing country was concerned with the distribution of effort within an educational materials development centre (EMDC) for maximum impact upon the educational systems of the entire country. A procedure was carried out involving the members of the planning unit (four persons) and a succession of managers from the separate departments of the EMDC (time demands prevented them all coming together for an extended series of meetings). The procedure can best be described as a combination of the brainstorming technique and the development of a cross-impact matrix. A third feature of the process was a continuing effort to share outcomes of work with the other managers not present during the planning exercise.

From three sequential brainstorming sessions, a set of critical factors related to the network in which the EMDC was involved were agreed to be:

The four major resources used by the EMDC: (A) professional time of the

staff; (B) professional competence of the staff; (C) managerial competence of the staff; (D) the facilities and materials used.

- The two major types of activity in the centre: (E) the preparation of educational media and their distribution, and (F) field work with and training for audiences.
- The two major classes of audiences: (G) school-related people, and (H) government decision-makers.
- The major feedback mechanism: (I) the amount of budget allocated by the government decision-makers.

These nine factors were arrayed down the side and along the top of a  $9 \times 9$  matrix; each participant was asked to indicate the relative strengths of all cause-effect relationships; their individual estimates were added, and the strongest of these were selected as the most important elements in the structure of the EMDC-education system. The resulting matrix, as in Figure 6, was given to all participants. The immediate outcome of this process (and the resulting matrix) was the realization that this matrix did reflect the present structure more accurately than any of their preceding



FIG. 6. A complete cross-impact matrix.

models (or personal thoughts), and that there were probably some serious weaknesses somewhere in that network which should be explored further. (Thus this illustration will be continued following the next type of modelling to be described, causal loops.)

Advantages of the cross-impact model arise from its requirement that the participants examine the relationships between the system's variables (critical factors) systematically, and individually, two at a time. Without some kind of structuring mechanism, most people find it very hard to concentrate on only a limited number of factors and relationships for long enough to express their current perceptions. When participants can see the entire matrix, and realize that they will be able to examine the entire structure, one cell at a time, they seem to gain confidence that this systematic procedure will permit them to express all of their ideas. Thus the discipline of the process appears to provide a large amount of freedom.

In addition, the requirement that people work alone tends to reduce the impact of social factors that might otherwise give too much weight to the perceptions of those with social or political power (or, at the other extreme, too little weight to the perceptions of those without such power). If the group environment permits it, the leader may ask each person to post the numbers in his/her matrix, so that large differences between perceptions can be discussed—but this requires a degree of openness to differences which some groups cannot provide.

Difficulties with this technique are similar to those mentioned above for types of verbal or prose model, and include:

Participant-centered difficulties, associated with the social, political and personal characteristics of the group members. The stronger these influences, the greater the competence and status requirements of the leader in guiding the group to a useful cross-impact matrix.

Technique centered difficulties, associated with (a) some persons' lack of familiarity with a matrix, (b) some persons' inability to limit themselves to considering only one interaction between two factors at a time, (c) the subjectivity of the estimate that is based on personal perceptions, (d) the relative sophistication of some persons who perceive these relationships as non-linear and impossible to describe by a single number in a matrix, and (e) the inability of the technique to consider and display time delays between actions in one part of the matrix and reactions in another region. Most people are willing and able to construct this type of a matrix without difficulty, provided the leader appears confident and in control of the social situation, and is able to encourage participants to complete this effort in order to see its outcomes and then consider the next steps beyond the construction of a cross-impact matrix.

#### CAUSAL-LOOP MODELS

These types of model are very similar to cross-impact models, except that the information is shown in a diagram of names and arrows. The factors are those that have been identified in the prose models as described above. The direction of each cause-effect relationship is determined in a process such as that described for the cross-impact matrix. For many persons, a diagram of factor-names and arrows that show the relationships among those factors is much more easily seen and understood than a matrix that displays the same information.

The technique for constructing a causal-loop diagram is:

- 1. A cross-impact matrix is developed, following procedures roughly described above, and shown in Figure 6.
- 2. Begin most simply by selecting those factors (only three or four) with the highest totals in both the rows and the column of the matrix (Fig. 6). Place those factors in the approximate centre of a sheet of paper; connect arrows between them pointing in the directions indicated by the strongest cause-effect links in the matrix. From Figure 6, the four most influential factors are (E) production and distribution, (F) customer services, (B) management competence, and (H) government decision-makers. These factors and their interactions are shown in Figure 7.



- 3. Additional factors are added to the diagram in rank order according to their combined totals. Initially, avoid adding arrows suggesting that two factors are both causes and effects for each other (include such reciprocal interaction later, only when the meanings of such mutual influence have been carefully explored). The second step in the preparation of the causal-loop would then yield Figure 8.
- 4. The remaining factors are added to the diagram until closed loops have been shown, stopping at the point where the diagram is both significant and still easily understood. If too many factors and arrows





FIG. 9. A simple, yet complete causal-loop diagram.

are combined within one diagram, it may become merely confusing, and may obscure some of the most important and fundamental causalloops. The diagram in Figure 9 shows all of the factors presented in the cross-impact matrix, Figure 6, yet retains a simple and easily understood character.

5. To make it easy for others to understand this diagram (particularly those who may not have been involved in its development), it is desirable to replace the symbols (A, B, etc.) with the names of the factors, and to omit the circles—the intent here is to use the fewest possible special symbols to display the important information that has been developed. At the same time, it seems to help if the diagram has as much 'symmetry' as possible and the arrows are curved to aid in suggesting the 'loop' nature of the influences. Figure 10 was developed only after considerable effort to make it easily understood by a variety of persons who were not involved in its preparation.

#### AN EXAMPLE OF A CAUSAL-LOOP ANALYSIS

One example has already been given, showing how the information in a cross-impact matrix was modified for a causal-loop diagram. This second example will illustrate how a causal loop was developed in another study, without the construction of a cross-impact matrix.

A college was faced with a shortage of financial reserves, and was forced to reduce its support of some internal activities. Should it reduce expenses for instruction, administration or research? There were strong arguments both for and against each possible plan for the re-allocation of scarce resources. A three-phase study was conducted to (a) gather data from the four college deans, eight department chairpersons and eight senior faculty members, (b) construct a model that attempted to show the most commonly perceived structure of the college as it influenced the effect of financial



FIG. 10. A complete causal-loop diagram, adjusted for ease of interpretation by others who had not been involved in its preparation.

allocations on administration, instruction and research, and (c) manipulate this model in ways that might suggest the probable impact of alternative plans for readjusting the flow of resources within the college. A simplified version of the activities and outcomes of this study is given below.

- 1. A small planning group conducted carefully structured interviews with the twenty participants (deans, chairpersons and senior faculty) to gather data on their perceptions of the goals and characteristics of the college.
- 2. The interview data were analysed, categorized and simplified in order to identify the factors considered the most critical within the college, as it responded to problems, while working toward its goals.
- 3. Each respondent was given feedback on these most commonly perceived factors, and given the chance to change or add to the list.
- 4. Because most had indicated that 'competence of graduates' was the central concern of the college, each respondent was asked to place a small circle (labelled 'Comp. Grads') in the centre of a large sheet of paper. Considering 'Comp. Grads' as a quantity that could increase and decrease, each respondent was asked to identify factors from the list that would act directly on 'Comp. Grads' to cause it to rise or fall, and to show those factors on the diagram as other circles with arrows leading to 'Comp. Grads'. Similarly, if 'Comp. Grads' was considered as a factor that influenced ('caused') other factors to increase or decrease, the respondent was asked to show those other factors as circles with arrows leading from 'Comp. Grads' to those factors. Thus, every factor was considered as both a 'cause' and an 'effect', in order to trace out a number of cause-effect relationships and loops.
- 5. The study team examined all of the respondents' diagrams, identified the most commonly perceived structures, and prepared a causal-loop diagram that included the respondents' perceptions in a symmetrical and meaningful manner. In this process, three successive causal-loop diagrams were prepared, each a little more complex than the earlier one. The diagram in Figure 11 was the most common core of all respondents' drawings, but appeared too simple to represent the complexities within the college. Figure 12 added a few details to the basic structure of Figure 11, yet still seemed too simple. Figure 13 added still more complexity to Figures 11 and 12, seemed to represent most of the interactions involved in the college's dynamic operations, and appeared to be at the limit of the study team's ability to analyse.
- 6. This causal-loop diagram of the school was then used in two ways: (a) to give feedback to the participants and encourage them to explore the implications of this model, and (b) to provide the basis for a dynamic model, such as will be discussed later in this chapter. Unfortunately,



FIG. 11. A simple causal-loop diagram of dynamic interactions within a college.



FIG. 12. An expanded causal-loop diagram of the dynamics within a college.



FIG. 13. A complex causal-loop diagram of the factors interacting within a college, relating the allocation of resources to the quality of programmes.

time pressures acting upon the study team led to a reduction of their efforts on the earlier stages of this modelling procedure (the formulation and manipulation of prose and structural models), and an overemphasis on the preparation of the dynamic model. This unbalanced effort resulted in an excellent dynamic model, but the participants were not sufficiently involved in the process to understand, accept and use it. This experience served to re-emphasize the importance of the overview of the entire modelling process as suggested by Figure 1, and the dependence of each successive stage upon a thorough development and manipulation of the models in each of the earlier stages.

Advantages of this causal-loop approach are based upon its use of relatively common symbols (names and arrows) to represent some rather unusual and complex characteristics of a problematique (a problem embedded in a social environment). When used appropriately, and based upon a foundation of good model-building at each of the earlier stages (see Fig. 1), the production of a balanced and symmetrical causal-loop diagram is often very satisfying to the participants, because it seems to express so much in such a compact and easily understood diagram.

Difficulties with this approach are the same as for the other techniques noted above, except that most people can interpret a well laid-out causalloop diagram more quickly and more completely than either a verbal description or a matrix. These basic difficulties are restated below, just as a continuing reminder of their potential influence:

- Participant-centered difficulties associated with the personal, political and social characteristics of the participants. In some situations, it has even been apparent that a deep resonant voice or a high degree of personal charm have added undeserved weight to that person's suggestions.
- Technique-centered difficulties associated with the trial-and-error nature of the process for constructing an easily understood and symmetrically balanced causal loop. This seems to require repeated trials and the ability to discard early efforts that may have led to satisfying results, in order to achieve an overall sense of flow and movement.
- Further limitations of this approach and type of diagram include (a) its inability to identify and display time delays between actions and reactions, (b) its inability to examine and display the nature of relationship between action-cause and its reaction-effect, (c) its inability to specify the manner in which several 'action-causes' combine when they act together on one 'reaction-effect', and (d) the severe limitations of the number of factors and relationships that can be included in a single study or in one understandable diagram.

### MANIPULATING STRUCTURAL MODELS

Structural models can be manipulated in much the same way as prose models: different groups can be asked to prepare alternative structures that can then be compared, and the basis for choosing among them can be developed.

A simple example of this kind of manipulation was demonstrated in a community's effort to improve relationships between the school system and the community. Because of some strong differences in the value systems of various groups, three different scenarios and causal-loop diagrams were prepared: one by the administration of the school system, one by the dominant ethnic group in the community, and the third by the minority ethnic group. Each group had quite different perceptions of the goals for the school, the factors that interacted to move the overall community-school system towards these goals, and the criteria that should be used in

making decisions on the allocation of the school's resources. Prior to the development of multiple models, each of the three groups felt wholly frustrated by the others' lack of understanding of the 'true' situation. Thanks to the strong desire of the top school administrator for mutual understanding, the groups were able to develop and distribute their models, to try to understand the others' perceptions, and to work towards a set of policies on the allocation of the school's scarce resources that was at least partially acceptable.

While it is possible to manipulate structural models to estimate the probable effect of changes in one part of the system upon other parts, this 'manipulation' is a mental activity of the analyst. The structural model itself makes only minimum contributions to this process because it is 'static' and cannot operate independently of the analyst to demonstrate the effect of changes upon its performance.

The final type of 'model' to be considered in this chapter will focus on two types of 'dynamic' model that can be set in motion and will operate on their own. The next section of this chapter will describe and illustrate such dynamic models.

# III: Dynamic models

Both prose and structural models are 'passive', in the sense that they cannot operate or act independently to demonstrate the effect of the system's elements and structure on its performance over repeated time periods. Dynamic models can be prepared and set into motion, in order to observe their performance over time. In the discussion that follows, two types of dynamic model will be presented: (a) computer models that operate within a computer, and for which the computer can print out a record of performance over as long a period of time as desired, and (b) human simulations in which individuals interact within prescribed roles and constraints in order to experience and demonstrate the effect of those roles and rules on their joint performances.

# COMPUTER MODELS

The essential characteristic of this modelling approach is the specification of the characteristics of a model with sufficient completeness and precision that the model will 'run' (operate) and display operation characteristics that duplicate (simulate) the performance of the problem situation (the problematique). An incomplete model will not run, thus demonstrating the presence of theoretical gaps that must be filled before the model is whole, internally consistent and balanced. An imprecise model cannot be 'understood' by the computer, thus demonstrating the need for unambiguous descriptions of the model's components and relationships.

The technique will be described below, in terms of the specific actions taken to prepare a dynamic model of the educational materials development centre (EMDC) that was presented above in the discussion of cross-impact matrices and causal-loop diagrams.

A prose model was developed with the co-operation of the major decision-makers who were concerned with the EMDC. Following techniques suggested in our section on prose models, a list of the significant factors involved and a description of their interrelationships was prepared.<sup>1</sup>

The prose model was translated into a structural model. Cross-impact matrix and causal-loop diagrams were very useful in this process. These generalized forms needed to be refined into a special type of structural model as a step in the preparation of the computer statements that constituted the dynamic model the computer could interpret.

When modelling the EMDC, the causal-loop diagram was expanded into a dynamic model that used the special computer 'language' of Dynamo. This required the initial preparation of a structural model using the special symbols of the engineering field from which Dynamo evolved. This is illustrated in Figure 14, where the overall pattern of the causal-loop (Fig. 10) is maintained, but the information is modified to show four types of quantity:

- Conserved quantities, which tend to endure over time unless some of that factor either flows into or out of what is stored. Such a conserved quantity is shown in the diagram as a rectangle (→□→).
- Flows of the conserved quantities, which are represented as a valve that controls the rate of flow along a solid line (□ ×).
- 3. Information, which is represented by dotted lines showing that data about conditions in one element of the model are taken to another part of the model, where they may be used to influence or control what is happening there (-----).
- Auxiliaries, where information is combined in some way (such as adding or multiplying two pieces of information) and this new, combined information is used for control purposes ((○)-->).
- I. There are advantages as well as difficulties associated with moving either quickly or slowly through this prose model phase of a study. The rapid completion of these first steps can be encouraging to the participants, giving them a feeling of achievement and progress—but may yield relatively superficial results. The best compromise may be to take on a relatively small problem in the beginning, move quickly through all of the phases of modelling, and return to the earlier stages whenever necessary to achieve greater precision or more detail. This approach is consistent with the general modelling motto: 'Start small and simple.'



FIG. 14. Translating a causal-loop diagram into a structural model using Dynamo symbols (simplified to illustrate only basic structure).

Knowing the meaning of these symbols may help in seeing the basic similarities between the two diagrams in Figure 14.

The information contained in the structural model was then expressed in a set of Dynamo statements (using the special 'language' of Dynamo, which the computer can interpret).

The Dynamo statements were entered into the appropriate computer, and the model was 'run'. While this is easy to say, it seldom happens quite this easily. Computers are very demanding aids, and must be given information in carefully thought-out and precisely stated form. Few of us who prepare computer models are ever able to write a good set of statements (the model) the first time we try. Usually, this process requires that we go back to the prose and structural models and refine them, before we are able to get the dynamic model to perform the way we think it should, to represent (simulate) the behaviour of the problem situation. However, the work and thinking required to prepare such a dynamic model often lead to new understanding (which enriches our mental models, as suggested by Line 9 in Figure 1), even before we manipulate the model.

When the basic dynamic model performed adequately, it was manipulated in order to find the place (or places) where small changes would lead the model to perform in desirable ways. In this way, several different but possible policies, decisions or actions were tested on the model to observe their effects on performance before those changes were applied to the realworld problematique. Experimenting with a good dynamic model is usually a much faster and cheaper way to test the probable impact of alternative decisions than actually acting on them in the real world and then trying to observe the effect of our decisions and actions.

#### AN EXAMPLE OF A DYNAMO COMPUTER MODEL

Building upon the prose and structural models that have already been presented, from the study of the educational materials development centre (EMDC) in a late-developing country, this section will present and discuss the performance of that dynamic model under three different conditions. It is also appropriate to indicate that the responsible decision-makers in the EMDC have taken steps to modify the EMDC in directions that tend to agree with those changes that produced desirable results in the model.

First, the basic computer model (developed from the diagram in Figure 14) was operated and found to lead to 'failure'. This can be seen in Figure 15, where the reputation of the EMDC was slowly dropping and the amount of money allocated to the budget was gradually reduced by the government decision-makers until the EMDC was no longer able to continue operation. During the preparation of this basic model, it was found

that approximately 95 per cent of the total resources of the EMDC was devoted to the preparation and broadcasting of educational television programmes, and only about 5 per cent of the resources allocated to customer and field services. It was agreed that the first changes in the model would focus on a redistribution of the resources between these two complementary functions.

Secondly, the same model was run with a different distribution of effort between broadcasting and service, and the computer results showed that the model's performance tended to stabilize, as shown in Figure 16. Although this represented an improvement over the performance of the basic model, further modifications and manipulations were attempted.

Thirdly, the effect of various types of education and training programme was explored, with small amounts of professional and managerial time being devoted to training. A combination of both changing the emphasis on broadcasting versus service and new training programmes produced the dramatic effects indicated in Figure 17, where both reputation and budget allocations rose significantly.

Further revisions in the model were proposed, because it was recognized that the initial model was very simple and had been developed very quickly by only a small group of people. But these further refinements were not made. The dramatic changes introduced into the EMDC during the months after the simple model was prepared suggested that the very scarce planning resources should be directed at facilitating the changes that were being introduced, rather than investigating the possible effects of other innovations.

Advantages of computer-based modelling arise from the extraordinary demands made upon the analyst-modeller and the participating decisionmakers for completeness and precision. Many assumptions that were hidden during the preparation of the prose and structural models had to be made public and stated with precision. For example, it had been assumed that the EMDC contributed to student learning in the school through a combination of the materials delivered to the schools and the various kinds of service provided for them. But just how did those two complementary influences combine to produce their joint effect? Were they additive, with equal weighting? Did they combine multiplicatively, so that the absence of one would completely neutralize any contributions from the other? Although there was little hard data to guide anyone in making this modelling decision, it had to be made specific and precise in order to permit the model to demonstrate the implications of our perceptions. Out of this difficult, detailed exploration of one's perceptions and theories (beliefs about how people learn from different influences in their environments), came both a new respect for the complexity of the system with which we



FIG. 15. The performance of a dynamic model indicating the probable outcomes of present trends in an EMDC serving a developing country.





FIG. 16. The personnel in the EMDC were shifted to provide a 40-per-cent emphasis on customer services.



FIG. 17. The resources of the EMDC were reallocated to provide both (a) an increased emphasis on customer services and (b) increased education of both professional and management personnel.

were working and a much clearer view of the probable impact of the everyday actions of the people within the EMDC. It was then quite apparent that the different groups of people simply must co-ordinate their actions in order that they impact together upon the schools and government decision-makers, for materials without training were worthless, just as training without appropriate instructional resources was also worthless.

Difficulties with this approach must be carefully considered and balanced against the possible benefits, before beginning a project to prepare a computer-based dynamic model. These difficulties include the following: Large and sophisticated computers are required, with the analyst/modeller

having easy access to the machine and the special program needed to enable the computer to interpret Dynamo statements. Although these are relatively common in large universities, large business and government planning offices, they are not immediately available in every location (although recent advances in the use of remote terminals that can connect to computers over normal telephone lines are making them much more available).

- Skilled personnel who can prepare the required computer statements and interpret the computer print-outs are relatively scarce. This difficulty is usually dealt with in one of three ways: by hiring a person with the skills and placing that person in a centralized planning unit; by sending a person from the planning unit to an appropriate centre or university for training; or by asking a skilled consultant or university centre to assist in the preparation and use of the model. None of these approaches should be taken without careful consideration of the alternatives and the implications of each alternative.
- Decision-makers' understanding of models and the meaning of data obtained from the models is involved. Dynamic computer models have an aura of mystery and magic that needs to be dealt with. The purpose of dynamic models is to further enrich the mental models of the responsible decision-makers (see Fig. 1 again), and considerable exposure to these models and assistance in interpreting them are needed before they are really useful to decision-makers. But the effort is well worth the trouble. It is generally recommended that the responsible persons be continuously involved in all of the steps indicated in Figure 1, in order that all of these kinds of model serve their intended purposes.

#### HUMAN SIMULATIONS

In this approach to modelling, people are involved in a simplified and simulated situation where they can experience the effect of carefully designed roles and rules over relatively short periods of time. This technique is deliberately presented as the last modelling approach in this chapter because it is recommended as one of the final steps in any modelling effort. The responsible decision-makers need to experience personally the scenarios, structures, and dynamics that have grown out of their thoughts and perceptions but that also need to be experienced (with appropriate thinking, feeling and acting) in order to develop the richest possible insight or mental model of the problematique.

The technique will be described below in very brief form, and will be illustrated in a later section.

An elaborate prose model should be developed, describing the groups and individuals who are involved in or who influenced the problematique. Initial efforts in constructing a prose model may deal with a selected number of decision-makers or knowledgeable persons, but later efforts should lead to field interviews and observations by the analysts/modellers in order to get a full sense of the complexities in the situation.

A structural model should be constructed in order to simplify and organize the basic factors identified and described in the prose models. This organized insight is an important foundation for the construction of an appropriate and effective simulation, and inadequacies at this point may be seriously limiting later.

A dynamic computer model should be developed and operated—even a very simple one—in order to gain assurance that the model being developed has the potential for describing and explaining the most important behaviours of the problematique. This step is very seldom taken. Most simulations are designed and played by persons who are seeking some form of general education, rather than to understand and take action to change some social problem. Here, we are not primarily concerned with educational goals and the means to achieve them, but rather we are focusing on aids to decision-making relative to serious social problems. From this point of view, the preparation of a dynamic computer model can provide additional data to support the validity of the conceptual base on which the human simulation is to be built.

The roles and rules for the simulation, and all of the supporting materials, facilities and personnel, are then specified. This is the stage of simulation design; and there is a delightful mixture of science and art, of discipline and drama, involved in the design of a simulation. An effective background for those who design simulations includes varied experience with different kinds of simulation prepared by others, exposure to the writings of those who have studied and taught simulation, modest skill in writing, a controlled flair for the dramatic, and a willingness to prepare something, to try it out on a group of people, and to modify it until it
simulates the critical characteristics of the problematique. In this way, a simulation is the same as a computer model because both must be tested and revised until they perform appropriately. Some specific suggestions on the design of a simulation would include the following:

- 1. Identify the key agencies and persons within each agency, carefully describe the actions which each agency and person take with respect to the problem, and write each set of actions in a separate 'job description' to be given as a role prescription to each person in the simulation.
- 2. Identify the important resources and limitations within which each agency and person must act if they are to experience the same types of environmental features in the simulation as in real life. Describe these in a scenario for each player, and provide the resources and restrictions as part of the environment of the simulation.
- 3. Prepare a 'simulation leader's guide' defining very carefully the actions the leader must take to begin the simulation, to guide it through its various stages, and to bring the people together at the end to examine their own behaviours during the simulation, and make suggestions for making the simulation more realistic.

Alternative sets of roles, rules and environmental resources and constraints can then be designed into subsequent simulations to experiment with the effect of alternative dynamic models (other systems structures) on the performance of the individuals and the entire system. During these manipusations of the model, it will be important to measure the critical behaviours so that the differences introduced by changing roles and rules can be evaluated.

#### AN EXAMPLE OF A DYNAMIC SIMULATION MODEL

The national agency concerned with the education of handicapped children had supported the development of a network of agencies to conduct research, develop and distribute educational materials that had been shown to be effective with specific types of handicap, and train teachers to use these materials and appropriate methods with particular children. Unfortunately, this network was not as effective as had been hoped, as indicated by the small number of materials and methods that were actually reaching the teachers in their classrooms.

In an effort to understand the dynamics of this system, as one step towards its improvement, a simulation model of this network was prepared. The key persons in the network were interviewed, the designer/consultant visited typical agencies throughout the network, and preliminary versions of the simulation were tested on persons who were uninformed about the network as well as others who were intimately familiar with it. Each

#### Modelling and simulation

participant was assigned to one of the typical agencies in the network: the national bureau for the handicapped; a university research and development centre; the national professional agency for educators of the handicapped; regional centres for the distribution of materials and the training of teachers; publishers of the educational materials; and the public schools (state as well as municipal) in which handicapped children were educated.

Each participant was given a notebook that provided the background for the simulation (the scenario developed early in the design process), and was assigned a specific title and role, with a number of required, permitted and forbidden actions. Physical facilities were arranged to simulate the difficulties associated with travel from one agency to another. Communications between persons and agencies were controlled. Tight time schedules were imposed for the completion of simulated tasks. Unexpected emergencies were introduced into the simulation at irregular intervals. Simulated money was provided, simulated materials were developed and sold with the hope that they would eventually reach the teachers, and certificates were awarded those who attended special training sessions. At the end of the simulation (after a full day of interaction), the behaviour of the participants was examined and the simulated materials and certificates for the acquisition of new skills were counted. It was generally agreed that the simulation replicated the critical performances of the real-world situation; everyone was very busy and very emotionally involved, but almost no materials and skills filtered down to the teachers. In the post-simulation discussion, many suggestions were made on changes that might be made in both the model and in the real network.

The simulation was then modified to produce three different versions, each representing a different proposal for the improvement of the network's performance: improved communication among the agencies; a network co-ordinator with responsibility to keep track of the flow of materials and skills, and the authority to expedite their movement; and a combination of both improved communication and a network co-ordinator. Each version of the simulation was conducted with groups of educators—but, unfortunately the types of person involved changed drastically between the successive simulations, and the highest level policy-makers were transferred to other projects or were distracted to other tasks, and therefore the final assessments were inconclusive about the effect of the different model structures on the simulation's performance (because too many other things had also changed).

This experience appeared to demonstrate that such simulation can be developed and conducted for the analysis of large and complex social problems, but that these activities require the continuing interest and support of the responsible persons. Advantages of human simulation as dynamic models come primarily from the effect of this experience upon those who become involved. Participating in a simulation is a vigorous and stimulating activity. When adequate preparation has been made (the participants come to the experience prepared, and are given a good introduction to the procedures and purposes), the experiences can stimulate a high level of motivation. During one experience with the simulation model described above (the network of centres for the handicapped), the thirty-five participants became so involved with the experience that they would not stop after eight hours but insisted upon continuing another two hours in order to see if they could achieve the goals they had set for themselves.

A second advantage that can come from a simulation arises out of the post-simulation discussion. Having experienced the impact of the simulation's rules and constraints on their progress towards the system's goals, the participants often come to new insights about ways of improving both the simulation model and the real situation that lay behind the model. This kind of a discussion acts as both a training device for the participants and a 'laboratory' from which both the analyst/modeller and the decisionmaker can learn.

Both of these advantages (motivational and educational) can be achieved in a very short time (a simulation usually lasts no more than from two to thirty-six hours) with an apparently small investment (only the paper needed to define each participant's role, and the minimum resources of a room and a few 'things' that are manipulated during the simulation). But...but...but...this simplicity tends to hide the long and demanding preparation that should come before the simulation (although some people do not go through all the modelling stages that have been described and recommended here).

Perhaps the most important advantage of a human simulation after all of the other kinds of model have been prepared and manipulated is the validation that a group of knowledgeable and experienced persons can provide through participating in one or more simulations that demonstrate the characteristics and implications of various models of some problematique in which they themselves are involved. For example, the managers of the EMDC (see the example of a cross-impact matrix in Section II, pages 131 to 133) could have been asked to participate in a human simulation of the three different structures that evolved from that study: the basic model of the structure that led to the simulated collapse of the EMDC, the model with a balanced effort between production and service, and the model that combined the production-service balance and the educational programmes. Their response could have indicated that the basic model was a realistic representation of the actual system in which they were daily involved—and the model could have been adjusted until the desired support was given. They could then have participated in testing the validity of each proposed change in the model. In the process, they might have learned a great deal about their own system, and might be emotionally prepared for the changes in responsibilities resulting from top management decisions to implement these new policies.

Difficulties encountered in designing and conducting dynamic human simulations include the following.

Special skills are needed to design and conduct this kind of group experience so that it is both meaningful to the participants and valid as a model of the problematique it is intended to simulate. Such skills are not common. Many designers of educational simulation-games would have relevant skills, but the two tasks are not exactly the same; the designer of this kind of dynamic model would need the skills necessary to design and analyse the earlier kinds of model that should precede the analysis of this more complex model.

The decision-makers' time and attention may be difficult to obtain. The most successful simulations of this type have been conducted when the highest levels of policy-makers are strongly interested and also wish to involve the other key persons in the system.

The evaluation of the initial problematique, and then the conduct of related evaluations before, during and after the simulation, can be a difficulty. Special skills are needed to conduct such careful evaluations in both real and simulated conditions. It is possible and desirable to obtain such skills either through (a) hiring persons with these skills, (b) sending individuals from a planning unit away to acquire these skills, or (c) employing agencies or consultants to provide these competencies as they are needed.

#### MANIPULATING A DYNAMIC MODEL

Several examples have already been given of ways of manipulating the dynamic models that have been described in the earlier sections. The essential characteristics of these manipulations have been:

Sensitivity analysis, or the effort to find the one or more places in a model where small changes are likely, in the long run, to result in desired changes throughout the entire network or problematique. A computerbased model is an ideal mechanism for a sensitivity analysis, because successive runs of a computer model can be achieved quickly and inexpensively (although finding the best places to make such changes and determining their direction and magnitude are neither easy nor inexpensive in terms of the mental effort of the participants). Structural analysis, or the exploration of the performance of alternative models of the same problematique. This kind of analysis goes beyond a 'sensitivity' analysis, which considers the effect of small changes within the same structure, and tries to examine the effect of changes in the structure itself.

As has been suggested repeatedly in these pages, the purpose of all of these manipulations is the enrichment of the mental models of the decisionmakers who are responsible for the policies, decisions and actions that will eventually have impact upon the real-life problematique. Their involvement in these activities will be a critical factor influencing their willingness to act upon the outcomes of any modelling effort.

#### IV: Summary and conclusions

Social problems are normally embedded in complex environments within which many factors interact to influence the behaviour of the components of the system (the problem as well as the desired outcomes) and the overall system itself. The complexities involved are usually beyond the capabilities of the human mind to understand in sufficient detail and precision so that correct actions have a high probability of being effective. Three different types of model can provide significant assistance to human decision-makers as they attempt to set policies, make decisions and take actions to solve problems and improve situations. These types of model are:

- Prose models, which use words (written or spoken) to express one's understanding of the factors and relationships that influence the problem.
- Structural models, which build upon the prose models and represent the system's components and relationships in the form of diagrams or other non-verbal symbols.
- Dynamic models, which build upon the structural models to achieve an operating, performing model that demonstrates the critical behaviour of the problem situation.

Each of these types of model has strengths and limitations, and requires particular skills to prepare, conduct and analyse. But the position is taken here that social problems faced by individuals, organizations and nations are sufficiently important and complex to justify both the time and skills required to prepare and use models. Those who acquire and use these new ways of approaching and understanding problems and social systems can make increasingly significant contributions to their countries and cultures.

# III

The technological approach

### Chapter 5

# The elements of communication technology

Bernard Webster

The rapid growth and spread of communication technologies pose unique problems and opportunities for planned communication systems. Technology of various kinds is available for a wide range of purposes, from village-level use of video-tape recorders to capture problems and feedback on development programmes to development of telecommunications systems for national-level development. Communication planners must understand the components of various technologies, and how technologies function, in order to develop and evaluate alternative technological solutions to communication problems. In fact, the element of technology—in particular large-scale electronic telecommunication technology—lies near the top of agendas in most communication planning situations.

In this chapter the basic components of communication technology will be defined, followed by discussion of important current trends in technology development and of international communication agencies. These discussions in turn form the basis for an analysis of various aspects of planning for communication technology, with special attention to the value of reviewing alternative technological systems for given planned purposes.

#### **Basic definitions**

The basic communication act consists of a message or signal emanating from a source and reaching a receiver. The message may vary in complexity from a single word 'Go' or its codified equivalent, a green light, to the complete works of Shakespeare or the data stored in a large computer. The message may pass to the receiver directly by means of sound or electromagnetic waves or indirectly by the physical movement of some medium on which it is recorded, e.g. paper, film or magnetic tape. The receiver may take many forms, ranging from a human ear to a television receiver or a computer.

#### COMMUNICATION SIGNALS

Analogue signals are electrical analogies of some physical variable. For example, when someone speaks into a microphone the instantaneous amplitude of the output voltage from the microphone varies with time in sympathy with the varying pressure produced by the speech wave.

A television camera also produces an analogue signal. In this case the instantaneous amplitude of the voltage or current corresponds with the instantaneous brightness of the picture at each point along one of its lines that is being rapidly scanned by an electron beam.

However complex an analogue waveform may be, it can be shown to be made up of a number of harmonically related pure sinusoidal waves of various amplitudes. An important characteristic of a sinusoidal wave is its frequency, i.e. the number of complete cycles occurring in one second.

Frequency is measured in cycles per second, which are called hertz. For convenience the following multiples are used: Hz (hertz), cycles per second; kHz (kilohertz), thousand cycles per second; mHz (megahertz), million cycles per second; and gHz (gigahertz), thousand million cycles per second.

Digital signals do not vary in sympathy with some message source but are a codified representation of it.

A series of on-off pulses of voltage or current can be arranged to represent letters or numbers. A message consisting of many letters or numbers can then be represented by a train of on-off pulses. A teleprinter is a typical example of a source producing signals in this form.

Each on or off signal is called a 'bit' or BInary uniT. An important characteristic of a digital signal is its bit-rate, i.e. the number of pulses per second. Similar multiples are used as for frequencies: bit/s, bits per second; kbit/s, thousand bits per second; mbit/s, million bits per second; and gbit/s, thousand million bits per second.

Digital techniques (i.e. the use of on-off pulse trains to represent any given signal) are not confined to machines such as teleprinters that produce the messages directly in this form. They may also be used to represent speech or television signals. The amplitude of the signal is sampled (or measured) at each successive instant of time, normally at a rate of several thousand samples per second for speech and several million samples per second for television. Its amplitude level at each sampling point is then represented by a code of pulses, e.g. a code using eight on-off pulses can define  $2^8 = 256$  discrete amplitude levels. These are usually sufficient to give an acceptable reproduction.

The bandwidth is the width of the band of frequencies that is needed to obtain a sufficiently accurate representation of the signal. In the case

Signal	Bit-rate (digital)	Bandwidth (analogue)
Television		
(625-line colour)	100 mbit/s	5.5 mHz
Music		15.0 kHz
Telephone speech	64 kbit/s	3.0 kHz
Data (medium speed)	48 kbit/s	<u> </u>
Data (low speed)	600 bit/s	
Teleprinter (telex)	50 bit/s	

TABLE I. Examples of signals

of analogue signals this depends upon the highest frequency present. Similarly, in the case of digital signals it is the required bit-rate needed to obtain a sufficiently accurate representation of the signal that is of the greatest importance.

Table 1 shows a few examples of signals and serves to draw attention to the important fact that the bandwidth of a television signal is more than a thousand times greater than that of telephone speech.

#### CARRIER WAVES AND MODULATION

Data, sound and television signals can all be directly recorded or sent live along cables to receivers. However, when it is desired to send such signals, whether in analogue or digital form, over long distances either by cables or by radiation, it becomes necessary to use a carrier wave. The carrier wave is a high-frequency wave onto which the relatively lower-frequency signal is impressed. This process is called modulation, and the reverse process carried out at the receiver whereby the original signal is extracted from the carrier wave is called demodulation. There are many different types of modulation and several methods of achieving each type.

#### TRANSMISSION AND DISTRIBUTION

The process of moving a signal from its source to a receiver can be divided into two main forms:

- Direct by electromagnetic radiation travelling at the speed of light (300 million metres per second) or very slightly slower by guided waves travelling along wires, cables, tube-like waveguides or glass fibres. For all practical purposes within a country all such transmissions can be considered to be instantaneous.
- Indirect by the physical movement of some form of recording. All signals can be recorded in some way. There are many recording methods and

many different recording media. Magnetic recording methods can be used with discs, drums, tapes of various widths, stripes on films, etc. Paper as an example of a recording medium can be used for recording a message by writing, printing, punching holes, etc. In every case time is taken first in making the recording of the message and then moving it from the source to the receiver by some means of physical transportation.

#### Direct transmission and distribution

There are two primary forms of direct transmission and distribution of signals—broadcasting and the use of 'guided' waves.

#### **Broadcasting**

Broadcast signals are distributed through transmitters located on the ground (terrestrial transmitters) and through a variety of technologies in which the transmitter is located above the ground to increase the range of broadcast (non-terrestrial transmitters).

Terrestrial transmitters fall into four categories, defined by the wavelength of broadcasting:

- Long wave: Very long range when high powered but only a limited number of frequencies are available. Only suitable for radio.
- Medium wave: Hundreds in use for radio. At night these waves are reflected by the ionosphere and strong signals can be received at a range of hundreds of kilometres. This can lead to interference problems.
- Short wave: These include long-distance radio communication systems such as ship-to-shore, as well as the amateur radio bands. They also include VHF and UHF television systems and radio navigation systems. The VHF (Very High Frequencies, 30-300 mHz) bands provide scope for many high quality radio stations and a fair number of television channels (thirteen are used in Europe). The UHF (Ultra High Frequencies, 300-3,000 mHz) bands are widely used for television, although in hilly terrain their range is severely limited.
- Microwave: These include radar transmitters using mainly the SHF (Super High Frequencies, above 3,000 mHz). They can be used for radio and television broadcasting, but when so used their range is strictly line-of-sight and they are considerably affected by atmospheric conditions.

Non-terrestrial transmitters include:

Aircraft: In the United States in the 1960s two high-powered transmitters were lifted in an aircraft that circled at 7,000 metres and provided educational television for an area of more than 326,000 km<sup>2</sup>.

- Aerostat platforms: A tethered balloon about the same size as a Boeing 747 aircraft is capable of supporting a large transmitter. This system has been tried experimentally in several countries and permanent systems are under development.
- Satellites: At present the power of a satellite-borne transmitter is limited by both the size of the launch vehicle and the capacity of the solarpowered electricity supply system. This means that when broadcasting to a whole country all of the receiving points need special dish aerials. This limitation will eventually disappear but only by shifting the costs from the receivers to the transmitter.

Another possibility with both satellites and aerostats is the use of a passive reflector driven from a powerful transmitter on the ground.

#### Guided waves

The quantity of signals that builds up over the paths between the main centres or nodes of a communication network leads to a distinction being made between transmission and local distribution. Transmission is equivalent to the motorways between cities, while local distribution is equivalent to the side streets of the cities.

The bandwidth and therefore signal-carrying capacity of different guided-wave systems varies greatly. It is therefore just as important to create a system whose capacity is appropriate for the anticipated demand as it is to build a road system appropriate to the anticipated traffic density.

The major types of guided wave system are:

- Lines and cables: Overhead lines and underground cables, in which there are a thousand or more pairs of wires all insulated from each other, are in extensive use for telephone and data distribution.
- Radio links: These have been in use for many years, mainly at very high frequencies, for crossing tracts of water or difficult terrain. By moving up into the microwave band, signals can be concentrated into narrow beams by means of dish or horn aerials, thus conserving energy. The bandwidth at these frequencies is sufficient to carry several television signals or thousands of telephone circuits. Exactly the same technique can be used with the aid of a satellite relay point to extend the links over vast distances.
- Coaxial cables: A single cable consists of a copper inner conductor insulated from an outer copper tube of diameter 2 to 8 mm. The larger cables are capable of providing bandwidths of up to 500 mHz. Several such cables may be bound together and are suitable for high-capacity intercity transmission paths for telephones, data and television.

Waveguides: These are used for long-distance transmission and are capable

of transmitting a wide band of microwaves sufficient to carry some 300,000 telephone channels or 250 both-way television circuits.

Glass fibres: As used for transmission and distribution systems these are comparatively new but under rapid development. Such systems use hair-thin fibres of extremely pure glass capable of transmitting light waves (which are the same type of wave as radio but of much shorter wave-lengths) over distances of several kilometres before the signals need to be reamplified. Each fibre can carry more than a thousand telephone channels and a cable of a hundred fibres is only I cm in diameter and is quite flexible.

#### Indirect transmission and distribution

An examination of the history of communication methods in which a copy of the recorded message is physically moved from source to receiver shows that the process is very dependent upon access routes. When many messages were being sent to the same places local postal services developed. Then gradually a distinction began to be made between intercity 'transmission' and local 'distribution'. Trunk roads, railways and airlines shortened the time taken on transmission paths. At intermediate stages of evolution mixed systems emerged, such as the telegram transmitted by wire to the nearest city and distributed by a boy on a bicycle to the individual receiver. A more recent mixed system involves the sending of complete copies of newspapers over telephone facsimile systems to provincial offices, where they are reproduced for local distribution.

Time, cost and convenience emerge as the factors determining the choice of transmission methods.

Currently a video recording of a television programme may be physically moved to another broadcasting station that wishes to use it or may be sent as a signal via cables or radio links, including satellites. Similarly, a computer program may be either physically moved on a magnetic disc or sent over a data link circuit. There can be little doubt that if the direct links were readily available and cheap enough, they would be used for all movements of 'recorded' material even if there was no urgency for its use.

#### COMMUNICATION NETWORKS

The first subdivision of networks is into two basic types: (a) star or open networks and (b) mesh or closed networks.

A typical national telephone system is likely to start as a star network gradually extending from the major to the minor population centres as in Figure 1.





FIG. 2. Mesh network.



FIG. 3. Simplified mesh network.



This arrangement provides no alternative routes, and a break in any one of the intercity links will cause a major disruption of the total service. Wherever possible communication engineers will aim to convert any star system to a mesh in order to create alternative routes that can act as stand-bys in case of breakdowns.

An arrangement such as that shown in Figure 2 provides two stand-by paths for every city. While such an arrangement may be technically ideal, it would normally be uneconomic. A simpler mesh arrangement, as in Figure 3, would be sufficient to provide one stand-by route for each centre.

#### Multiplexing techniques in networks

The links between the population centre or nodes of a network may be made using any of the guided-wave methods described above. In order to increase efficiency and reduce costs it is necessary to maximize the number of signals that can be sent over each path. The process of sending two or more separate signals at the same time over a single communication path is called multiplexing. There are two main forms:

Frequency-division multiplexing (FDM): The available bandwidth is subdivided into several adjacent channels, each accommodating a separate signal, as in Figure 4.



#### FIG. 4. Frequency-division multiplexing.

Complete train of pulses passing along the path





Time-division multiplexing (TDM): In this arrangement the whole of the available path or system bandwidth is used to accommodate a single train of high-speed pulses. Individual communication channels are created by interleaving groups of pulses at regular intervals of time, as shown in Figure 5.

#### Switching techniques in networks

The ability to connect any given source to any chosen receiver both quickly and reliably through a switching system is an essential facility of any communication network. But this is often more difficult to achieve than the actual transmission of a signal from point to point. Switching techniques may range from a single switch used in selecting one from three channels on a television receiver to the hundreds of switches that may come into action when one telephone is connected to another of the 350 million that are available via the world network.

In the earliest telephone networks all switching was done manually by human operators. Although automatic switching is now accepted as the norm, operators are still retained in a variety of situations, although their role as advisers is generally more important than their involvement in the actual switching process. As networks grow and continue to bring in more people who are inexperienced in the use of the system there will continue to be a need for some human operators.

As networks grew the demand for switching automation was met by the very widespread use of the two-motion Strowger switch, an electromechanical device. The process of dialling a number generates series of pulses. The first digit dialled causes a radial contact arm to be stepped vertically through a number of levels equal to the number of pulses. The contact arm then hunts round a horizontal semi-circle of contacts to find an unoccupied second switch, when the process is repeated for the second digit dialled, and so on.

Despite its relative simplicity and low first cost, the Strowger switching system is now seen as more of a burden than an asset. Owing to the need for considerable maintenance, and above all to the relatively slow speed of operation, the exchanges using this system are being replaced by more modern arrangements. This, however, is not easy. The capital investment in Strowger exchanges in many countries is very great. In the meantime it is necessary to use a great deal of engineering skill to make new developments compatible with these existing systems.

Modern switching systems are either partly or fully electronic. The former include those that use matrices of fast-operating reed-relays or cross-bar switches. In major switching centres and exchanges the control equipment, which automatically operates all the various switching, routing and charging processes, often assumes the character of a computer. These technologies exploit the hardware concepts of computers through the use of electronic logic, integrated circuits and storage devices. They also employ software concepts, particularly for stored program control (in which programs of instructions written in standard computer languages are used to carry out a range of predetermined operations).

Trends in the development of communications technology

#### THE SEARCH FOR MORE BANDWIDTH

As the total quantity of information to be transmitted has increased, so also has the demand for frequency space in which to transmit it. The demand is always for more bandwidth, whether for direct transmission by broadcasting or by guided transmission along some predetermined path. As the lower frequencies in the electromagnetic wave spectrum are used up the required bandwidth for broadcasting can only be found by continually moving farther up the frequency spectrum. As the spectrum diagram below reminds us (see Fig. 6) electromagnetic waves, which all have the same characteristics except for their frequency or wavelength, are commonly divided into discrete bands. However, the waves of different bands are increasingly used for similar purposes.

Each increase in usable frequencies that has been won has required long, intensive and costly research and development, not only in the fields of propagation characteristics and transmission media, but also in the equally difficult areas of new types of amplifier, modulator, etc., to work at each higher frequency.

#### Very High Frequencies (VHF 30-300 mHz)

These are used both for direct broadcasting for radio and television and also for coaxial cable television systems of nine or more channels. Frequencies in the VHF band are also used for trunk telephone coaxial cable links. These use the Frequency-Division Multiplexing system, which allows groups of channels to be combined into progressively larger groups before transmission over the main routes.



FIG. 6. The electromagnetic wave spectrum.

165

#### Ultra High Frequency (UHF 300-3,000 mHz)

This band of frequencies is now extensively used for television broadcasting. For example, in the United Kingdom it provides a nation-wide four-channel capability, although owing to the limited range of these frequencies, hundreds of low-power repeater transmitters are needed to provide more than 96-per-cent coverage of the population.

#### Super High Frequencies (SHF, above 3,000 mHz)

Waves of these frequencies are also known as microwaves. Wavelengths range from several millimetres down to fractions of a millimetre. When used for open-air radiation their range is strictly line-of-sight, and they are considerably affected by atmospheric conditions, particularly rain. However, they can be guided by either rectangular or circular cross-section pipes called wave-guides. These pipes have extremely critical mechanical tolerances. Devising methods of making long lengths with sufficient accuracy has kept several teams of research and development engineers busy for the past few years. To achieve success, complex combinations of precision steel tubing with copper-wire helices and suitable insulating materials have had to be produced. In addition to work on the guides themselves, much effort has gone into the development of microwave generators and special transistors to work at these extremely high frequencies.

#### Visible light frequencies

At even higher frequencies the development of lasers, which are coherent light-emitting sources, and the means of modulating their output have opened up new possibilities of very wide band transmission paths. For example, a recently announced laser modulator, which incidentally is of extremely small size, is capable of modulating into a single laser beam of red light 25,000 radio or 20 television programmes. Since light beams can be completely cut by fog or snow, development has concentrated on guiding these light waves along optical fibre cables. By the use of ultrapure glass or low-loss liquids inside fine glass tubes the loss of the transmitted signal can now be kept very low. A complete cable having fifty conductors with all the necessary padding and protection is only 1 cm in diameter. It is much cheaper than a comparable cable using copper. It is mechanically flexible and therefore capable of being laid round bends and in existing ducts.

#### SMALLER, CHEAPER AND MORE RELIABLE EQUIPMENT

By the mid 1960s the use of the transistor in place of thermionic valves was producing extensive changes in the electronics and telecommunications industries. Products were becoming dramatically smaller, lighter and cheaper, while at the same time becoming more reliable and consuming less electrical energy.

The amazing fact is that this process is still continuing. The growing importance of digital methods replacing long-established analogue systems and the concurrent demand for higher speed circuits for use in digital computers have encouraged the development of integrated circuits. The techniques used in the development of printed circuit boards, now used instead of separate wiring for each circuit component, combined with the use of semiconducting materials from which transistors are constructed have made it possible to put thousands of complete circuits on to a tiny chip of silicon. This technique is known as large-scale integration (LSI). With careful design the distances between circuit elements can be kept down to a few millimeters using LSI. This produces very fast-acting circuits. These in turn can perform a large number of simple operations in a sequence that is equivalent to some more complex operation previously performed by an analogue circuit.

With such circuits it is quite easy to include arrangements that indicate when a fault occurs and also to have stand-by circuits. The result is that the mean time between equipment failures becomes very long. The nature of monitoring, servicing and maintenance tasks is thus markedly different from the days when large numbers of individual components were wired together to form a system.

#### COMPUTERS IN COMMUNICATION SYSTEMS

Over the last decade the computer and its digital circuit techniques have made many inroads into all branches of communication systems. A point has been reached where it is now debatable whether communication is a branch of computer technology or vice versa. Computers relate to communication systems in a number of ways.

#### As subscribers

The need for computers to 'talk' to each other has led to the establishment of data networks of very high bit-rate capacity, because modern computers with large memories may wish to 'say' a lot to each other in a very short time.

#### As controllers

In large complex systems computers are now being used to control operations such as switching, routing and charging in telephone exchanges. The development of intercontinental communications and national broadcasting via satellites would never have been possible without the use of computers to carry out essential tracking and controlling operations. In smaller, less complex systems the newer microprocessors (or special purpose minicomputers) are beginning to be used for similar controlling purposes.

#### By the introduction of new techniques

Many new types of circuit suitable for large-scale integration (LSI) have been developed for use in computers. These find application either directly, or with a minimum of modification, in a wide range of communication systems.

#### As generators or terminals

Early data links over telephone lines used teleprinters that were derivatives of electric typewriters. More recent techniques of producing and displaying messages used with line transmission, such as the British Post Office's Viewdata System, are derived from a combination of computer and television circuitry.

#### As editors

The first editing technique applied to a video tape was an exact copy of the method used for film—cut and join. Before long all-electronic techniques became the norm. More recently digital circuit techniques developed for computers have enabled each separate frame of the recorded television picture to be identified, and from there it has been a short step to using a special purpose computer to control all editing processes.

#### As a standards convertor

Perhaps the ultimate incursion of computer technology into the world of television is the Independent Broadcasting Authority's system for converting colour television signals back and forth between the 525-line 60-Hz American and the 625-line 50-Hz European standards. This is achieved by producing a purely numerical representation of one signal and then reprocessing it within a computer to produce a numerical representation of the same signal conforming to the other standards. The computer then provides an output that is the new signal converted to the normal analogue form.

#### As aids to planning, modelling and design

With the aid of computers a communication network of the future can be simulated and the effects of varying any number of parameters such as system quality, reliability and charges can be studied.

#### THE POTENTIAL OF LIMITED BANDWIDTH SYSTEMS

In the development of printing, line drawings and text were followed by monochrome and later by colour pictures. It is interesting to note that in the production of a book a publisher will make a careful choice from the whole spectrum of presentation techniques available to him. As a result he will not use photographs unless they are essential and the cost of including them can be fully justified. Line diagrams, cartoons and a variety of type faces all find a place alongside photographs in a single edition of a newspaper.

If we compare this situation with that of electronic communication by cable or broadcasting, we find that there are conspicuous gaps in the spectrum of presentation techniques that are only just starting to be filled. The enthusiasm for television has been the root cause of this uneven development. While there has been no shortage of experiment leading to laboratory models of many different communication devices, there has been very little commercial exploitation of the technically less demanding and therefore less expensive systems that can operate on bandwidths far less than those required for television.

If we look at direct transmission systems in most developed countries we find that the patterns of both radio and line transmission display a range of latent signal carrying capacity. By comparison with printing the equivalents of line drawings and cartoons are conspicuously absent.

# International bodies that influence national communication systems

All national communication systems will come directly or indirectly under the influence of a number of international organizations, which may recommend or in some cases require certain behaviour from national bodies. Participation in any satellite linkages will obviously require exact conformity to certain technical standards and codes of practice before any communication can take place. Less obvious but just as important will be technical compatibility between the nodes of any telephone cable or microwave links crossing national frontiers. Preventing serious interference between the radio transmissions of neighbouring countries by the careful allocation of and adherence to suitable frequencies is a prerequisite for successful communication.

Some of the more important international regulatory and collaborative organizations and their functions are indicated below.

The International Telecommunications Union (ITU) has a membership of 148 states. Its origins can be traced back to 1865. In 1947 it became one of the United Nations' Specialized Agencies. Its purposes are defined as:

To maintain and extend international co-operation for the improvement and rational use of telecommunications of all kinds; to promote the development of technical facilities and their most efficient operation with a view to improving the efficiency of telecommunication services, increasing their usefulness and making them, so far as possible, generally available to the public; and to harmonize the actions of nations in the attainment of those common ends.

The International Frequency Registration Board (IFRB) records assignments of radio frequencies throughout the world and advises the ITU on interference problems between stations.

The International Radio Consultative Committee (CCIR) and the International Telegraph and Telephone Consultative Committee (CCITT) both study technical, operating and tariff questions and issue recommendations on them.

In recent years the ITU has developed a role complementary to its traditional function as co-ordinator and standards-maker for international telecommunications services. This new role is specifically supportive of developing countries and is manifest in many different ways, ranging from the production of tutorial technical handbooks to assisting the creation of the Pan African Telecommunications Network.

#### National and regional broadcasting and communication organizations

Many regions of the world have well established organizations whose main purpose is to arrange both technically and administratively for the free flow of communication between member countries, states or constituent organizations. Each area of the world has its own form of organization, with some variation in status and terms of reference. The following are typical examples: United States of America: the Federal Communications Commission (FCC) and the National Association of Broadcasters (NAB).

- Europe: the European Broadcasting Union (EBU) and the Eastern European broadcasting union (OIRT).
- Middle East: the Arab States Broadcasting Union (ASBU).

Asia: the Asian Broadcasting Union (ABU).

Africa: the Pan African Telecommunications Network (PANAFTEL) and the Central African Group of Telecommunication Administrations (CAPTAC).

In 1925 the first organization, the International Broadcasting Union, was formed in Geneva. This was discontinued after the war, and the subdivision into continental organizations started with the formation of the EBU at a meeting in Torquay, United Kingdom, in 1950. The EBU now has some seventy broadcasting organizations from thirty-one countries in full membership and as many again from another forty-six countries in associate membership. The ABU was founded in 1964 and already has twenty-seven members and eighteen associates.

It is seldom easy for a developing country to release key personnel to play an active role in such co-operative ventures, but it is vital that planners recognize the importance of such activities and provide for a staffing level that makes them possible.

#### A national communications structure

Developed countries seldom provide good models for the structure of a national communication system. For various political, financial, technical and other matters of expediency long since lost in the pages of history, systems have evolved to the present stage and post hoc rationalization is used to justify the continuation of structures in which many people have vested interests. For example, in many ways it is quite illogical to put line and broadcast transmission under separate authorities. It is equally illogical to maintain that the production and distribution of broadcast programmes must be the responsibility of the same organization. However, many such arrangements exist and therefore tend to get copied by developing countries.

Before the level of capital invested, the staffing structures, and other factors lead to the creation of a rigid system with inbuilt resistance to major changes, it is timely for a developing country to review its developing communication system. This review should look to the present and probable future social impact of the system, and its efficiencies. The success of the system in meeting needs, its compatibility with developments external to the country, and its total costs and comparative cost-effectiveness are important factors. Some planning activities must occur at the national level. External links for both telecommunication and broadcasting must be provided. These may involve a selection from, or mixture of, lines, cables, radio links at various frequencies and satellite relays.

Frequency planning for national and local broadcasting, for telephone and data links (usually by microwaves) and for other services such as police, railways, pipelines, gas, water and electricity must be accomplished within the international regulation framework.

National broadcast coverage must be planned in response to social requirements. Depending upon the geographical features of the country, choices will have to be made between small numbers of high-powered terrestrial transmitters or larger numbers of lower-powered ones. Other alternatives may also be considered owing to advancing technology, e.g. the possibility of using a national satellite or large tethered balloons (aerostats), which are equivalent to low-level satellites and can carry a heavy electronic pay-load and may prove more economical.

Main telephone/data network nodes and links should be planned nationally. As we shall see below this planning should attempt to achieve a balance between technical and other requirements.

Co-operative arrangements between separate communications organizations are normally achieved only at the national level. Joint planning between broadcasters and telecommunications authorities may lead to considerable savings by the sharing of buildings, towers or, as is in Nigeria, aerostats for carrying a variety of aerials and other equipement for both relaying and broadcasting.

Technical broadcast performance standards will need to be developed if individual regions of a country have independent broadcasting units in order to weld these into a national service. These will have to be agreed to and imposed nationally. These considerations are not likely to present many problems for sound broadcasting, but if a small local television unit wishes to contribute to a national network this may be impossible unless it has equipment up to the required standard.

#### COMMUNICATION SYSTEMS AND PURPOSES

Before technical planners start to set up rigid structures to meet early demands for communication facilities, which always seem to be concentrated in the areas of the mass media and one-to-one interpersonal communication, national policy-makers need to consider the overall correlation between types and purposes of communication systems and how they are likely to develop in the near future.

Communication systems have three basic parameters, whether they are

considered from the viewpoint of the government or from that of another authority which provides them. These parameters are purpose or function, numbers of people involved, and degree of interaction between people involved (user's role).

Consideration of the three aspects of communication systems can be facilitated by use of the three-dimensional matrix (Fig. 7). In the matrix the two continuous dimensions of number of participants and the user's role have each been divided into four convenient bands, although it is recognized that these merge and overlap.

Many of the elements in the matrix are common. For example, the element 'educational-passive-masses' describes educational radio broadcasting. Others are only found in a few countries: for example, 'politicalinteractive-individual' could describe a phone-in response to a party political broadcast. Some exist only in very limited areas: 'social-participatorycommunity' might describe video programmes being made by community action groups for transmission over local cable networks.

The matrix also makes possible identification of systems that are either not yet available or possibly unwanted: for example, 'political-feedbackmasses' might describe the use of a domestic telephone or television cable network with push buttons allowing voting in national referenda.



FIG. 7. Communication systems matrix.

Developing countries with educational priorities may be interested in ensuring that their communication system will provide an 'educationalsmall groups-interactive' element. This might be implemented in practice by ensuring that the technical planners provide telephone exchanges that will allow-groups of students to have telephone links with remote tutors.

At different stages of development other priorities may emerge. A government's desire to bring its people to a more advanced stage as quickly as possible may focus on the 'political-community-participatory' element. In technical terms this might represent a mobile television recording and replay facility. Although financing might be provided by the central government, technical planning might be more appropriate in the local situation where knowledge of environmental factors is greatest. Such factors may include terrain, climate, roads, electricity supply and ability level of staff in technical planning; and local religion, culture, language and social history will all influence the production style.

#### TECHNOLOGICAL ALTERNATIVES

It was only a few years ago that engineers produced a unique solution to some problem that was presented to them. Now the difficulty is to persuade planners that their problems are not unique and that engineers generally have a number of solutions on the shelf to meet any technological task the planners identify.

This situation is certainly the case in the engineering/technological aspects of communication. There is often an urgent need for top-level decision-makers to understand enough about each of the technological alternatives to enable them to appreciate the sociological implications of each potential choice.

To facilitate the development of this technological awareness it is necessary for top managers to save time by such means as avoiding waste on pointless high-precision forecasting based on inadequate demand data; decentralizing the management of activities that can safely be left to local managers, often with local community involvement; and not reinventing such things as computer aids to planning that already exist and are adaptable to any country.

Top-level planners of technical facilities in a developing country are often subjected to intolerable pressures arising from simplistic and sometimes unrealistic general policy decisions. Such decisions are usually concerned with time and people. For example, two countries might have similar targets for communication system development by the end of a five-year planning period. The first may decide to import all necessary equipment and technical expertise for installation and initial operation from abroad. The second may have a totally different approach, based on a firm policy of very rapid indigenization, of developing the system with full reliance on local equipment and expertise.

Such extreme policies may both lead to serious difficulties. Under the former policy there is the likelihood of acquiring unsuitable equipment, and national staff will not have the opportunity to gain experience from planning and installing the system they will eventually have to operate. In the second case, the few staff with planning skills may be hopelessly overworked. Technically trained installation and operational staff may be quite unobtainable in sufficient numbers in the time available because of inadequacies in the educational system (which are quite beyond the control of the technical planners). A judicious balance of internal development with external help and an early concentration on technical staff training will give a much better chance of both meeting targets for installation of new systems and keeping the systems operating efficiently in the future.

Sometimes technical planners may be so concerned with producing the most technically elegant solution to a planning problem that they may distort the picture of alternatives presented to the decision-makers to whom they report, or worse still, they may blind themselves to the alternative solutions, without any conscious decision being taken. Installation may proceed on the technically 'obvious' solution. This point may be illustrated by reference to an imaginary communication network connecting five nodes, as illustrated in Figure 1.

As we have seen, a good engineering solution effecting a compromise between economics and reliability would be as shown in Figure 8. Let us assume that 'A' is the national capital and that the main geographical features are as indicated.

Other development plans might include development of industries along the coastal strip between 'A' and 'E' with a road system to service it. Urgent attention might also need to be given to improving the agriculture around the many small villages on the plain between 'E' and 'D'. These two developments would call for a network as in Figure 9 to be given priority.

This would leave 'C' rather isolated, with a long route to the capital and no stand-by route if the lines between 'B' and 'C' developed a fault or were stolen for the copper, as sometimes happens.

'C' and 'D' are state capitals and, both being in mountainous areas, have much in common. The leaders in these two areas are keen to work together and believe that the health of the country depends on strong regional government, a view which is shared by a majority of the national government. Therefore the political priority for the network would be to provide efficient and reliable links between 'C' and 'D' and between 'D'



and 'A' and also between 'C' and 'A', thus demanding a network as in Figure 10, which involves expensive microwave links over the mountains between 'C' and 'D'.

Thus we have three alternatives, which for convenience we may call by the main considerations which determine them:

- (a) engineering (Fig. 8);
- (b) commercial (Fig. 9); and
- (c) administrative (Fig. 10).

If (a) is adopted the five cities will get a more reliable service as soon as possible at minimum cost. (Also, the village development will be made easier.)

If (b) is adopted industrial and commercial interests will get the most help. Maintenance costs on the 'EA' link will be minimized when the road link is complete. 'E' will grow in importance while 'C' and the concepts of regional administration will have a hard time.

If 'c' is adopted it may lead to more bureaucracy in a country whose financial position will deteriorate as it needs to import more food and has no corresponding growth of exports, since industrial developments are being frustrated by poor communications.

There is no right answer to this problem. Top-level policy-makers should understand each of the alternatives and make a deliberate choice between them, bearing in mind the capital and operational costs and the longer-term consequences involved in each case.

Other network planning questions will occur, such as where to locate the international gateway exchange. Historically this may have been started on a small scale at the capital 'A', but if satellite links via Intelsat are to be used in future it might be better to locate the ground station near 'B', which is more central to the network.

In most developing countries capital for development is in short supply. Whatever communication system development is planned, there will almost certainly be a need to achieve all possible economies. It is very difficult to effect any substantial savings on transmission equipment by special design for limited channel capacity and relatively low usage without accepting substandard performance for transmission quality and circuit availability. However, it is possible to effect significant economies in system planning as indicated above and by judicious choice of switching equipment. To do this effectively help and training are needed. Much of the ITU's efforts are directed towards providing such information, advice and assistance. A number of other organizations, including some universities in the more developed countries, also provide educational opportunities and research assistance for communication projects appropriate to developing countries.

#### THE THEORY AND PRACTICE OF COMMUNICATION SYSTEM PLANNING

Every country in the world has some form of communication service, however limited it may be. This will automatically form a starting point for the planning of future communication services, which will include both the mass communication element of broadcasting and the interpersonal element of telecommunication (both internal to the country and external to the rest of the world).

Simple extrapolation from existing levels of service and known demands may lead to a very false picture, as Figure 11 attempts to show.

The basic forecasts of projected demand may be very inaccurate for two main reasons. First, in areas with little or no previous experience of a particular communication service (such as television broadcasting), the only demand will come from a minority of the population, who have either travelled to or read about areas that already enjoy such provision. Secondly, the large gap between supply and demand for a particular service (such as private telephone services) may be so well known that many people do not bother to put their names on waiting lists, thus leaving a distorted picture of true demand. The shortfall of supply will also lead to high intensity of use of equipment, leading to high malfunction rates, which in turn lead to more call attempts failing. Failure leads to more repeated call attempts. The result is an even higher intensity of system use but no increase in revenue or demand.

A number of other factors will affect projected demand. Economies of scale or the development of new techniques may cause costs and therefore charges to consumers to drop, and this in turn will increase the latent demand for services. Also, early systems may offer very poor services, convincing many people that they are not worth having. Thus any improvement in quality and reliability of services must also be recognized as another factor that will in turn increase the future demand. Finally, the provision of new services, whether new in kind (such as adding facsimile transmission to a telephone network) or new in application (such as provision of time checks or weather forecasts on an existing telephone system) will also lead to an increased demand for even more services.

Potential demand can be limited to some extent by tariffs and license fees or other methods of charging for services. This applies mainly to initial payments (connection 'charges') for telephones, which have been set at several hundred dollars in some developing countries. Although use charges to customers also vary considerably because of variations in operating costs and other factors, there is little likelihood of these having much direct effect on demand. However, they will typically pro-



179

duce only some 40 per cent of the financial resources needed for systems development.

Many communication services are planned to meet estimated (or forecast) demand from clients. However, in many developing countries demand-based planning is difficult to implement fully for a number of reasons; each of which is shown in Figure 11.

Government decisions to override economic considerations in order to provide communication services in support of educational, community health, agricultural and similar development programmes may create increased demands for both broadcasting, telecommunication and journalistic services, particularly in rural areas. Similar demands may arise from national political decisions to provide the whole population with some minimum access to emergency services, broadcasts, or telephone services at the same basic charge irrespective of the actual costs.

Industrial or commercial developments, particularly those related to the exploitation of oil or mineral deposits, may exert irresistible pressures for unplanned and rapid development of all types of communication service. Expansion of the service may be possible only by sacrificing expansion of other services, thus distorting the shape of future services.

Capitalization of system development can also alter the shape of the system. As we have noted earlier, the revenue from operational services may be expected to provide around 40 per cent of the capital required for development. The remainder will have to found elsewhere, mainly in the form of loans from various sources and aid grants from foreign governments. Much of the capital investment will need to be on sophisticated equipment that cannot be obtained from indigenous sources, and therefore foreign-exchange requirements may cause further problems, which will at the least delay developments and cause more distortion of the future programme.

Staff problems are likely to prove one of the biggest causes of failure to meet planned growth. All communications systems require a range of qualified and highly trained staff extending from craftsmen to professional engineers and including journalists, artists, producers and directors for the various media. Even if the communications authorities assess their staff needs correctly and follow this by setting up adequate training units in good time, their endeavours may be frustrated by the lack of suitable infrastructures in terms of schools, technical colleges and universities or polytechnics to provide the required output of educated manpower able to undertake the specialized training.

In addition to these factors represented on the diagram, there will be others that are likely to limit growth rates, such as internal inefficiencies within the communication organizations and their lack of sufficient autonomy, leading to cumbersome and protracted negotiating patterns with other interested bodies or departments.

Despite all the limitations described, the actual growth, although slower than planned, may be very high in absolute terms or when compared with that in developed countries. This leads to an unstable system in which even technical choices are highly interactive and engineering problems cannot be solved separately from each other or from related socio-economic problems, and it introduces further uncertainty into the relationship between what is planned and what is likely to be achieved.

Similar influences on planning precision to these at the national level will also be found at the local level. The following example illustrates the alternatives that might arise in planning expansion of telephone services, assuming limited capital or foreign-exchange availability.

#### Typical development alternatives in a small area

Let us assume a rapidly growing city, a few towns and many small villages in a selected area. In the city we shall find telephone waiting lists for both domestic and business properties. In the rural areas other development plans for such things as agriculture, health, housing and local industries may all need access to the telephone system so that they may make the progress planned for them.

To favour either urban or rural development to the exclusion of the other would obviously be wrong, but deciding the best balance between the two, given the inevitable limitations imposed by lack of capital and/or lack of skilled manpower, is a far more difficult task.

Communications facilities for business houses in the city may be essential to the development of external trade and the in-flow of goods required for national development programmes, but beyond a certain level of development the capital investment required for exchange improvements will be very high, and the same investment spread more thinly over a large number of rural areas could at least provide public telephone facilities that could give the majority of the population some measure of access to other services.

Even if a policy decision has been taken to improve telephone services in rural areas and an allocation of capital has been made to do so, the alternatives are immediately apparent:

- 1. Spread the capital over a large number of villages by buying many relatively cheap manual exchanges. Advantages: more villages connected and jobs as operators created for the underemployed rural labour force.
- 2. Install automatic exchanges in fewer villages. Advantages: chosen

villages will have a dial service, and there is no need to train operators and therefore no delay while this is done.

If I is chosen because of a desire to provide a service (even if of marginal quality) to the maximum number of people (a choice that will also provide a quick income from connection charges), it may well lead to serious congestion because exchange and trunk routes become overloaded. This in turn will lead to a low income and so fail to provide capital for future growth. If 2 is chosen a high-quality service will be provided that should produce adequate income for growth in the future, but initially it may be a politically unpopular approach.

#### Summary of planning points

It has been noted that while many planning problems are similar at the national and local levels some solutions are best sought nationally and others locally.

Common strands that have emerged are:

- The difficulty of accurately predicting future demands from information about existing services.
- The close interrelationship between technical and socio-economic planning, which is greater when the pace of development is high.
- Maximum conformity to international standards needs to be encouraged from the earliest stages of development.
- Plans for the provision of sufficient trained personnel are just as, if not more, important than those for equipment.

### Chapter 6

### Short- and long-term forecasting for communication technology planning

John Spence

Organizations involved in communication technology have varying planning needs—ranging from requirements for a demand forecast for the subsequent twelve months to requirements for a strategic plan for the next twenty-five years. Often, because of the long economic life of communications equipment and the long time-lags involved in implementing new communication technologies, there is more interest in the longer-term future so that strategy can be formulated and a communication development programme established.

Making 'forecasts' of such long-term events is rightly viewed with scepticism. However, while the difficulty of prediction may militate against the usefulness of any planning methodology whose output is a unique forecast, the strategic importance of communication demands that systematic planning be undertaken. One approach that has been found effective for long-term planning is that of assessing the implications of alternative futures. This is a satisfactory and practical basis for decision-making and is recommended here.

The particular requirements of short- and long-term communications planning are discussed in more detail in the following two introductory sections. We then go on to discuss forecasting methods for both shortand long-term communication technology planning in some detail. The chapter closes with a section on our conclusions.

# The requirements of short-term communication technological planning

All technological planning is to a large extent oriented to decision-making. Forecasts of communication needs, whether for the short-term or for the long, are ultimately geared to policy assessment or to the adaptation of the planning organization to the effects of innovation and development, to the extent that these effects are forseeable from the perspective of present-day possibilities and applications.

It may be argued that the objective of any form of planning is to make decision-making in an organization more consistent and purposeful. Plans are not formulated in isolation but rather evolved from the interaction of policy and forecast, and from an assessment of the organization's capabilities and objectives. They arise therefore from the questions raised within an organization, depending ultimately upon the organization's perception of its goals and opportunities.

Short-term planning, however, owing to the long time horizon of many technological innovations in communications, concentrates on the opportunities for growth and expansion within an existing structure of products and markets. It shows significantly less concern with the potential discontinuities arising from major changes in communication technology or in market structure.

In many cases, this narrowing of perspective to short-term goal assessment reduces to a specific concern with the evaluation and selection of individual projects in communication technology or in short-term economic planning. However, more generally, short-term planning has an important role in the assessment and revision of long-term organizational or technical goals in the light of the extrapolation of short-term trends and developments.

Since all decision-making involves risk, if only of lost opportunity, it is clear that the role of the communication planner does not end with the mere identification of opportunity but must extend to its assessment.

To achieve each of these objectives, the planner must accomplish three primary tasks: (a) identify all of the feasible alternative courses of action (or alternative directions of technological change), (b) assess the consequences of these courses of action (or technological developments) within a suitable time horizon, and (c) determine the degree of risk or uncertainty involved in these alternatives.

For the short-term planner, these tasks become quite specific. His concern with the social or technical effects of innovation or technological development reduces to consideration of these effects in relatively continuous phases of their development; his concern is usually with the time span between major discontinuities due to innovation; and one of his most frequent concerns is with the development or investment cost of the alternatives to be considered.

A paradigm for short-term planning thus involves four major phases: technological assessment, data acquisition, forecasting and decision-making.

The first of these phases, technological assessment, involves a relatively simple application of many techniques common to long-term forecasting,
such as the development of likely long-term innovations in communication technology. More details of these techniques may be found below.

However, technology assessment applied to short-term communications planning involves three steps:

- First, it is necessary to analyse the theoretical potentials of a selected communication technology and its present state of the art, including estimates of development for each potential innovation. The result of this analysis will be to identify any possible discontinuities of growth within the time horizon of the project.
- Secondly, the demand pressures influencing the development and growth of the communication technology must be isolated. The factors considered should include short-term analyses of economic activity, of existing social policy, and of sectoral variations in demand resulting from differences in application or in market penetration.
- Thirdly a time horizon for the forecasts must be defined over which it is feasible to assume a continuous development of both market and technology.

The second phase of the process, data acquisition, will normally require a significant amount of field research to enable the planner to establish a data base upon which forecasts may be validly constructed.

The precise physical structure of the resulting data base itself may range from the complexities of a computer-based, dedicated hierarchical retrieval system, to more traditional structures such as sequenced manual records or specialized single-application computer files. The main determinant of this structure will be the scale of communication planning involved.

Whatever the ultimate form of the data base, its construction draws on two primary data sources. These are (a) the existing records (such as sales, rental or maintenance files) of communication equipment or of its users, obtained from the planning organization itself or from other suppliers of communication services, and (b) a marketplace survey of one or more categories of user, to acquire new forms of data or to update existing records.

The forecasting phase involves the production and assessment of forecasts of demand over the selected time horizon. These forecasts are based upon the data acquired by field research in the data acquisition phase, and take into account the constraints and growth opportunities identified in the initial phase of technological assessment.

Finally, on the basis of these forecasts a decision must be made on the policy options confronting the planner. These decisions in turn become factors modifying the original technological assessment, and this process gives rise to the essentially recursive nature of short-term communication planning.

# The requirements of long-term communication technological planning

In today's environment attempts to forecast even a few years ahead are regarded with considerable scepticism. Attempts to forecast the long-term future, say twenty-five years ahead, are frequently viewed as lacking practical value. Yet, because of the time-lags involved in developing communication technologies and because of the often long economic lives of the various facilities once they are installed, many strategic planners in the communications industry must deal with the long-term future.

However, noting the huge amounts of resources in question and the enormous costs of modern mistakes, successful attempts to grapple with this planning problem would promise significant benefits if uncertainties could be minimized. Most sections of the communication industry cannot adopt the stance of simply letting economic and technological forces run their course because they are no longer able to respond quickly enough to change. Therefore, the impossibility of predicting economic, technological, social and political forces twenty to twenty-five years into the future does not invalidate the need for long-range planning.

Indeed, since decisions must be made 'today' about the longer-term future, it should be noted that long-term planning results are a critical input to sound short-term decision-making.

Valuable long-term planning results are achieved by the systematic examination of alternative likely futures, based on combinations of identified trends and possible events. In other words, instead of attempting to predict a single unique future—either nationally or internationally—planners analyse uncertain future events to gain an understanding of the implications of a particular circumstance, or set of circumstances. In this way an organization improves its capacity to deal with change by being able to identify new emergencies and opportunities at the time they start to appear on the futures horizon, and plan accordingly. More importantly, leading indicators are identified and monitored, and contingency plans and strategic actions continually updated as new trends or developments emerge.

The refinement of traditional forecasting techniques does not necessarily provide communication planners with this ability. This is because traditional extrapolative forecasting techniques do not allow for the incorporation of unexpected developments into the planning process. Extrapolative techniques generally rely on building a picture of the future starting from the present and using known trends from the past or known plans for future developments.

Consequently, such techniques lose much of their usefulness and power

- . - (\_\_\_

in a field such as communications, where, because of new socio-economic developments and rapid technological change, the amount of relevant history on which to base projections is usually very small.

The inadequacy of the present as a basis for predicting the future is highlighted by a consideration of the multiplicity of technical developments either at present under way or expected in the communications industry, their impacts on one another and their interactions with social, political, and economic events and trends. For example, consider the problems that existed in 1950 in forecasting out to 1975. In 1950 the transistor had just reached the market and the computer was scarcely known; communication satellites were a concept only; the first manual copying machines had only just come on to the market; instant colour television reporting from anywhere in the world was only a dream; the prospect of landing men on the moon, the potential importance of uranium, the impact of the environmental movement and the world population explosion were not appreciated.

Planners in 1950, using traditional extrapolative forecasting techniques, would not have been able to predict the particular combination of conditions that existed in 1975. Moreover, they would have missed many of the most significant developments of the period—by having neglected to consider extreme conditions, unpalatable events and trends, or 'irrational' extrapolations from the then present; or perhaps through tending to analyse only economic data or events with high expected probability.

In approaching long-term forecasting, we are proposing that the examination of alternative possible futures is a more realistic and practical way of identifying strategic actions than is the application of a single future forecast. It has been our recent experience to apply these forecasting techniques to studies concerned with the future of communications in Australia, and the forecasting and evaluation of several alternative possible futures proved invaluable.

The procedure can be broadly described as follows. For each set of likely future developments, a 'picture' or 'scenario' is developed to describe various social, economic, political, and technical factors. The factors are chosen on the basis of their relevance to the external environment in which the particular organization will be expected to operate. Along with these external scenarios, alternative choices or strategic options for the organization itself are developed and then examined in the light of the scenarios. This examination is made through the use of a 'cross-impact analysis' technique, as described on pages 129–33.

The examination of strategic options against descriptions of the several alternative futures enables the implications of significant competitive developments to be isolated and consequently the appropriateness or the inappropriateness of particular options to be assessed. In the case of communications planning, the strategic options are usually technological in nature, and so we have used the term 'technology prospects' to describe those alternative technology strategies that need to be considered.

Futures research, using scenarios, technological prospects and crossimpact analysis, offers an approach that ameliorates some of the shortcomings and difficulties one is faced with if one seeks to apply traditional forecasting techniques to a study of the long-term future.

The important outputs from a planning viewpoint are not so much a particular result from a particular scenario as the implications of various results derived from a consideration of several scenarios. An organization employing this approach benefits in the following ways:

- 1. Plans include detailed quantitative forecasts and an assessment of the implications of these.
- 2. Non-economic variables are explicitly included.
- 3. Critical planning lead-times are identified.
- 4. New opportunities are often identified.
- 5. Potentially threatening events are easily considered in the planning process.
- 6. Monitoring and contingency planning are concentrated on trends and developments that are identified as critical.

Forecasting methods: key elements and techniques in a communication technology planning methodology

The methodology used for communication planning will of course vary from problem to problem. Common elements and techniques, however, are likely to occur and be used. As a basis for discussing some of these, an overview of a methodology for communication technology planning which we have successfully used is provided in Figure 1. This shows the major outputs of the methodology, intermediate results derived to achieve these outputs and the techniques employed in conducting the research. The diagram divides the elements of the overall technology into short- and long-term applications.

The remainder of this chapter will focus on describing and discussing the usefulness in communications planning of what we have referred to as the 'intermediate research outputs' in the diagram and the techniques used to derive them.

### SHORT-TERM COMMUNICATION TECHNOLOGY PLANNING

Four key phases of short-term planning were described above: technological assessment, data acquisition, forecasting and decision-making.

A restatement of these phases is provided in Figure 1. In this new scheme, technological assessment and data acquisition form the basis of the planner's understanding of present conditions and influences; projections, derived from the short-term interrelationships established by data analysis and model-building, lead to forecasts regarded as valid in the short term. From these forecasts, decisions are made, and the validity of long-term strategic assumptions assessed.



FIG. 1. Outline of the elements of a methodology for communications planning.

This section considers in more detail the techniques involved in each of these phases of the short-term planning process.

### Phase 1: technological assessment

There are two aspects of this phase of planning: (a) a preliminary literature search (covering technical publications as well as manufacturers' literature) to establish a background to other forms of data gathering in the project, and (b) an analysis of the current state of the art to predict the direction of technical innovations in the technology under study, and to establish a broad time horizon within which short-term methods of analysis are valid.

### Phase 2: data acquisition

The communication technology planner requires several classes of information for short-term forecasting. These data include historical details of equipment usage in the technology under consideration, historical details of traffic levels corresponding to this equipment, adequate details of present-day equipment installation, traffic levels and application areas in the technology under consideration, historical and forecast indicators of economic activity, and information regarding particular social, economic, or technological constraints which may have affected these historical measures, or which are likely to influence forecasts of future growth.

These quantitative and qualitative data form the data base for the forecasting project, and are used in the formulation of models and in the calibration of forecasts in the next phase of planning.

It should be noted, however, that the term 'data base' refers simply to the total collection of data to be used, and does not imply the format or structure of the information, although its use in the literature has come to refer to a structured, typically hierarchical, retrieval system based on a computer. More correctly, however, such a system should be referred to as an information retrieval system, or data base management system (DSMS).

More generally, a data base consists of information occurring in a variety of forms. For example, in raw or processed manual records or in other hard-copy forms (e.g. equipment sales or rental records, subscriber lists, billing records); in electronic or magnetic media records (e.g. magnetic tapes, discs or cards) perhaps associated with existing computer retrieval systems; as tabular data (e.g. origin/destination traffic matrices); in graphs; as mathematical formulae (e.g. existing mathematical models or predictive relationships verifiable from independent sources). Most of these data will already exist within the planning group or in the records of the administrative body.

Where existing records are used, a considerable effort is required to collate records to a form viable for the selected forecasting techniques. For example, communications records frequently relate to the sale, rental and maintenance of communications equipment (telexes, telephones, facsimile machines, etc.) supplied or controlled by the planning organization, or else occur as usage records maintained on a regional and subscriber basis, for the purpose of subscriber billing.

In each such case, records must be centralized, standardized and collated into the forecasting data base; each of these operations frequently involves extensive manual intervention, with a considerable risk of transcription or fatigue-induced errors. Because of these factors and the resulting need to validate information and to correct errors, manpower requirements for the collation of existing records are considerable.

Additional information, particularly on present and planned usage of the communications technology, can only be obtained from field surveys. There are several reasons for this:

First, reliable forecasting requires some knowledge of the strategic intentions of users and non-users of the communications systems under study.

- Secondly, since maintenance or even accounting records are often incomplete, owing to their specialized nature, they are often unsuitable for forecasting purposes, or, where complete, may provide unsuitable measures of communications traffic or equipment types, for example.
- Thirdly, existing records seldom indicate the actual applications of the communications equipment and the users' plans for new applications; they also seldom identify the sector of the economy in which the user operates, and thus restrict even indirect forecasts of applications.
- Fourthly, no central record of some forms of communications equipment may exist at all, owing to their recent introduction or limited use.

There are several forms that such field research can follow. The major types are:

- A manufacturer survey, in which records of equipment types, technical standards, roles and market share are sought.
- A national systems survey, in which major public sector communication networks are studied. While these will vary according to national or regional policy, they should include networks associated with defence, scientific research, public health and government administration.
- A large networks study, in which significant non-government users of communication technology are studied. These systems include (where relevant) booking services (such as used in hotels, motels, airlines and

other transport services), funds transfer (whether by telex or data transmission), and special purpose communications sytems relevant to the technology under study (e.g. computer service bureaus, message-switching networks and other common carrier utilities).

A user study, in which a wide-ranging market survey is conducted of all significant users of the communications technology under study throughout the economy.

In more extensive planning tasks, several or all of these surveys may be required.

In each case, the design and execution of the survey itself follow fairly conventional lines. The main steps involved are the formulation stage, involving identification of the market structure and questionnaire design; sample selection (frequently on economic sector and regional grounds); detailed survey planning (including a pilot study and planning fieldwork management); and the execution of the fieldwork.

Details of the design and management of each of these steps may be found in any standard reference work on survey design, although a communication study usually demands considerable alteration to the definition of the sample frame, and to the creation of an adequate sample.

Whether data is obtained from existing records or from field studies, the data-gathering process and its validation will place heavy manpower demands upon the planning team.

Further demands on manpower arise from the need to manipulate data for information retrieval purposes, or for the creation of specialist data files, tabulations, or a complete data base management system. These tasks are of course in addition to the requirements of the forecasting methodology itself, which demand significant amounts of cross-reference among the data items and other forms of data manipulation.

Because of the complexity of the data-gathering task in a major communication study, we have found it preferable to approach data-gathering, from the very inception of the project, with the ultimate uses in mind—such as the creation of an ongoing information retrieval system.

Strictly speaking, the analysis of the data base is inherently part of the model building and forecasting phase of any study. However, there are three forms of preliminary analysis that may be derived once the data base is established. These three are elementary statistical analysis, trend analysis and regression analysis. Although initially done prior to forecasting, the latter two analyses will most probably be repeated a number of times in the course of refining the forecasts.

The immediate results of these analyses are sample statistics that may be used for the calibration of analytical models; measures of trends and their statistical significance; and the isolation of possible causal factors and interrelationships, indicated by regression equations.

### Phase 3: forecasting

Although the preliminary analyses of the data base form the basis of initial model building, there are several basic forecasting techniques capable of application to certain classes of short-term forecast. These are:

- Judgement, in which subjective estimates of predicated values are estimated by persons qualified in the area concerned.
- Trend, in which statistical curve-fitting techniques are used to project future figures based on the growth trends indicated in historical values.
- Econometric, in which a relationship is established between values of the predicted variable in any year, and the values of other factors  $x_1, x_2 \dots x_n$  where it may be felt that future values of these factors can be forecast more accurately than the predicted variable itself. Frequently such relationships are derived from regression analyses, such as time series.
- Normative, where the stated intentions of surveyed organizations or individuals regarding communications usage are used directly as an estimate of future values.

These techniques and other extrapolation methods derived from them are widely described in reference works on statistical and numerical analysis, and need not be discussed here in any technical depth. Their use, however, will vary according to the amount of data acquired in field work, and the length of time the technology concerned has been established. For example, in a fairly new communication technology, such as commercial facsimile transmission, many countries and regions would be restricted to judgemental or normative forecasts of equipment demand or traffic.

On the other hand, we have found that established technologies such as telex or telephones are very successfully modelled on an extrapolative (trend) model, or even by econometric means alone.

In each case, the use of the method requires a careful calibration of the technique against measures derived directly from the data base. For example, on the basis of considerable historical information, it may be possible to derive a trend model using the traditional S curve. Following an approximate curve fitting, various types of S curves would then be examined to determine the best fit to the historical measures. In some cases, different S curves may be found to be applicable to different industries, or to different types of communication service application, depending upon the technology under study.

Other factors, such as economic trends, can later be correlated to this basic predictive curve, and more complex models developed.

Finally, when each type of model has been successfully validated against historical trends, forecasts of equipment, traffic or application are made, and the error in these forecasts estimated for each industry, region or communications service under study.

This later stage of the modelling process will determine actual forecasts of usage over the time horizon considered, estimate the sensitivity of these forecasts to errors in the parameters of the model, and forecast the degree of saturation of the communication medium under study, according to its industry of application, as a guide to forward planning.

### Phase 4: decision-making

The precise impact of forecasts on policy and decision-making varies with the planning needs of the organization. From our experience, there are three broad classes of decision to which forecasts are applied. These are investment decision-making, policy assessment, and 'early-warning' planning.

The first of these classes of decision, investment decision-making, applies to those plans relating to pricing policy or to capital investment. These problems frequently require the use of more complex models derived from the elementary forecast models, and the use of such techniques as discounted cash flow analyses or financial and accounting models.

Decisions involving policy assessment are more complex, owing to their lack of general structure. Typically, however, these require more stringent simulation models, or the application of other operations research, technological or econometric techniques. Decisions falling into this class include the technical assessment of network development or technical innovation.

The final broad class of problem relates to the ongoing use of forecasts as an 'early-warning' guide to management for forward planning. It refers to the use of forecasts to isolate key industries or communication networks nearing critical levels of performance. These include, for example, unusual load patterns on specific channels or networks, regional variations in demand, saturation loading or unusual deviations in growth or usage.

In each case, the forecast itself is used merely to signal the likely future behaviour, although the implications of the forecast are assessed and used directly in forward planning.

### LONG-TERM COMMUNICATION TECHNOLOGY PLANNING

The primary objective of the long-term communication planning methodology outlined in Figure 1 is to provide a research package that is repeatable and enables longer-range futures to be examined and assessed. The detailed descriptions of these alternative futures are not, at the conceptual stage, derived from the present. The link between the uncertain futures and the need for making definite strategic decisions in the present, or shorter-term future, is provided by identifying the implications of these futures and by establishing a monitoring procedure.

The first two major results of the methodology build pictures of how various societies might communicate in the future and isolate the effects of these environments on the particular communication system under consideration.

To achieve the first of these results, a method is required for developing detailed descriptions of environments that from our present standpoint might well appear highly unlikely; and then for exploring within each of these environments the impact that developing technologies may have. (In addition, the methodology must examine the potential ways in which the environments might influence the development of the technologies.)

The way we have found to meet these two requirements is by postulating future environments of strategic relevance to the planning organization. These socio-economic-political environments, or scenarios, include conditions which are generally considered to be very favourable. Technical developments in communications can then be brought together with the various future environments described in the scenarios. This enables us to determine if the technology would develop in that future and what its general use is likely to be, given the other technologies that might also be expected to be implemented in the situation described in the scenario.

The order of this methodology might be reversed. For example, scenarios could first be written for communication technology, and developments of an economic, social and political nature then introduced into these technical scenarios. Either way, the methodology fulfils one of the distinguishing features of futures research used for planning purposes—the ability to make dynamic 'experiments' in a future that is not connected with or wholly derived from the present. The 'experiments' are intellectual in nature, but the results are quantified descriptions of the likely degree of implementation of the various technology prospects.

Long-term futures work of this type is most relevant when carried on in conjunction with more traditional short-term studies aimed at gaining insights into present communication behaviour. Techniques such as market surveys, regression analysis, simulation, dynamic modelling, and trend extrapolation provide a quantitative data base that can also be used to guide the translation of future situations into statements of demand for a particular communications service.

A most important feature of any communication planning methodology is that the methodology itself be repeatable. At any particular time, long-range research must be effective in exposing strategic issues that should be appropriately faced at that time or a given future period. However, if futures research is an ongoing process, new developments will be identified earlier than otherwise and fewer surprises can be expected. Inherent in the methodology is the implication that a range of futures and technical possibilities is always being explored. Repeating the methodology at intervals, when either circumstances have changed or new factors and information have come to hand, enables the direction of development and the implication of this direction for the organization to be assessed.

One of the key outputs of any long-range planning of this nature is the structuring of a monitoring programme within which key future developments are identified early and incorporated immediately into the strategy formulation process.

### Scenario design, construction and use

A scenario is a clear statement about postulated future environments of significance, and usually embraces social, political, economic and/or technological possibilities. The use of scenarios is a way of deriving a range of futures that do not have to be wholly based on perceived trends and influences in the present.

Although it is possible to describe the steps in a procedure for the development of scenarios in fairly straightforward way, such a description is liable to make the scenario technique appear far more simple than it is in practice. The process of development of scenarios is highly iterative and interactive, and the material must be revised many times throughout its evolution.

Because of this, the steps in scenario building described below will vary from case to case. They are, however, discussed for their illustrative value.

The development of non-technical scenarios for use in communications planning starts by selecting socio-economic-political themes around which the scenarios can be developed. It is important to emphasize that these themes should not be chosen because of the likelihood of their emergence as an accurate description of the future, but rather because they expose those extreme conditions that would evoke special economic, technical and social circumstances of relevance to the planning organization. Thus the themes chosen would include circumstances that might appear at present to offer a favourable environment to the organization, or an unfavourable one, or simply a set of possibilities that are different from the present although the direction of their actual impact is not readily apparent.

A number of techniques can be employed to help establish these themes. The simplest is probably just to survey opinion within the planning organization (or in relevant areas outside it) in order to establish broad conditions under which the organization's responses might need to be different. A procedure such as this might be supplemented with a simple normative or goal-oriented forecasting approach. This is one in which goals for the future are hypothesized and the circumstances to realize them are assumed to evolve. (This technique can range from simply the recording and expansion of an individual's view of the future to a highly structured reasoning process. Its use in the development of scenario themes is best kept simple and used primarily to generate ideas.)

Once the scenario themes have been decided on, initial drafts are prepared. The preparation of these drafts provides a major guide to the analysis that next needs to be undertaken to refine the draft. For example, long-term trend analysis might be required to refine those factors which have been included in the scenarios because they are critical to the themes. Since useful scenarios should be quantified as far as possible, this trend analysis is a critical element and can also be used to ensure consistency within the scenario.

To refine and provide another check on consistency, the whole scenario and selected parts of it should be subjected to outside review. This review typically takes the form of comments submitted by experts or a more structured survey of expert opinion on more specific aspects of the scenarios. One example of the latter method is the Delphi technique. This technique, described on page 122, is based on the polling of expert opinion in such a way that the source remains anonymous, the respondent's reply is as precise and quantified as possible and the aggregated results of the poll are then fed back to the panel members, who are required to challenge or support predictions that fall outside the interquartile range.

Each of these procedures yields refinements that require further analysis and checking, and perhaps final incorporation into revised scenarios.

Up to this stage the format and scope of the scenarios are kept fairly flexible to enable recognition and assessment of additional factors of importance. At some stage, however, to facilitate further use of the scenarios, a consistent format should be developed.

A quantitative summary of the scenarios, in the form of a scenario range table,<sup>1</sup> facilitates the choice and use of critical factors and is a device for ensuring internal consistency. A range table also helps check the relationships between themes and quantitative data on the one hand and between themes of different scenarios for further analysis.

Scenario building is an iterative procedure. Techniques of the type

I. A scenario range table sets out all the critical factors used in the development of the scenarios and shows the ranges and actual values where applicable for each factor for each scenario.

described above should be used to refine the scenarios until they present pictures of social, economic and political conditions at a point of time in the future which are internally consistent and which together encompass a range of conditions significant to the planning organization.

By enabling the study of a range of alternative futures, the scenario technique broadens the horizons of events being considered and therefore heightens the likelihood of identifying developments that would have an impact on the planning organization, requiring it to take strategic actions.

### Methods for including the effects of new communication technologies

### Overview of approach

The analysis of existing and emerging communications services is a fundamentally important aspect of communications planning. To be able to plan for long-term communications demands necessitates an hypothesis of the communications services to which governments, businesses and individuals will have access. For each potential service, statements about likely demand levels need to be made.

The approach to long-term communication technology planning being outlined in this chapter assumes that the study of communications technologies can be initially separated from the other sections of the planning problem (which are concentrating on the possible future societal developments). This assumed separability for the purposes of initial study is more valid in some countries where the new communication technologies are likely to be imported. Even in highly technologically developed countries, the approach has merit in that technology research is proceeding in many areas at any one time, and it is econo-politico-social conditions that eventually determine the extent to which particular technologies are implemented.

A critical problem in examining the effects of new technologies is the structuring of the analysis—there is an enormous range of research effort and experimentation under way in several countries at any one time, and a single study in one country or for a communication authority in a part of a country is unlikely to be of sufficient size and budget to cover all these emerging technologies or research efforts.

We have developed a technology 'tree' approach to this part of the planning problem, feeling that the main thrust of any technology study in the communication field should be to identify and describe future communications services. For example, technological advances in the fibre optics field will make wide-band distribution networks around central business districts and into homes a reality. While this is certainly of extreme interest, for example, to telecommunication network planners, the extent to which the new network would be used is possibly of higher importance. That is, we have found that the most useful area of the technology field to address first is that of communications services. For these services, there will be various devices and systems that are used in forming them. Similarly, for the devices and systems there will be components that shape their format and capability. Figure 2 illustrates how this type of analysis is structured.

This 'tree' approach to the problem is somewhat related to the typical morphological analyses that are discussed in technology forecasting texts.

It was stressed earlier that the separate consideration of communications technologies was only appropriate as an initial stage of the planning process. The likely demands for the communications services identified can only be examined in the light of the broader societal conditions that are postulated for that time horizon. These alternative societal conditions are the 'scenarios' discussed earlier, and the combination of the communications services with the scenarios is done in an analysis that consists of two phases: First, each service is examined within the context of each scenario to

determine the likely degree of acceptance or implementation that would occur under these societal conditions.

Secondly, all the services that pass this first examination (and are therefore candidates for implementation within those societal conditions) are examined one against another to determine likely interrelationships and impacts that may occur within the group.

These two steps are referred to as a 'cross-impact overlay analysis'—the communications services are analysed by overlaying them on the scenarios, and the individual services are examined for the cross impact of one on



FIG. 2. Structuring the communications technology study.

another. In both stages of this analysis, the studies of the various systems and devices and the various components are used to augment the overlay analysis, but the analysis concentrates on the communications services.

This procedure is represented in Figure 3.

The outputs from these various analyses of communications services within societal scenarios are forecasts of the likely levels of implementation and demands for the communications services.

In one of the studies that we have recently completed, the results of these cross-impact overlay analyses were used as inputs to a computer model, which in turn was used to produce the quantified forecasts of demand. The incorporation of such a computer model into the planning programme provides an extremely useful way of examining many alternatives and of identifying and quantifying the effects of changes in any of the subjective assessments that may have been made.

The cross-impact analysis and the format of its outputs are structured so that they fit easily into the computer model that has been constructed.

This has summarized in broad terms the methodology that we have used on a number of communications planning problems. The remaining part of this section discusses in more detail some of the aspects of developing the study of communications technology.

### Developing the communications study

We have already described the 'technology tree' approach that we have found to be a most useful classification and structuring technique. In this section, the approach for developing the relevant pieces of information about each communications technology is outlined.

Because of the range of possibilities that could be studied, even within the confines of a structured 'tree', it is necessary to ensure that the technical treatment of each prospect for communications technology is sufficiently deep to allow informed conclusions to be drawn, and that the research of each prospect is not so detailed as to require large manpower efforts and budgetary drains and result in only a small number of technology prospects being considered. In other words, in each study it is necessary to consciously trade off the detail of individual prospects against the overall coverage of the total package of prospects.

An organization can minimize the inherent risks in such a trade-off exercise by developing a tightly structured research paper on each communications technology prospect. Consistent with the breakdown of prospects for communications into services, devices and systems, or components, the research papers for the 'services' category contain the most detail and the papers for the other two categories contain whatever support material



### 2

Examine the impact of one or another for those services that pass the previous examination



# Short- and long-term forecasting

seems appropriate. Within each research paper, the same structure is used so that subsequent analysis can also be systematized. For example, one set of topic headings for the technology prospect research papers is:

1. Proposition for the Year 2000.

A short statement, identifying concepts and some of the variations that may appear during implementation of the communications prospect for technology.

- 2. Present State of Development. A statement and history of the present stages of development of the technology prospect or the elementary forms of it.
- 3. Known Plans (if any) for this Technology. Statements about any known plans that may exist anywhere in the world relating to the development of the technology—whether the prospect is for a service, a device/system or a component.
- 4. Trends.

Any attempt to look many years (say twenty-five) into the future must also look as far into the past as possible for similar introductions or technology developments. Therefore this section should establish, where possible, information about performance, substitutions of alternative services, times between stages of development in the past, and an identification of analogous introductions or developments of other devices/systems—particularly where direct information on the communications technology being examined is not yet available.

- 5. Technical Problems to be Overcome. This section identifies particular technical problems that need to be overcome before the technology (as the prospect outlined in Topic I above) would be available.
- 6. Likely Implementation Problems.

This section seeks to identify the particular social, political and nontechnical implementation problems that would need to be addressed before the stages of development for the communications prospect could be postulated. Other matters in this section include legal questions, potential organizational constraints, regulatory problems and standardization needs.

7. Probable Costs and Timing of Availability.

As far as possible, costs and availability estimates need to be included in the research papers. Most of the base information included in the papers on a communications prospect relates to its development without considering specific constraints that may apply in the country or region under study. This is consistent with the overall methodology by which the local constraints and effects are considered when the communications prospects are cross impacted with the scenarios. We have found that these research papers are very useful reference documents, even before being analysed in the light of the societal scenarios. Their development draws on a number of forecasting and research tools, including:

- Analysis of expert opinion: Groups of experts in particular fields of interest can be used to identify likely ranges of development and potential implementation and technical problems that may arise. In larger studies, we believe a two-round Delphi study is a useful way of further quantifying the expert opinion. (In our work, we have had a structured approach to the group discussions, but our second round of discussion was with a different group of experts. This innovation to the normal Delphi techniques proved most useful and avoided some of the 'pressures for consensus' that are typically present in Delphi work.)
- Trend extrapolation: Where trends are available, various trend extrapolation techniques can be employed. These include analyses of historical substitution effects, in addition to step function analyses and the usual curve projections.
- Morphological analysis: Where future communications services could take alternative forms and use different component parts, morphological analysis can be used to help clarify and identify the areas on which to concentrate by establishing the system format that appears most likely to occur.
- Literature search: For the various technology prospects being studied, literature searches should be conducted. These often yield authoritative data bases on which to perform trend analyses and with which to guide the group discussions with experts.
- Contacts with world authorities: Direct contact with recognized world authorities can prove most helpful in identifying potential future directions for communications technology. As part of one of our studies in this area, one of our consultants and an officer from our client organization toured some of the major research and technology consulting organizations in the United States, Canada, United Kingdom and France. These exchanges of information and direct contacts proved to be valuable in the particular study.

The research papers should be reviewed by panels of experts from both within the client organization and also from outside organizations (equipment manufacturers, designers, universities and other appropriate groups). These discussions are of particular value in identifying likely timings of commercial introduction and/or availability.

The research papers about each communications technology prospect are then used with each scenario in the cross-impact overlay analyses described earlier. This process also involves elements of subjective assessment, but with the structure of the analysis quite tightly defined. This is very important because it allows repeatability and reproducability of the various analyses.

The outputs from this stage of a study typically include statements about the new communications services—their likely format, their likely level of implementation, their likely interactions (positive and negative) and an identification of any key societal or technical variables that need to be monitored in order to have improved knowledge of a particular communication service's future.

Part of the output from this stage of a study is a diagrammatic representation of the new communications services. An illustration of this is contained in Figure 4. This shows how individual services impact others (+/-) on the connecting arrows) as well as likely levels of implementation for each service over the study time horizon.

### Incorporating forecasts of communications demand

In both short- and longer-term planning, a quantification of results in terms of communications demand is desirable.

In a study of the longer-term futures, such as discussed in this section of the paper, the thrust of the research is on exogenous variables or matters external to the organization offering the services. In other words, the typical study seeks to identify the likely demands for new services if they were made available, and is therefore concentrating on the 'demand' side of the usual 'supply/demand' relationships. In some of our studies we have implicitly introduced 'supply' side considerations via the proposed relative tariffs for the various communications services. For example, an alternative form of supplying a service (via a move to automation, for instance) can be incorporated into the study by alternative tariff/pricing assumptions.

To allow quantification to be made easily, and in order to allow the impact of changes in any subjective assessments to be easily judged in quantitative terms, a computer model of the particular communications industry under study can be valuable. Such a model, static or dynamic, has as its variables some of the key societal and technical factors identified in the other parts of the planning study. New values for such variables and new relationships between them come from the scenario development research and the results of the communications services/societal scenarios 'overlay' analyses.

Figure 5 illustrates the sort of results one might expect for an organization's overall demand, with and without taking advantage of new communications technologies to augment the services they offer.



FIG. 4. Representation of important new communications services in a future characterised by an extrapolation of societal trends scenario (illustrative sample only).

### MONITORING

Our discussion of short- and long-term communication planning has included an identification of those important factors that appear likely to impact particular aspects of the communication future.

The monitoring of these factors and the updating of the study results are an implied feature of the approach and should not be viewed as an optional activity. If there is no monitoring, the value of the 'futures' data base that will have been developed in the course of the study will be quickly eroded because of the rapid changes that are confronting communication planners. Therefore, these changes in technologies, as well as changes in social and political directions, must be quantified and incorporated into an organization's strategic thinking.



FIG. 5. Illustration of typical quantification of demand forecasts.

There are two basic approaches to monitoring. These are to periodically examine new developments in each critical factor that has been identified, and to continuously search for new developments across the whole critical factor field.

There are various ways in which this monitoring can be achieved, particularly in the case of technology variables. These include researching selected journals and establishing contacts with selected organizations around the world who are known to be working on technologies of interest.

A useful way of assessing the likely behaviour of critical factors throughout a state or nation is to conduct 'intervention experiments' with a selected sample of the population. The intended communication strategy is set up on a pilot scale and then its progress and acceptance are monitored. Key factors that have been identified in the analysis phase as affecting the likely implementation of the strategy are in this way tested in practice. Actual rates of acceptance, for example, are determined from such a monitoring approach.

In practice, a study's results suggest different monitoring approaches for different critical elements—be they social, economic or political factors, new technical systems or expected technological developments. Some factors would require more detailed tracking than others.

Once the monitoring programme is established, it should not be regarded as an ad hoc activity; rather, significant resources should be devoted to these tasks on an ongoing basis.

### Conclusions

This chapter has outlined details of short- and long-term forecasting methods that are appropriate to communications technology planning. We have also identified the best ways in which the various techniques can be applied.

In the case of longer-term forecasting, the approach discussed is fairly new to the communication planning field, but it is one that we have recently proved in practice in Australia with two major studies for the Telecommunications Commission and the Postal Commission, in addition to confidential studies for private-sector organizations.

Many of the problems associated with single forecast extrapolative techniques are overcome by adopting the approach we have discussed —using 'scenarios' to describe the relevant external conditions and 'technology prospects' or 'alternatives' to describe the possible technical options. By employing this approach, the uncertainties of long-term forecasts are minimized because attention is focused on the implications of different futures rather than there being a dependence on the accuracy of a single prediction.

Possibly the most significant output of the approach is the development of a systematic monitoring procedure for tracking those factors that are isolated as the most critical.

All of the methods discussed in the chapter make a valuable contribution to the communications planning process. The choice of method and the degree to which it is implemented will vary from case to case, depending on the particular needs of the organization initiating the work, the time frame of the study and the time and budget available for the completion of the work. IV

The economic approach

# Chapter 7

## Economic analysis and communication decision-making

Steven Klees Stuart Wells

The nature of economic analysis

Communication, as a field of study and policy concerned with a particular sector of a system of societal organization, can and should draw on many disciplines for an understanding of the technical and social phenomena with which it is concerned. Although most approaches to date may be characterized more as multidisciplinary than as interdisciplinary, the more integrated effort associated with the latter approach has become of increasing importance when dealing with the interdependent complexities of the modern world. This is especially true for communications systems planners and decision-makers concerned with furthering national and international development. In this chapter we examine those concepts and technologies of economic analysis that we consider most useful for understanding socially appropriate communications systems policies and investment decisions.

The increasing global emphasis in recent years on private and publicsector economic rationality has resulted in decision-makers at all levels being called upon to justify their choices in terms of such criteria as profitability, efficiency, cost-effectiveness, and cost-benefit. Unfortunately, these criteria have been rendered obscure by the rather technical and esoteric nature of much of the economics profession and its associated language. Thus most planners and decision-makers who are being held accountable for economic rationality often have little understanding of its basis.

Furthermore, economics is not a unified social-science discipline. Criteria for economic rationality have grown out of one dominant strain of economic thinking—neoclassical economics. This theory is based on analysis of the role of the private and public sector where most societal resources are allocated through competitive markets. Critiques of this approach are numerous, and at least one coherent alternative paradigm of economic thought exists. It is important for planners and decisionmakers to understand the basis for these critiques and alternatives to neoclassical theory if they are to respond sensibly to demands for economic rationality. This is especially important for the field of communications, in which many of the current debates mirror those within economics.

In the first part of this chapter we begin by setting out the assumptions, logic, and approach of neoclassical economic theory. We then examine how this theory leads to the commonly encountered analysis of economic costs, effects, and benefits for decision-making. In the next section we turn to the question of how neoclassical theory has been applied to national development, which is clearly central to our topic. We then consider some of the criticisms of and alternatives to this mode of analysis. In the second part we then go on to examine how economic analysis applies specifically to communication systems planning and decision-making, and in the third part we offer some concluding observations.

Our intent is not to make economists out of communications specialists. However, in a world where specialization and expertise are the norms, we consider it essential that those individuals who are supposed to manage, consume, and apply fragmented knowledge have a good grasp of the bases, strengths, and limitations of the various concepts and tools they must integrate. It is to contribute towards this end that this chapter has been written.

### THE FRAMEWORK OF NEOCLASSICAL ECONOMIC THEORY

What is usually called 'economics' would be more accurately labelled by any one of a number of expressions—'competitive market theory', 'free market theory', or 'neoclassical economics'. We use these labels equivalently to refer to a body of literature, theory, and applications that traces its intellectual history from the work of Adam Smith (*The Wealth of Nations*, 1776) in the mid-eighteenth century. This body of thought has emphasized the workings of the 'invisible hand' of the market economy. Although primarily directed towards analysing capitalist systems based predominantly on the private ownership of resources, it has also been applied to the workings of socialist systems, with public ownership, that base resource allocation on (public) profit-oriented market operations (e.g. Lange and Taylor, 1938; Lange, 1965; and Vanek, 1970).

Competitive market theory begins with some assumptions about how the economy operates. Chief among them are: (a) that firms who produce goods and services are out to maximize their own profits; (b) that people who consume goods and services are out to maximize their own utility (economist's jargon for happiness); (c) that these profits earned from

### Economic analysis

production, and utility gained through consumption, are unaffected by the production or consumption of others; (d) that there are many buyers and many sellers of each type of good or service, among whom competition is solely on the basis of price; (e) that none of these individuals or firms can affect the market price by themselves; and (f) that there is complete information available to producers and consumers as to prices, market opportunities, and the alternative technologies of production.<sup>1</sup>

Assuming that these conditions hold, a number of conclusions have been shown to follow. Under such a system firms would produce those goods and services most valued by consumers. These firms would produce and sell these goods and services at the lowest possible cost. Human labour and other inputs would be paid at a rate equivalent to the value that they add to the firm's output. To sum it all up, an economy operating according to the above assumptions would not 'waste' anything. The resources society has at its disposal would be used so efficiently that no reallocation of those resources could yield a 'clear' improvement. That is, the condition that neoclassical economists label Pareto efficiency or Pareto optimality would hold: the resources of the society could not be reallocated in such a way as to make even one person better off without making someone else worse off. Since neoclassical economists do not believe that they can make comparisons of welfare among individuals, such a global efficiency criterion of 'no waste' in the allocation of resources is their main emphasis.

Both producers and consumers are assumed to act in their own economic self-interest, which is the basis for their choices. Perfect competition relies critically on prices to act as signals that convert this private self-interest into social gain. In a competitive system no single actor can affect price-price is determined in the aggregate as an equilibrating index balancing supply and demand. The price of a good or service thus reflects both the relative value that consumers place on a good (demand), and the relative value of resources that go into producing that good, given our knowledge of production techniques (supply). Prices are thus viewed by market economists as societal opportunity costs. That is, in a market system they are a measure of the value of the resources that society had to use up or employ in order to produce the particular good or service in question. Given the existence of competitive forces, if those same resources were more highly valued by consumers in another endeavour, then smart profit-maximizing entrepreneurs would bid the resources away from the production of a less valued good or service.

1. Actually (a) and the last part of (d) are not assumptions, but are conclusions that can be shown to follow if the other assumptions hold. See any standard microeconomics textbook for more detail on these and other points in this subsection (e.g. Henderson and Quandt, 1971).

It is interesting, especially for communications specialists, to view this system as a very economical societal information system in which price efficiently embodies all of the relevant information from an atomistic bargaining process (often assumed to be instantaneous) that occurs throughout society (or the world, depending on how far you wish to apply this theory) for each resource input, good and service produced. Consumers only need to know which goods and services make them happiest, and their prices. Producers only need to know the alternative ways resources can be used to produce some good or service, and the prices of those resources. Thus, economists, for the sake of global Pareto efficiency, are interested in efficiency at the micro level in terms of the behaviour of producers and consumers.

This concern provides linkages between economics and communication (as well as with education), since from this perspective communications can affect the efficiency with which producers and consumers behave. Neoclassical economists examining the behaviour of any production unit concern themselves with the ability of producers to maximize profits (since according to the above framework higher private profits lead to higher net social benefits). They usually distinguish between two types of efficiency at this micro level: allocative efficiency—the purchase of lowest-cost combination of resource inputs to achieve that output level which maximizes profits; and technical efficiency—the ability to organize these market-purchased human and physical resources (along with any non-purchased resources) in a manner that maximizes profits. Both types of micro efficiency are necessary for the system as a whole to be Pareto efficient, and both can be affected by communications systems, as we will discuss in the second part of this chapter.

Given the above framework, the role of the public sector is to engage primarily in those activities where the market, if left to its own devices, would not perform efficiently. That is, the main reason for government to intervene in the economy is when any of the assumptions (a) to (f) listed above do not hold. To understand how economists view communication systems planning and policy we must briefly examine the main economic rationales for such interventions. First, the technological nature of certain production processes may exhibit economies of scale, which means that as a firm produces a greater volume of output, its average cost per unit of output declines. If this holds true, as for example in certain aspects of the telephone industry, then one firm would come to dominate the market, yielding a natural monopoly and violating assumptions (d) and (e) above. This is an efficient arrangement in one respect, leading to production at the lowest possible cost (since the monopoly is still assumed to be a profit maximizer). But it is inefficient in another respect, in that the monopoly

### Economic analysis

position of the firm allows it to influence the price charged and the total volume of output produced. Neoclassical economists argue that, in the case of a monopoly, less output would be produced than that desired by consumers, and thus government has a right to intervene in the market place to achieve Pareto efficiency, usually through some form of regulation.

Secondly, all of the benefits or costs of a production or consumption activity may not be completely captured by the producers or consumers involved in the market transaction. That is, a firm's production or a consumer's consumption may yield costs or benefits to other parties. For example, a production process may pollute the environment, affecting individual well-being or the production costs of another firm. Alternatively, the consumption of certain goods and services by one individual, such as education or information, may yield benefits to others. In the case of these externalities, assumption (c) above is violated: there is a divergence between private and total societal costs or benefits, and therefore the private market, if left to the workings of the 'invisible hand', will not produce efficiently. Market-determined prices do not accurately embody either (a) all the relevant value attached to production. Therefore the public sector needs to intervene, usually by correcting market prices through taxation or subsidy.

Thirdly, there may be certain goods or services that are not profitable for the private market to produce and sell, but are still valued by individuals—and thus by society. These goods or services have two salient characteristics: (a) once they are produced, an individual cannot be excluded from receiving the benefits; and (b) an individual's consumption of what is produced does not diminish the stock of the good. Thus, once produced, producers cannot sell such goods to individual consumers. Examples of such public goods usually given by economists include national defence, lighthouses, roads, bridges, and parks. Thus the public sector, to ensure the production of the goods and services most valued by consumers (which is necessary for Pareto efficiency) must intervene by producing them or subsidizing their production.

There are a few other circumstances, relevant to the communications sector, that call for public sector intervention. Neoclassical economists usually assume that there is some role for the government in maintaining socially satisfactory levels of inflation, employment, and economic growth, since the extent to which the competitive private sector can do so is unclear. Additionally, the absence of complete information for all producers and consumers (the violation of assumption (f) above) may dictate that the public sector play a role in information provision.

Finally, there are all the questions that revolve around social equity. The general position taken by neoclassical economists is that they really cannot

make prescriptions about equity, although their statistical tools may be used to examine the impact on equity (e.g. in terms of income or wealth) of various economic policies. Competitive market theory assumes that there is an initial societal distribution of wealth at any given moment and that marketplace transactions are entered into freely by individuals, each of whom benefits from such transactions. Otherwise, they would not purchase a particular good or service, or would not sell their labour or physical resources to a particular producer. Whether the initial distribution is 'fair' in some societal sense is beyond the judgement of economists. They argue that if the society as a whole decides it values a more equitable distribution of goods and services, then the public sector, as an agent of the society, can institute policies designed to bring it about. However, such economists often go on to argue that equity should not come at the expense of efficiency. All the other public-sector interventions mandated by the failure of the competitive market should be based on efficiency considerations; if greater equity is desired it should be treated as a separate concern and arranged for through the taxation mechanisms by which a government finances its activities. This efficiency/equity dichotomy of neoclassical economics is the focal point for considerable criticism, which we will discuss further below.

### COSTS, EFFECTS, AND BENEFITS

Given the framework of neoclassical economics, the competitive private market is left to itself, with each production unit maximizing private profits. Private profits are considered to be social profits, since output and input prices are measures of social value, as long as all relevant benefits and costs are captured within the market transaction (i.e. there are no externalities). That is, profits in this theoretical system represent a signal to the producer that the output they produce is more valuable to society than the societal resources that are used in the production process.

If the private sector is to use profits to guide its activities, what are the relevant criteria to guide public-sector decision-making? How should a nation's government decide which alternatives before it are most socially valuable? The theoretical and methodological procedure devised by neoclassical economists to answer this question is cost-benefit analysis. Costbenefit analysis involves the valuation of the outcomes of public-sector investments in monetary terms and the comparison of these social benefits with the costs of the investment (see Mishan, 1972 for a detailed treatment). If the monetary benefits exceed the monetary costs, the investment is deemed worthwhile, and those investments with the highest net social benefits (benefits minus costs) are the most worthwhile. The costs of any activity are viewed as the total social value of the resources used in that activity. Since prices are measures of social value, resources purchased on the market may be easily costed. However, non-purchased resources should also be costed if they have alternative uses that are considered to be of social value, that is, if they have an opportunity cost. For example, the time students spend in a formal educational system, the time audiences spend in reading, listening, or viewing various communications media, or the time of project volunteers, should all be costed if the time of these individuals has alternative uses of social value. From this economics perspective, decision-makers must consider the total resource-use costs of any activity, not just those they pay for out of their particular budget (see Klees and Wells, 1978, for a detailed treatment of cost analysis applied to education decisions, but generally applicable to communications as well).

Cost analysis alone cannot be the basis for choice among alternative courses of action, without assuming something about or investigating the benefits of each alternative. The benefit of any activity is viewed as the total social value of the outputs produced by that activity. It is possible that some public-sector outputs are sold in a competitive market. But more often they are either sold by a public monopoly, in which case price bears no necessary relation to social value (e.g. public utilities), or not sold at all (e.g. education or government-owned television). Neoclassical economists often attempt to value such outputs indirectly; for example, since education graduates and dropouts take jobs on the labour market, their additional earnings due to education can be viewed as a monetary measure of the social productivity benefits of education, given that wages for human resources, like prices for physical resources, measure social value.

If, as is likely, the outcomes of many public-sector activities cannot be translated into monetary value, economists use cost-effectiveness analysis to compare alternative courses of action. Effects are measured in whatever units are considered relevant and possible—e.g. cognitive achievement, numbers of students enrolled or graduated, police arrests, media exposure, incidence of disease, number of missiles, etc. The alternative that gives the most effectiveness for a given budget is chosen. Alternatively, the course of action that achieves a certain specified level of effectiveness for minimum costs is chosen. Of course, with this approach it is impossible to choose between alternatives that yield different outputs or a mixed ordering on the same outputs (i.e. one alternative yields higher outputs of some kind than another alternative, and lower of other kinds), unless society or its decisionmakers can weigh and thus compare the value of different system outputs. In neoclassical economics, prices act as such weights, which is why such economists favour the use of cost-benefit analysis. Economics actually brings little special expertise to cost-effectiveness analysis, other than the approach to calculating social costs and the suggestion of effectiveness measures that may be associated with social productivity benefits.

Both cost-effectiveness and cost-benefit analysis often involve the comparison of alternatives whose costs and benefits extend over more than one year. Economists bring a unique framework to the evaluation of impacts over time through viewing a societal rate of time preferences as a price. This price is determined by demand based on people's preferences to consume goods and services now as opposed to later; and supply, based on the availability of profitable investment opportunities through which forgone consumption today will be turned into greater future consumption benefits. This social interest rate or social rate of discount can then be used to translate future costs or benefits into their present value. Alternatively, the rate of return of benefits over costs internal to the particular activity can be calculated and compared to that of alternative investments (see Dasgupta, Sen and Marglin, 1973, for a good discussion). Non-monetary effectiveness measures should also be discounted if effects now are more socially valued than effects later (see Jamison, Klees, and Wells, 1978, for an example applied to educational system enrollments, and Klees and Wells, 1977b, for a general discussion).

Both cost-effectiveness and cost-benefit analysis focus on the neoclassical economist's concern with the efficiency of any production process. The selection of the course of action that yields the highest net social benefits (or greatest effects for resources used, provided that society values the effects more than the resource inputs) yields, by definition, an allocatively and technically efficient production process. It also contributes to global, Pareto efficiency if everyone affected by the process benefits. That is, neoclassical economists argue that if the social benefits are higher than the social costs of an activity, undertaking that activity could result in a Pareto improvement, since any possible losers could be compensated and thus everyone affected could be better off (see Rothenberg, 1974, for a discussion of the whole conceptual approach to cost-benefit analysis). Even if losers are not compensated, some neoclassical economists would argue that an activity with positive net social benefits should be undertaken anyway, since redistribution can be taken care of separately.<sup>1</sup>

Such reasoning is closely related to the efficiency/equity dichotomy posed by such theory. As long as benefits exceed costs, there is no necessity

<sup>1.</sup> Actually this perspective is of doubtful validity within such economic theory, since the reluctance of such economists to compare welfare between individuals forces them to advocate Pareto improvements only; any other change would mean comparing one person's monetary benefit with another's monetary cost. This is inadmissible in the theory.

for the public sector to ensure that those who benefit do in fact pay, as the private market does. Efficiency has to do with what is produced (that which is most valued), how it is produced (at least cost), and who receives it (those who value it most), not how such production is paid for. This is why neoclassical economic theory has been seen as applicable to a marketbased form of socialism—the ownership of resources and the distribution of wealth is irrelevant to economic efficiency (except as it may affect motivation). Within capitalist society the financing of public-sector activities is thus seen as generally irrelevant to economic efficiency considerations. As long as the benefits exceed the costs of a proposed activity, one may finance it by making the beneficiaries pay, by making those with the most wealth pay, or by any other means desired from an equity point of view.

Finally, both cost-effectiveness and cost-benefit analysis rest on the ability to determine the causal impact of any public-sector undertaking. Given the difficulties of conducting controlled social experiments of any magnitude, economists rely on quantitative modelling and statistical methods, chiefly regression analysis, for testing their theories about such impacts (see Kerlinger and Pedhazur, 1973; Rao and Miller, 1971).

### NATIONAL DEVELOPMENT

Competitive market theory, based on what is supposed to be the socially efficient 'invisible hand' of supply and demand, is not well suited to examine questions of economic growth. As we have seen, competitive market theory is based on a world where the operation of a relatively few behavioural and structural assumptions yields a system that allocates its resources efficiently. It describes a system in a stable equilibrium and how that equilibrium is restored when forces yield changes in supply and demand. Changes in consumer tastes and preferences lead to changing demand. Changes in the techniques of production lead to changes in supply. In competitive market theory such forces are usually considered to be exogenous. To account for economic growth, neoclassical economists have developed a whole area of study called macroeconomics, which attempts to place the microanalytic, individual unit behaviour perspective of competitive market theory in the context of a set of dynamic forces that affect growth, employment, and inflation.

Their view of development is usually synonymous with economic growth, defined as annual increases in the real value of the total goods and services produced by an economy, generally measured by indices such as gross national product (GNP) or GNP/capita. The source of economic growth is seen by these economists to be generally a function of investment. Some resources are taken away from the production of consumer goods and services now, and devoted to activities such as the construction of machines to be used in future production, or education, research and development. This in turn improves production techniques so as to generate even greater output of consumer goods and services in the future (see Ranis, 1968, for a succinct review of the literature).

Policies after the Second World War reflected the development strategies implied by such theory. Wealthier nations viewed the problem of poorer nations as one of the barriers to capital formation. The poorer nations had plenty of labour available, and a wide range of production technologies was available for choice internationally. The problem was one of technology transfer. To break out of poverty, a poor nation needed to tighten its belt, forgo some present consumption, and invest its resources in these new technologies (this whole approach is quite analogous to that taken to poverty within wealthy nations). Alternatively, an infusion of foreign aid might also be able to break the cycle. The focus of economic growth was thought to be industrial development, and the requisite of industrial development was, at least initially, felt to be investment in physical capital, which was seen as the driving impetus behind the industrial revolution of the past three centuries. In the late 1950s and early 1960s, disappointments with the apparent effects of loans and gifts from rich to poor nations for dams, factories and physical machinery, combined with the development within the neoclassical perspective of theory that focused on the critical importance of investment in human resources to economic growth, yielded a marked concern with the formation of human capital through education and training as a necessary complement to physical capital investment.

From this perspective, rural development and agricultural productivity were of secondary importance. Industry, especially capital-goods industry, was viewed as the leading growth sector of the economy (see Hirschman, 1958). Rural society in low-income countries was viewed as inherently static, economically stagnant and culturally tradition-bound. The solution to the problem of agricultural stagnation was seen to lie, in the first place, in uprooting traditional society, which would happen over time as a consequence of the forces of industrialization, as it had in the wealthier nations. The process could perhaps be speeded through the spread of education and mass communications. Both were seen to be of key importance in remolding traditional values and attitudes, particularly in teaching rural people to perceive their economic self-interest, and in motivating them towards change. Secondly, technological inefficiencies in agriculture could again be solved by direct transfers of technology from the industrialized countries.

### Economic analysis

More recently, a strategy of development based on expanding GNP through investment in heavy industry has appeared to many as increasingly unacceptable. Concerns with system equity, brought on at least in part by rising and often frustrated expectations of the poor, have brought into question the uneven growth policies of the past. Food shortages have in some cases proven too stringent a constraint to permit a continued bias of investment against agriculture in favour of industry (see Schultz, 1965). Even in countries that have not found growth inhibited by food shortages, industry-based growth has frequently manifested undesirable equity and employment affects (see Mellor, 1976). Moreover, even where industrial growth has been most rapid, it has not usually expanded at a rate sufficient to provide employment for an exponentially growing population, let alone a population migrating out of the countryside into the cities. Furthermore, given the relationship between low income and high fertility, a high unemployment rate appears increasingly intolerable if only because of its accelerating effect on population growth (see Kocher, 1973). In the agricultural sector the benefits of technology transfer have accrued in many cases to the larger and more prosperous farmers, while small farmers appear in some instances to be relatively or absolutely worse off than they had been a decade before (see Freebairn, 1973).

The general response to this shift in development perspective among policy-makers has been to focus their efforts more on the agricultural sector, and with more attention paid to specific societal groups that may have received inequitable treatment or insufficient attention, such as the rural poor, the urban poor, or women. A somewhat analogous response within the ranks of neoclassical economists has given considerable attention to agricultural productivity, with some of it focused on poorer, smaller land-holding farmers (see Schultz, 1964). The potential of new technologies, brought on by the so-called 'green revolution', to yield agricultural productivity increases, combined with the problems experienced with direct technology transfer, has brought up the question of the most 'appropriate' technologies (see Hyami and Ruttan, 1971). To a neoclassical economist, the technology of production has always referred to both the physical and organizational aspects of a production process, and improvements in technology are considered the basis of economic growth (brought about through savings and investment). The concentration on appropriate technologies fits in with neoclassical economists' concern with efficiency; the efficiency of a particular production process will depend on the resource-use opportunities and their costs within the particular local environment to which it is being applied.
#### CRITIQUES AND ALTERNATIVES

For the communications system planner and policy-maker to understand the applicability of the conceptual tools of conventional economics (and their methodological derivations), she or he must also understand their limitations. In this section we briefly examine some criticisms of neoclassical economics. We will discuss the alternative framework posited by Marxist economists.

First, we have seen that neoclassical economics depends on certain assumptions about individual behaviour and the structure of the economy to reach these conclusions: (a) that prices reflect social value; (b) that competitive markets allocate efficiently; and (c) that the public sector should be guided by cost-benefit analysis. Since it is clear that no real economic system operates strictly according to the few, but rather stringent, assumptions of perfect competition, a critical question is what happens when the behaviour of the system deviates from these assumptions. Some theoretical arguments (see Baumol, 1972) imply that when even small deviations occur, it is doubtful that the system operates near efficiency. In an interdependent economy one distorted price may create ripples, distorting other prices. In this situation, where prices do not reflect measures of social value, decisions based on monetary analysis of costs, benefits, or profits have no necessary relationship to wise social choice.

This type of argument is used by economists like Schumacher (1973) to extend the quest for 'appropriate technology' beyond the domain of monetary efficiency. Economists have argued for the transfer of more appropriate technologies to developing countries largely on the grounds of efficiency. Analysis of local conditions, markets, and prices is considered necessary to devise the most profitable technology for any given use. Schumacher, however, makes two further points: (a) that the prices of our dwindling supply of natural resources do not reflect their true value to society, in part because of societal short-sightedness and in part because of the relatively immediate profit orientation of large-scale industry; and (b) that large-scale industry has sufficient negative social externalities (in the form of social and technological control, pollution, etc.) to warrant converting to smaller-scale technologies. His conclusion that small-scale technology use and natural resources conservation need greater attention thus rests on his implicit argument that prices, as they are currently determined, offer a poor guide to truly efficient societal resource allocation.

Secondly, competitive market theory assumes that the tastes and preferences of consumers of goods and services (demand) guide resource allocation, yet treats these preferences as outside the influence of economic activities. Communication systems, which can be viewed as an economic

#### Economic analysis

activity that may change individual preferences as much as inform them, clearly contradict this assumption, and the same could be argued for other aspects of our modern economic system (see Gintis, 1974, on this issue). Under these circumstances, consumer sovereignty and efficiency become ambiguous criteria, and the question then becomes whose preferences, and preferences at what point in time, will guide resource allocation decisions.

A related issue concerns the weighting of these consumer preferences. How do these preferences become aggregated? The theory of perfect competition essentially yields the rule of 'one dollar, one vote'. The preferences of those who have the most money exercise the most influence over what goods and services are produced by society. As we have said earlier, to a large extent competitive market theory ignores any such equity considerations.

These latter points raise questions that are difficult for cost-benefit analysis, as at present formulated, to deal with. The economists' concept of 'society' as some abstract entity that receives all the benefits and incurs all the costs of any particular investment activity is called into doubt for three reasons. First, 'society' as an aggregate of individuals has different preferences at different points in time, and these are affected by earlier resource allocation decisions. Secondly, 'society' consists of individuals and groups with unequal power; thus even though total benefits may exceed total costs, some individuals or groups may be hurt by a decision. Thirdly, it is recognized that because unequal power exists, decisions may be made leading to costs that exceed benefits as long as the costs are imposed on a different group than that which receives the benefits (and presumably makes the decision).

The problems discussed above, taken together, form a critique of the basis of most neoclassical economic thought and, consequently, the costbenefit analytic framework that is most commonly used. First, it is not clear that monetary values represent societal values, and secondly, it is not clear what 'society' as an aggregate concept really means. Alternative economic viewpoints have been advanced, perhaps the most coherent centering around the works of Karl Marx and subsequent additions to the theories he initially expounded. This body of literature is more difficult to summarize justly than neoclassical economic theory. Nevertheless, the basics are essential to understanding an important alternative view of the economic role of communications.

Fundamentally, the Marxist economist views resource allocation in most societies as a struggle between competing groups with unequal power. From this perspective capitalist social-resource allocation is seen as a function of the conflict between two broad social classes: capitalists, who own or control the non-human productive resources; and workers, who must sell their labour to capitalists for survival. Although neoclassical economists would say that such a structure is motivated by individual self-interest, Marxists would argue that there is nothing necessarily optimal or efficient about the overall social outcome.

For one thing, Marxist theory argues that although capitalists are motivated to accumulate wealth, they do not seek maximum profit. Capitallists are equally interested in maintaining their social position within a whole system of social relationships. Thus they use hierarchical organizational forms, not because such forms are necessarily efficient, but because such forms maintain social distance between and within classes, and act as legitimation mechanisms for status and wealth differences. They favour capital-intensive technologies (even when not the most profitable technology) because these require fewer workers, and because fewer workers are easier to control. Systematically, this behaviour results in a continuing body of unemployed and underemployed workers (what Marx called a 'reserve army') whose very existence acts as a threat to help control employed workers. Such a perspective explains the persistence of sexism and racism over time, which to neoclassical economists is 'irrational' since everyone could be made better off (i.e. a Pareto improvement) by hiring the best person for each job. Marxists argue that such prejudice is useful to capitalists because it fragments the working class (see Gintis, 1976, for a detailed discussion of the above points).

The capitalist class has had considerably greater power than the working class in the past. This condition is made worse by the greater concentration of power in the present monopoly stage of capitalism. Workers in the structure are seen as alienated from the product of their work, from the work process itself, and in their daily social relations with others. However, this does not mean that the working class is powerless, and indeed major social change from this perspective occurs from class conflicts, brought about by the contradictory forces emanating from within the capitalist structure itself. One major argument is that the capitalist accumulation process is sufficiently contradictory to human desires that workers' consciousness of their situation and its causes will increase, and with it their power (see Ollman, 1973). Most Marxists do not seem to view social change as coming inevitably through economic or historical forces, as they are often caricatured. They in fact believe social change can be achieved by collective political action (again see Ollman, 1973).

Marxists view government in capitalist society as a conscious or unconscious agent of the capitalist (or other ruling) class. Government also embodies the contradictions and class power conflicts of the capitalist economy. The state and its organs (such as education and communication policy) by and large act to maintain an ideological superstructure conducive to the accumulation of wealth by capitalists and the maintenance of classbased social relationships. Thus one could expect that government regulation of the private sector would by and large further capitalist interests (e.g. see Miliband, 1969). In the United States the employment of former industry executives as government regulators of their former industry is seen as a case in point. However, at the same time, one can view the contradiction between the political democracy of some capitalist states and their industrial authoritarianism as again leading to increased consciousness within exploited classes and perhaps as a source of meaningful social change.

The Marxist view of national development is explained within a global structure at present dominated by monopoly capitalism. Capitalists in advanced industrial nations, it is argued, search for profitable investments and the ability to maintain the high-level consumption of luxury goods at home. Thus they expand across the globe to exploit cheap labour and obtain raw materials. The limited industrial expansion and relative agricultural stagnation in poorer countries are seen to have evolved historically from the international relations of the colonial era. These were characterized by unequal power, and hence exchange, between producers of raw materials and producers of industrial capital goods (see Frank, 1969). These relations sustained a distribution of wealth and power within poorer nations that made it possible for élites to adopt patterns of consumption that are similar to those in wealthier countries and are dependent upon imports. These consumption patterns and import and capital goods dependencies distort industrialization in such economies, leading to the production of luxury goods. This in turn yields internal terms of trade and returns to investment that have negative impact on food production in agriculture (see de Janvry, 1975). Thus autonomous national development is seen to be constrained by this centre/periphery model, in which the structure of monopoly capitalism places the poorer nations in a peripheral state of dependency and uneven development, responding to the demands of the wealthier centres.

The Marxist economics perspective described above generally rejects the efficiency/equity dichotomy of neoclassical economics and the assumptions that lead to the conclusion that price is a valid measure of social value. Thus monetary cost-benefit analysis is not usually considered a legitimate criterion on which to base social-policy decisions (although it may be useful to look at the inequities of private gain among social classes). Although the potential of collective action, short of violent revolution, to achieve significant system change is widely argued among Marxists, many would advocate public policies directed towards strengthening the workers' control over and consciousness about production. This perspective results in proposals for industrial democracy, labour-managed farm co-operatives, Freire's (1973) emphasis on 'conscientization' in education, and economic and political alliances of Third World countries. Such strategies can perhaps lead to more equitable and non-alienating societal development.

# Applications to communications systems

We have discussed neoclassical and Marxist economics at some length because we believe that for communications planners to understand the use of the conceptual and methodological tools of economics, they must understand the frameworks within which these have been developed. Indeed, the frameworks themselves are perhaps the most useful tool that neoclassical or Marxist economics brings to the analysis of communications policy and planning.

In what follows we combine our analyses of neoclassical economic theory, its critiques, and the Marxist approach to examine four topics of relevance to communications system planning: (a) the product of communication activities; (b) the industrial structure of communication activities; (c) the role of communication in national development; and (d) the analysis of communication sector decisions. In the first three sections we focus on economic factors that call for public-sector intervention of some sort. In the fourth section we discuss the economic criteria for determining how the public sector should intervene in such cases.

## COMMUNICATION AS A GOOD OR SERVICE

One can view the primary products of communication activities as information, ability and attitude formation or change, and entertainment (although clearly any particular activity may involve the joint production of any or all of those attributes). To the extent that we are examining entertainment activities, neoclassical economists would argue that entertainment should be treated as any other good or service intended for private consumption. That is, the government should adopt a 'laissez-faire' attitude, leaving such production to the private sector. Competitive forces of supply and demand should determine the types of entertainment provided and their price, unless, as we will discuss below, there is something about the technology of production that requires public-sector intervention.

The informational aspect of communication activities, however, has a rather special place within neoclassical economic theory. Competitive market theory, in its basic form, usually assumes that producers and consumers have complete and accurate information as to market production

possibilities, opportunities, and prices. To the extent that this assumption does not hold up in practice, there is a rationale for some sort of publicsector intervention. Thus neoclassical economics, despite its lack of attention to communications systems per se, has devoted considerable attention to the economics of information. Imperfect or incomplete information has been shown, within this theory, as often leading to an inefficient allocation of societal resources (see Stigler, 1961). Furthermore, it is argued that the private market, left to its own devices, is likely to produce an insufficient amount of socially desirable information, and that often such information will be biased (see Lamberton, 1971, for a variety of relevant discussions). Thus public-sector intervention, for example to force truth in advertising through regulation, can be justified. One could argue further that the government should subsidize information dissemination; and perhaps even disseminate certain types of information that the market supplies inadequately, such as information on negative product attributes. Warnings on cigarette packages are an example, but one does not need a public health hazard to justify such activities from this perspective. The negative features of any product are rarely publicized, and the 'perfect information' assumption of the neoclassical framework provides a rationale to encourage more efficient mechanisms for the dissemination of such information than trial and error and word of mouth.

The informational activities of the communications sector obviously include considerably more than information directly relevant to commodity production and consumption. Within neoclassical economics there is some generalized discussion of the social value of a 'free' marketplace of ideas and information, in which good ideas drive out bad ones and in which anyone should be permitted access to any information source they desire (see Coates, 1973). This economics perspective is closely related to the current ideological debates among communication specialists on the international 'free' flow of information, and we will return briefly to this question in our discussion below on communication roles in national development.

To probe more fully into the neoclassical economics view of the product of communications sector activities it is useful to consider the analogy between communication and education, since economists for the past two decades have focused much attention on the latter field of study (see Schultz, 1963, 1971). The conception of education as an individual and social investment in human beings, yielding the formation of 'human capital' that can be a significant factor in production, can easily be applied to communication. Communication systems, like education systems, disseminate information and can produce skill and attitude changes in individuals.

The work of Welch (1970) has been particularly influential in this

area in probing into the mechanisms through which education (and by analogy communication) is supposed to affect productivity. Welch distinguishes between three potential effects of education: (a) the improvement of the skills and knowledge of workers, which makes them more productive; (b) the improvement of managerial ability, fostering the purchase of those inputs associated with efficient, least-cost technologies; and (c) the improvement of managerial ability to take purchased (and also non-purchased) inputs and organize them into the most efficient, least-cost production techniques. Thus education is viewed as increasing both worker productivity and the allocative and technical efficiency with which producers maximize profits. That is, education is believed to produce certain skills, usually undefined but thought to be cognitively-based, that may enhance ability on the job, as well as provide certain aptitudes and attitudes, also usually undefined by economists, that better enable industrial or agricultural decision-makers to obtain, process, and evaluate information relevant to production, and thus to allocate better their available resources. To the extent we believe this to be true for education, it seems likely that it would also be applicable to communication. We will return to this perspective in our discussion of development below.

Welch's analysis offers a human-capital explanation for why individuals might demand, and pay for, educational or communication services. Presumably, individuals would be willing to pay to acquire valuable skills. Although from a neoclassical perspective such private benefits are also social benefits (since higher profits and wages are assumed to be measures of more valued output and higher productivity), the reasoning above, as far as it goes, does not provide a rationale for public-sector intervention. Communication-sector activities that yield private benefits can and should be bought and sold on the private market like any other consumption or investment commodity. However, there are some additional reasons for government involvement with the communications sector.

First, it is likely that the information, attitudes and abilities gained by one individual will confer benefits on other individuals—work associates, friends and family. In economic terms these benefits are externalities present in the private market transaction. As discussed earlier, neoclassical economists argue that when such externalities are present the private market will not yield a socially (Pareto) efficient allocation of resources. If others than those involved in the market transaction receive benefits (or incur costs) then too few (or too many) resources will be devoted to communication activities. In this case there is an economic rationale for the public sector to subsidize (or to tax) those communication activities that involve externalities (see our discussion of some such instances in our other chapter in this volume). Secondly, as is argued with respect to education, there may be publicgoods aspects to the provision of certain communications system services, in the technical sense defined above, such that the private market alone could not provide them. For example, a literate, informed populace may be considered necessary to the workings of a democratic government, and it is unlikely that this type of good will be produced by the private sector left to its own devices. To the extent that the communication sector contributes to the provision of such socially valued public goods, there is a reason for the public sector to subsidize communication activities.

Thirdly, the private sector will not necessarily yield an equitable distribution of goods and services. Thus if society values greater equity then there is also a role for the public sector. Although neoclassical economists would argue that equity considerations be examined separately from the decision to supply goods and services, many such economists recognize that in practice this grand redistribution scheme is unlikely to be sufficient. They therefore argue that equity considerations are relevant to any decision. If communication activities have a significant influence on individuals' access to social rewards (such as through increased job skills) then equity becomes important in determining communication policy. Such concerns are heightened by evidence of the existence of the 'communications gap': empirical studies show that higher socio-economic classes gain more knowledge from communication than lower classes (see Donohue, Tichenor, and Olien, 1975).

The discussion above presents the framework neoclassical economists bring to an evaluation of products of the communication sector and the primary justifications from this theoretical perspective for public-sector involvement. However, this framework is based on the particular economic reasoning of competitive market theory. A Marxist economist takes a quite different view of what is produced by the communication sector.

Fundamentally, Marxist economists see the outcomes of the communication sector in a capitalist society in a manner similar to the way they see the outcomes of such institutions as education, politics, or religion. These are all tied together to form an ideological superstructure which is conducive to the accumulation of wealth by capitalists and the maintenance of their social position over a period of time (see Gintis, 1976, or Carnoy, 1974, for an analysis of education from this perspective). That is, communication in a capitalist society serves by and large to transmit the information, skills and attitudes that are useful to capitalist interests.

From this perspective, the communications effect gap is a functional part of a capitalist system in which larger rewards go to the wealthier classes. Also, from this perspective we may take a quite different view of the product of communication sector activities. For example, where communication media are supported through advertising, a Marxist might focus on the target audience as the product that is being sold, with the advertisers as consumers. This twist in orientation from that of the neoclassicists<sup>1</sup> provides a different insight into the economic role of communication-sector activities. Instead of looking at the characteristics and message impact on the individual behaviour of receivers, it focuses on the power of the sources of media messages to establish and legitimize the rules of the game, to decide the boundaries of communication activities, and to direct the attention of the public towards the interests of the prevailing powers.

#### COMMUNICATION AS AN INDUSTRY

Economics is concerned not only with what is produced, but also with how goods and services are produced. In the previous section we focused on the nature of communication goods and services produced. Here we examine the economic structure of the industries that produce them.

We have seen how, in a neoclassical economics framework, the nature of communication system outputs may require public-sector intervention, even if such outputs are produced in a competitive marketplace. To the extent that such outputs are not produced in a perfectly competitive system, there is an additional rationale from this perspective for public-sector intervention. In particular, as we have explained, the technology of certain communications industries may exhibit economies of scale, in that the cost per unit of output produced decreases as the volume of output produced is expanded. If this is the case, neoclassical economists reason that one firm will come to dominate the market, forming a so-called natural monopoly. As we said earlier, such an arrangement is considered efficient in terms of production at the lowest possible cost (assuming the monopolists are profit maximizers), but the control the monopoly has allows it to produce less than the socially (Pareto) efficient amount of output.

In the communication sector, for example, the telecommunications industry is often viewed in this way, leading to an economic justification for government intervention of some sort, usually in the form of regulation. Such natural monopolies may be on a local or global level, as well as a national one. For example, in small towns or cities, a single newspaper may be able to saturate demand at the lowest possible price. This is as much a natural monopoly as a national telecommunications industry and thus equally justifies public-sector regulation.

<sup>1.</sup> The validity of the capitalist position is at least casually confirmed by looking at the trade journals of capitalist media managers, such as *Advertising Age*.

#### Economic analysis

A natural monopoly or oligopoly (a situation with few sellers) can also develop when the state of the art of production technology naturally restricts the number of competing firms in an industry. For example, in communications the natural scarcity of the broadcast frequency spectrum, combined with our knowledge of how to utilize it, yields a situation where only a few channels of broadcast television may be possible in any given place. Unless there is considerable competition from close substitutes (such as cable television), an oligopolistic situation will result, and the goods and services resulting are not likely to be socially efficient. Here again there is reason for government intervention in the workings of the private market, which should involve more than the sale of frequency-use licences. Such sale merely reflects a societal situation in which the radio spectrum is considered public property; it does not guarantee that private operation will be in the public interest.

It should be noted that the reasoning above applies not only to those communications industries supplying services to the consumer, but also to any intermediate goods or services. For example, if television receivers or satellites are not produced in perfectly competitive industries, then there is an economic rationale for the public sector to intervene. Furthermore, such intervention does not depend on the existence of natural monopolies or oligopolies, but should be a response to any deviation from perfect competition, regardless of whether its cause is inherent in the technology or not. However the nature of the appropriate intervention from the neoclassical perspective will be different depending on its cause. If we are dealing with natural monopolies or oligopolies, then efficient (least-cost) production demands their continued existence. But these need to be regulated so that they operate in the interests of social efficiency. However, if the industry is not competitive for other reasons, related perhaps to its historical development, then neoclassical economists would recommend breaking up the monopolies or oligopolies to achieve a more efficient competitive marketplace.

A related issue concerns the relatively large physical capital investment necessary for many communication industries. In these instances many decisions are 'lumpy' as opposed to incremental, in the sense that once they are made it is not usually economical to change them for a relatively long period of time. Thus in cases where public-sector intervention is called for owing to any of the reasons above, there is a strong necessity for careful planning (see Melody, 1973, for a good discussion of the issues above) since the decisions made today may prove binding for some time to come.

From a Marxist economics perspective, again the issue does not centre on the role of the public sector, since the performance of private profitoriented firms is never considered to be in the public interest. What is of interest is the nature of the forces that direct it and govern the internal and external relationships of the communication industrial structure. The power of capitalists to accumulate wealth is seen to be a primary driving force, in large part determining what communications investments are undertaken. Since capitalists own much of the natural and constructed physical resources of society, it is the returns to these resources that will be maximized, with a consequent bias towards capital-intensive (as opposed to labour-intensive) technologies. From this perspective the emphasis on technology that has often been critiqued stems not from systemic mindlessness but from capitalist class interests that strongly influence resource allocation.

The desire of capitalists to maintain their social position over time leads, from the Marxist perspective, to a set of hierarchical, oppressive, and alienating social relationships within the production process. This is of central interest, since a key component of the Marxist conception of social welfare focuses on worker control of production activities. In the communication industry, as in other industries, there has been some smallscale experimentation with alternative work structures and products in such areas as news services, newspapers, radio and television programme production, and, at the local level, radio and television stations. Public policies that encourage such innovations are worthwhile from this perspective.

### COMMUNICATION'S ROLE IN NATIONAL DEVELOPMENT

As we discussed above, the neoclassical economics view of the development process has changed somewhat over the past several decades. Nonetheless, the primary thrust views development (i.e. economic growth) as the result of a process through which a society chooses to take resources away from satisfying present consumption demands and devotes them to investment in physical and human capital that, combined with the adoption of new technologies, will yield greater future productivity. Within the human capital perspective, communication, like education, can add to the stock of available human capital and thus contribute to national development.

While economists have devoted little specific attention to the role of communication systems (other than education) in development, some communication theorists have given direct attention to such economic development considerations. Their approach to development, however, relies as much on disciplines such as anthropology, sociology, psychology, and political science as on economics (see Rogers and Shoemaker, 1971, and Frey, 1973). Most generally this approach views the essence of develop-

#### Economic analysis

ment as a process of technology transfer, similar to the perspective of neoclassical economists. However, the focus is on an aspect of this technology transfer process usually neglected by economists, namely how such innovations spread. Economics, with its assumptions of free and perfect information and rational, maximizing economic actors, mostly sidesteps this whole question of diffusion of innovations.

Part of this literature theorizes that the adoption of new technologies requires institutional and structural changes in traditional societies to alter the role of the family, religion and culture in a way that will yield the specialization, mobility, and rationality necessary for a market economy and growth (see Etzioni, 1963; Smelser, 1968; and Golding, 1974). Communication systems are seen in this light as quite similar to educational systems, in that they provide information and produce attitude changes that are thought to be conducive to the necessary modernization of societies. The work of Lerner (1957, 1958, 1963) has been most influential, and centres around an elaboration of the relationship between urbanization, literacy and mass-media exposure to development, which is conceived of as some aggregative societal participation concept involving political, economic and psychological manifestations. Lerner hypothesizes that the urbanization process leads to a rise in literacy, which in turn fosters the development of mass media (exposure to which may increase literacy). He argues that the oral-based traditional communication system, characterized as hierachical, status-based, prescriptive, and small-group oriented, must give way to modern systems that can be characterized as more professionally based, descriptive, and mass-oriented. Social development in the form of greater GNP/capita, political participation, and other supposed indicators of modernity then follow from the diffusion of innovative ideas and attitudes via urbanization, literacy and mass-media exposure. Within this framework, the rural, agricultural sector of the economy exists primarily as the population and food source for growing urban areas. This perspective closely parallels that of the economists' framework of industrialbased growth described earlier, with a similar role envisioned for the rural population.

The process Lerner describes can be viewed as operating not only at the societal level but at the individual level. The focus in much recent communication research is on the individual level, looking at the process by which individuals adopt innovations (see Rogers, 1962). This literature has been influenced considerably by supporting sociological and psychological theories, the former looking at stages in the adoption process (e.g. awareness, interest, evaluation, trial, and acceptance), and the latter looking at the operation of role models and personality attributes that foster innovative behaviour (see McClelland, 1961, and Hagen, 1962). Communication researchers in this area focus on the impact on adoption of innovations of sets of variables that describe either types of communication contact or types of individual psychological orientation, such as empathy, cosmopolitan-ness, or the number of newspapers in the home. With the more recent interest in agricultural development, the focus of study for the communication researcher was easily shifted; given an approach that examines the adoption of innovations by individuals, one can easily study the communications system properties and psychological traits that foster the adoption of agricultural innovations (see Rogers, 1969).

This approach to the role of communication in development fits well with that of human capital theorists, and even with early neoclassicists who pointed to the important role of entrepreneurship, but never began to examine the process by which entrepreneurs arise. The more recent focus of economists on the effects of an individual's allocative abilities on productivity and development (see Schultz, 1975) is clearly supplemented by a communication literature that investigates the characteristics of successful allocators, assumed to be those who are early adopters of new technologies.

Nonetheless, neoclassical economists by and large would probably not place much reliance on the results of this communication-research literature, primarily because they would consider the theory and methods used to be inadequate. To separate out the causal impact, if any, of communication on development, one needs to test empirically a model that hypothesizes all the relevant causal factors and their relationships. Thus the correlational approach usually taken by communications researchers, without simultaneous consideration of the host of other economic factors affecting productivity, is not considered by economists to offer much meaningful support for development hypotheses. The regression analysis approach that neoclassical economists would advocate is discussed briefly in the next section.

Despite these criticisms, neoclassical economists would likely believe that communication systems can contribute to economic development; this is reasonable given their hypotheses about the development process. This yields an additional rationale for public-sector concern with the communication sector, since economic growth is not necessarily accounted for by the competitive marketplace.

The Marxist view of national development within a global system dominated by monopoly capitalism was discussed above. Marxist economists argue, basically, that neoclassical theories are inadequate because they fail to take into account the implications of the institutionalized maldistribution of power and wealth within low-income countries and between the so-called developed and developing countries. From this

#### Economic analysis

perspective, Lerner's theory is viewed as a mechanistic manipulation of structural variables with little understanding of the forces that affect urbanization, literacy, and the rise of mass media (see Golding, 1974). The diffusion theorists' focus on audience response to communications systems is seen as based on an unrealistic theory of social change, in which new ideas are the motivating force, and which for the most part neglects the fundamental role of historical, structural system properties.

Instead, Marxist economists argue that much more could be understood about communications system impact through studying the material base within which capitalist interests are particularly influential. The international power structure distorts markets for inputs (including capital, credit and new technologies) in favour of larger, more prosperous and powerful industrialists and farmers. The capitalist desire to reduce labour management problems through the creation of a reserve army of unemployed or marginally employed individuals leads to the diffusion of capital-intensive, labour-saving and land-extensive development technologies.

Again, the focus of communication research from this perspective should be on the sources of media messages, not the audience. The 'free flow of information' ideal is seen primarily as a tool of the powerful to maintain a structure of dependency for the weak. Communication systems, as part of the capitalist ideological superstructure, operate to maintain the vast inequalities between and within nations; any 'development' such a structure yields is likely to be to the interests of the controlling élites (see Schiller, 1971, 1973, 1976). This is not to say that the Marxist framework does not allow for the possibility of public-sector communication (or other) policies to foster positive societal development; it is just considered unlikely. Possibly, progressive policies can result from exploitation by communication planners of the contradictions within capitalism, thereby hastening its demise.

#### ANALYSING COMMUNICATIONS-SECTOR DECISIONS

In the three sections above we have examined how neoclassical and Marxist economists view three critical aspects of the communications sector: communication as a good or service, communication as an industry, and communication's role in development. We have seen that even in a competitive market economy there are some strong reasons for the public sector to intervene in communication systems. Public-sector communication system planning and policy-making are necessary even in a privateenterprise economy, since under certain circumstances a market governed by private profit considerations is not socially accountable. To the extent (a) that information plays a critical role in a market economy, (b) that communications transactions are characterized by externalities, (c) that there are public-good aspects to some communications activities that will not be undertaken sufficiently by the private sector, (d) that monopolies or oligopolies exist within parts of the communication industry, or (e) that the communication sector can foster national development, there is likely to be an important role for government.

Within neoclassical economic theory, the most appropriate publicsector communication policies depend upon which particular deviation from the assumptions of a perfectly competitive market is being corrected. Taxation and subsidization policies are usually recommended to correct private decisions for social externalities and to influence positively national development. Government regulation of private industry decision-making is usually recommended for natural monopoly or oligopoly situations, while breaking up these concentrations of power is called for when such a situation results from other than production efficiency. The actual publicsector production of communications services is usually considered a stratagem of the last resort by neoclassical economists, although within the theory such interventions may be justified when socially valued communication activities will not be undertaken by the private market (the publicgood aspects discussed earlier), or when certain communications activities are especially important to overall economic efficiency and development, and are inadequately handled by the private sector.<sup>1</sup>

According to neoclassical economic theory, the guide to any of these public-sector decisions is efficiency, specified as best as possible by explicit cost-benefit analysis. If the total benefits of an activity exceed its total costs, the activity should be undertaken, since this can result in a Pareto improvement, with everyone made better off (if losers are compensated). Equity considerations can be left primarily to general government financing mechanisms, the prime concern being to design communications financing so as to minimize its interference with selecting the most socially efficient alternative.

Thus the primary issue from a neoclassical economics framework is measuring the total social benefit and costs associated with any communication sector activity. As we have discussed, this is not always easy (even given a belief in prices as measures of social value), especially when outputs are not marketed. Perhaps the simplest situation from an analytic standpoint (but certainly not from a practical one) is when a natural monopoly causes public-sector intervention. Profit maximization by a communication system

<sup>1.</sup> See our other chapter in this volume for discussions of specific communications applications of these public-sector policies.

#### Economic analysis

monopoly will yield less than the socially efficient amount of output produced. Public-sector communication planners, to determine the socially efficient level of production, must estimate the public's demand curve for the output in question, and the cost curve for supplying such output. Given this information, the efficient strategy is to expand production as long as the additional social benefits received (measured by the price individuals are willing to pay, given no external benefits) is greater than the additional social costs incurred (measured by the cost of production, assuming no external costs). Government policies can then be prescribed to force monopolists to expand production to this socially optimal point.

Supply curves can be estimated given a good understanding of the (most efficient) technology of production, but demand curves are usually much more difficult to estimate, as they depend on consumers' tastes and preferences, which are not easily discovered (see Phillips, 1969, and Kahn, 1970, for a general discussion of this standard approach to regulations). It should be noted that from a neoclassical economics perspective, the current trend in communication and other fields towards 'needs assessment' is inadequate. One must know not only what people want, but how much they are willing to pay for it, since the aggregation of this latter aspect is the principal criterion of social valuation. Need is a quite different concept from demand, the latter implying a price-quantity trade-off, and it is unclear what meaning, if any, need has within such an economics perspective.

When communication-sector outputs are not directly marketed, demand curves are especially difficult to estimate. For example, when communication media are paid for, in part or entirely, by advertising revenues, it is quite difficult to discover what consumers would have paid for the output had it been marketed. Some economists attempt to measure such demand curves nonetheless; for example, in the United States, Noll, Peck, and McGowan (1973) attempt to infer the value consumers place on commercially supported broadcast television by analysing the amount individuals have spent for television services in communities where cable and pay television are available.

It is also possible to estimate the benefits of various communication activities by attempting to look at the production side, as opposed to the consumer-demand side, of the economic system. This is done by looking at the impact of communication outputs on industrial or agricultural productivity, profits, or an individual's earnings (the latter two being assumed to correspond to social productivity). Economists have given considerable attention to examining industrial and agricultural production functions, in which the relationship of production process inputs to outputs are specified and estimated. They have also studied individuals' earnings functions, in which the relationship between individual characteristics and their earnings are specified and estimated. With the widespread acceptance of human capital as a source of productivity, education variables are now often included in estimates of such production and earnings functions (see Griliches, 1976; Psacharopoulos, 1974; Klees and Wells, 1977a; and Lockheed, Jamison, and Lau, 1977). Given our argument earlier that communication systems can yield increases in human capital and production efficiency, communication variables (e.g., exposure to general or specific communications media, knowledge gained from such exposure, etc.) should also be included. Regression analysis estimation of a model which simultaneously includes all relevant variables would yield coefficients which supposedly measure the impact of communication variables on productivity. On an aggregate, national level this procedure can be used to look at the impact of communication on economic development (i.e. productivity growth), and would be considered a far more reliable procedure by neoclassical economists than the partial correlational approach usually taken by communication researchers.

When benefit externalities are present (that is when one person's exposure to communication-system outputs positively influences another's earnings, productivity, and/or well-being) it is even more difficult to estimate the total social benefits of an activity, if indeed such is even possible. Nonetheless, when externalities are important, as they may be for certain communication activities, neoclassical economists argue that they need to be included in the consideration of the proper social choice, even if the best estimate that can be made is the subjective judgement of the decision-maker. Such judgements also become important in examining alternative communication strategies that have general public-good effects, such as increased literacy, knowledge, political participation, or cultural appreciation, which may be difficult to measure in monetary terms. In such cases, economists generally argue that cost-effectiveness analysis is the appropriate approach to choosing among alternatives, at least from the perspective of a more limited (than Pareto) view of efficiency. One chooses the alternative that achieves a given objective (measured in some nonmonetary unit) at the lowest cost, or that for a fixed budget achieves the greatest effectiveness. Again, empirical estimation would call for regression analysis applied to a model that included all relevant variables, including those of communication-system interest, on the effectiveness index selected.

Much of the technical literature on modern communication-system alternatives abounds with the words 'optimize', 'maximize', and 'efficiency', usually viewing them as synonymous with the lowest monetary cost per information circuit. However, such engineering efficiency does not necessarily imply economic efficiency (even given that price is a measure of social cost), unless the effects or benefits (intended or unintended) of each

#### Economic analysis

alternative are identical, and the amount of output (information circuits) necessary to achieve minimum unit cost is socially justified in terms of the resulting effects or benefits. Moreover, engineering concepts of system reliability, usually achieved through redundancy of components for use during breakdowns, can be quite expensive (e.g. satellite back-up systems). Decisions as to desired system reliability and redundancy should not be left to technicians, but must be considered by communication planners in terms of their effects and benefits relative to their costs.

As before, the criticisms of this approach to analysing communication decisions are numerous. One aspect that we have underemphasized before is the reliance of cost-effectiveness and cost-benefit analysis on empirical estimation techniques for measuring the causal impact of communication activities. We have mentioned the economists' critique of the correlational methodology of communications researchers, but the regression analysis approach of econometrics also has some very serious failings. Given the difficulties of specifying the complete functional relationships of communication with various 'impacts', and the difficulties of rigorously quantifying the relevant variables, it should be far from surprising that few consistent relationships appear in any social-science literature, at least beyond those associations known from simple correlations. Moreover, one needs to believe in the magnitude of the coefficients (which measure effectiveness or benefit impact) in order to make decisions, not just whether they are positive or negative influences. And there is hardly any consistency between studies on such causal impact measures (see Phillips, Costner, and Fennessey, 1974 for a discussion of more fundamental objections to the standard communication research methodology).

Aside from such methodological questions, there are biases in estimating the costs, effects, and benefits of alternatives, such as those examined by Carnoy and Levin (1975) in a critique of instructional technology evaluations. Some errors in estimation arise out of the inherent problems in projecting cost, effectiveness, and benefit impacts into an uncertain future. But some consistencies, such as the substantial underestimation of the costs of new technologies, are significant. This may result partly from the lack of experience with these technologies, and a consequent likelihood of overlooking some cost components. But Carnoy and Levin also argue strongly that bias often results from the structure of the evaluation process, in which agencies hire researchers who have a framework and personal attitudes which are consciously or unconsciously favourable to these new technologies.

This last criticism of Carnoy and Levin generally stems from a Marxist perspective that views the outcomes of the research and decision analytic process as being as much a function of class conflict and unequal power as are private-sector decisions. Moreover, when prices are outcomes of monopoly capital power, they cannot automatically be considered measures of social value. Thus the whole monetary benefit-cost analytic framework that the neoclassical economist would bring to objectify the examination of communication-sector decisions is useless from the Marxist perspective, except as a tool for analysis of the private gains to different classes of individuals. This is of interest from an equity and political conflict standpoint. Neoclassical economics, which is supposed to simplify decisionmaking by allowing the aggregation of cost and benefit impact by a single measure of social value, no longer can do so. Even cost analysis is meaningless, since total monetary cost is not a measure of total social cost if resource prices are not determined by competitive markets. Thus the Marxist economist generally rejects the neoclassical philosophy that public-policy conflicts result from errors, and can be resolved by and large by the choice of the most efficient alternative, which can in turn make everyone better off.

With respect to communication-sector policy and planning, a Marxist economist would argue that these are unlikely to be effective in remedying conditions that are primarily the result of the historical dynamics of an entire economic structure, such as inequitable national development. Levin's recent examination of this point with respect to educational policy and planning, equally applicable to communication, is worth quoting at length:

To return to our original theme, the failure of the educational and training programs of the War on Poverty to alleviate poverty should not be surprising. Poverty and inequality in income are an integral part of our system of monopoly capitalism...

More generally, educational policies that are aimed at resolving social dilemmas that arise out of the basic functioning of the economic, social and political institutions of the society are not amenable to solution through educational policy and reform. The leverage available to the most benevolent educational reformer and policy specialist is limited by the lack of a constituency for change and the overwhelming momentum of the educational process in the direction of social reproduction of the existing polity...

Rather than face the sobering reality of these limits as mirrored in the failure of past attempts to create educational reforms for solving social problems, we have become captives of our own rhetoric. That is, social dilemmas are continually referred to as the ones that educational policy should address. Volumes are written on how educational policy is the key to their solution, and the discussions are couched in technological terms that assert that we have substantial control over social outcomes. In fact, the discussion of educational policy often has a wishfulfilling aspect in which the rationality of the policy process is treated as evidence of its reality. That is, we are seduced into believing that if we can talk this way,

that it must be true. Thus, planners and policy-makers constantly use such terminology as *change agents, managed change*, and *planned change* as if the use of that language and the logic of rational change imply a control of the change process itself. But, a review of the literature on educational reform and innovation suggests that the rhetoric of reform is probably its most important manifestation rather than the existence of the changes that it claims to produce (Levin, 1977, p. 26–8).

It is probably reasonable to conclude that as long as qualitative educational reform and planning remains at the rhetorical stage, it poses no danger to the established institutions and serves a valuable function for the ruling groups by blaming the lack of justice, equality, and freedom on the failures of education rather than on the military and political dictatorships that serve the interests of multi-national capital and state-socialist hegemony. That is the danger with suggesting that such educational approaches have promise for drastic social change when at the very least such premises are completely untested (Levin, 1976, p. 20).

Again, this does not mean that the Marxist perspective disallows the possibility that public-sector communications policy-makers can contribute to meaningful social change. Levin argues that public policy-making should cross sectoral lines:

What I am suggesting is a new role for educational policy, one in which it considers the basic functioning of the society in calling for changes outside the educational sector as *preconditions for changes in education itself* (Levin, 1977, p. 29).

Moreover, from a Marxist perspective even changes within the communication sector could be socially beneficial to the extent that they exploit the contradictions internal to global and national monopoly capitalist economics structures. That is, monopoly capitalist interests are not all-powerful, and social change is seen to result from basic contradictions within the economic structure, which are manifested in the operations of all sectors of the economy. To the extent that communication policy and planning can foster meaningful consumer-control of decisions on goods and services to be produced, worker-control over the production process, and increased consciousness of exploited groups as to the structural causes of their plight, then there is the possibility that communication can contribute to positive social development. Clearly this implies communication policy-making and planning that go far beyond the 'needs assessments' and token user committees that are commonly called for to legitimate public-policy decisions. Fundamental change in the decision-making structure, away from a professional managerial class and decisions by expertise, towards a more collective, locally controlled decision-making process (informed by expertise) seems necessary.

# Concluding remarks

Communication system policy-makers do not and indeed cannot plan in a vacuum. Data collection and analysis must be guided by an explicit or implicit theoretical and methodological framework. The model-building and forecasting approaches to decision-making of the emerging managementscience field (which in large part has developed from economics) are useless without something to guide their use. Management science tries to avoid this problem by essentially saying to decision-makers, 'use any objectives you consider relevant, we have a grab bag of techniques with which to analyse the extent to which they are being met'. Unfortunately, public-sector decision-makers are usually far from certain about the appropriate criteria by which to judge societal decisions. In the absence of clear theoretical criteria, they will, given the incentive structure within which they operate, often choose ones that are politically expedient. Furthermore, even if clear criteria are available, adequate theory is necessary in order to model and estimate the impact of communication-policy interventions.

Neoclassical economics meets this problem head-on by providing a beautifully complete theory and methodological approach by which to conceptualize and estimate social valuation criteria. Unfortunately, the theory often seems to fail to reflect a realistic description of how economic system structures actually behave. Marxist economic theory seems to provide a more realistic alternative conception of economic systems, but offers a less precise guide to social welfare and only minimal hope for beneficial public-policy structure. Furthermore, its more historical and structural mode of analysis opens it to the criticism of being more subjective and tautological than the seemingly objective, more rigorously quantifiable approach of the neoclassicists.

Nonetheless, communication policy-makers and planners must choose some framework by which to define and evaluate their options. In a very real sense the comparison of the neoclassical and Marxist approaches to economic analysis that we have undertaken in this paper parallels a broader, more basic split that seems to permeate all of social science as well as research methodology. In Kuhn's (1962) perspective this would likely constitute two fundamentally different ways of explaining social phenomena. One focuses on the characteristics of individuals in terms of policy impacts and their resultant social costs, effects and benefits; the other looks at the relationship and properties of structural systems as the focal point of causal impacts.

It seems unlikely that such fundamentally different perspectives can be resolved by an appeal to empirical observation, as positivist theorists generally argue. It is of course a truism that both paradigms have some validity, and both neoclassical and Marxist theories clearly admit that individual characteristics and systemic properties are interrelated. But the two approaches differ significantly in terms of where each sees the focal point for societal change. It is possible that the nascent developments in the field of general systems theory (not to be confused with its much more restricted relative systems analysis), with its emphasis on viewing systems as organic, circuitous sets of relationships (whether stable or runaway), and its sensitivity to the telling question of where one draws the boundaries around the system one looks at, may eventually shed more light on such paradigmatic interconnections.<sup>1</sup>

For now at least public-sector communication decision-makers will have to make difficult choices with little in the way of clear guidance from economics, or any of the other social sciences for that matter. Perhaps the most important aspect of economics for such individuals is its ability to provide coherent, alternative perspectives from which to select potential alternatives, decide how these are to be valued, and even determine the most important policy issues to consider. Although this may be less guidance than policy-makers might like, in many respects the definition of the most relevant alternative choices, about which both economics paradigms have much to say, is the most important component of policy-making. Perhaps the greatest social waste is caused by the failure to consider enough alternatives.

It is important that communication planners and policy-makers should not get too easily seduced by the current fervour with terms such as 'efficiency', 'optimizing', 'maximize', 'cost-effectiveness analysis', and 'cost-benefit analysis'. Furthermore, it would seem that research from a Marxist perspective has received considerably less attention in comparison with neoclassical research. It is likely that decision-making would benefit from a more balanced approach to empirical observation.<sup>2</sup> In our opinion, less reliance on the judgements of expertise and greater reliance on understanding the underlying frameworks behind such expertise is of greatest importance in applying tools of economic analysis to communication policy and planning. This seems of greater and greater necessity in understanding the multitudinous and often conflicting recommendations of experts, and it is towards this end that we hope we have contributed.

- 1. See M. C. Bateson, 1975, and G. Bateson, 1972, for some excellent discussions of the framework, although from a Marxist perspective this literature is also somewhat idealist-oriented and lacks an integrated view of the associated materialist base (see Ollman, 1973).
- 2. Clearly this situation would be no surprise to a Marxist view, since research, like other public-sector activities, is strongly influenced by monopoly capitalist interests (see Gandy, 1975, for a discussion relative to communications), although it should surprise those who advocate a 'free' marketplace of ideas.

## References

- BATESON, G. 1972. Steps to an Ecology of the Mind. New York, Ballatine Books. BATESON, M. C. 1975. Our Own Metaphor: A Personal Account of a Conference on the Effects of Conscious Purpose on Human Adaptation. New York, A. A. Knopf.
- BAUMOL, W. 1972. Economic Theory and Operations Analysis. Englewood Cliffs, N. J., Prentice-Hall.
- CARNOY, M. 1974. Education as Cultural Imperialism. New York, David McKay.
- CARNOY, M.; LEVIN, H. 1975. Evaluation of Educational Media: Some Issues. Instructional Science, Vol. 4, p. 385-406.
- COATES, H. 1973. The Marketplace of Ideas. American Economic Association Papers and Proceedings, May.
- DASGUPTA, P.; SEN, A.; MARGLIN, S. 1972. Guidelines for Project Evaluation: UNIDO Project Formulation and Evaluation Series, No. 2. New York, United Nations.
- DE JANVRY, A. 1975. The Political Economy of Rural Development in Latin America: An Interpretation. American Journal of Agricultural Economics, August, p. 490–8.
- DONOHUE, G. A.; TICHENOR, J. J.; OLIEN, C. N. 1975. Mass Media and the Knowledge Gap: A Hypothesis Reconsidered. Communications Research, January, Vol. 2, No. 1, p. 3–23.
- ETZIONI, A. 1963. The Epigenesis of Political Communities at the International Level. American Journal of Sociology.
- FRANK, A. G. 1969. Capitalism and Underdevelopment in Latin America. New York, Monthly Review Press.
- FREEBAIRN, D. K. 1973. Income Disparities in the Agricultural Sector: Regional and Institutional Stresses. In: T. T. Poleman and D. K. Freebairn (eds.), Food, Population, and Employment: The Impact of the Green Revolution. New York, Praeger.
- FREIRE, P. 1973. Education for Critical Consciousness. New York, Seaburg Press.
- FREY, W. F. 1973. Communication and Development. In: I. de Sola Pool,
  W. Schramm, F. W. Frey, N. Maccoby and E. B. Parker (eds.), Handbook of Communication. Chicago, Rand McNally.
- GANDY, O. 1975. The Economics and Structure of Bias in Mass Media Research. Stanford, Calif., Stanford University Institute for Communication Research. (Unpublished paper.)
- GINTIS, H. 1974. Welfare Criteria with Endogenous Preferences. The Economics of Education. International Economic Review, June, Vol. 15, p. 415-30. -. 1976. The Nature of Labor Exchange and the Theory of Capitalist Pro-
- duction. Review of Radical Political Economy, Summer, Vol. 8, p. 36-54.
- GOLDING, P. 1974. Media Role in National Development: Critique of a Theoretical Orthodoxy. Journal of Communication, Summer, p. 39-53.
- GRILICHES, Z. 1976. Notes on the Role of Education in Production Functions and Growth Accounting. In: W. L. Hansen (ed.), Education, Income, and Human Capital. New York, Columbia University Press.
- HAGEN, E. 1962. On the Theory of Social Change. Urbana, University of Illinois Press.
- HENDERSON, J. M.; QUANDT, R. E. 1971. Microeconomic Theory. New York, McGraw-Hill.
- HIRSCHMAN, A. O. 1958. The Strategy of Economic Development. New Haven, Yale University Press.
- HYAMI, Y.; RUTTAN, V. W. 1971. Agricultural Development: An International Perspective. Baltimore, Johns Hopkins University Press.
- JAMISON, D.; KLEES, S.; WELLS, S. 1978. The Costs of Educational Media: Guidelines for Planning and Evaluation. Beverly Hills, Calif., Sage Publications.

- KAHN, A. E. 1970. The Economics of Regulation: Principles and Institutions. New York, Wiley.
- KERLINGER, F. N.; PEDHAZUR, E. J. 1973. Multiple Regression in the Behavioral Sciences. New York, Holt, Rinehart and Winston.
- KLEES, S.; WELLS, S. 1977a. Agricultural Development and Human Capital: The Impact of Education and Communications. Paper presented at the joint annual meetings of the Latin American Studies Association and the African Studies Association, Houston, Tex.
  - ----. 1977b. Cost-Effectiveness and Cost-Benefit Analysis for Educational Planning and Evaluation: Methodology and Application to Instructional Technology. Washington, D.C., United States Agency for International Development.

----. 1978. Cost Analysis for Educational Decision-Makers. Washington, D.C., United States Agency for International Development.

- KOCHER, J. E. 1973. Rural Development, Income Distribution and Fertility Decline. New York, The Population Council.
- KUHN, T. 1962. The Structure of Scientific Revolutions. Chicago, University of Chicago Press.
- LAMBERTON, D. M. (ed.). 1971. Economics of Information and Knowledge. New York, Penguin Books.
- LANGE, O. 1965. Problems of Political Economy of Socialism. New Delhi, Peoples Publishing House.
- LANGE, O.; TAYLOR, F. M. 1938. On the Economic Theory of Socialism. Minneapolis, University of Minnesota Press.
- LERNER, D. 1957. Communications Systems and Social Systems. Behavioral Science, p. 266-75.
- -----. 1958. The Passing of Traditional Society. New York, The Free Press.
  - ----. 1963. Toward a Communication Theory of Modernization. In: L. W. Pye, (ed.), *Communications and Political Development*. Princeton N.J., Princeton University Press.
- LEVIN, H. M. 1976. A Skeptical View of Qualitative Educational Planning. Paper prepared for the International Institute of Educational Planning, Paris, for a volume on qualitative educational planning.
- —. 1977. A Radical Critique of Educational Policy. Journal of Educational Finance.
- LOCKHEED, M. E.; JAMISON, D. T.; LAU, C. J. 1977. Farmer Education and Farm Efficiency: A Review of the Literature. Unpublished paper, Educational Testing Service, Princeton, N.J.
- MCCLELLAND, D. 1961. The Achieving Society. New York, Van Nostrand.
- MELLOR, J. W. 1976. The New Economics of Growth. Ithaca, N.Y., Cornell University Press.
- MELODY, W. H. 1973. The Role of Advocacy in Public Policy Planning. In: G. Gerbner, L. P. Gross and W. H. Melody, *Communications Technology and* Social Policy. New York, John Wiley and Sons, Inc.
- MILIBAND, R. 1969. The State in Capitalist Society. New York, Basic Books.
- MISHAN, E. J. 1972. Cost-Benefit Analysis. New York, Praeger Books, Inc.
- NOLL, R. G.; PECK, M. J.; MCGOWAN, J. J. 1973. Economic Aspects of Television Regulation. Washington, D.C., The Brookings Institutions.
- OLLMAN, B. 1973. Alienation: Marx's Conception of Man in Capitalist Society. New York, Cambridge University Press.
- PHILLIPS, C. F. 1969. The Economics of Regulation. Homewood, Ill., R. D. Irwin.
- PHILLIPS, D.; COSTNER, H.; FENNESSEY, J. 1974. The Madness in Our Methods: A Three Author Exchange in Four Parts. Sociology and Social Research, April, Vol. 58, p. 225-36.
- PSACHAROPOULOS, G. 1974. A Survey of Earnings Functions. Unpublished paper, London School of Economics, Higher Education Research Unit.

- RANIS, G. 1968. Economic Growth: Theory. In: D. L. Sills (ed.), International Encyclopedia of the Social Science, Vol. 4, New York, Macmillan.
- RAO, P.; MILLER, R. L. 1971. Applied Econometrics. Belmont, Calif., Wadsworth Publishing Co.

ROGERS, E. M. 1962. Diffusion of Innovations. New York, Free Press of Glencoe.
 —. 1969. Modernization Among Peasants: The Impact of Communications. New York, Holt, Rinehart and Winston.

- ROGERS, E. with F. SHOEMAKER. 1971. Communication of Innovations. New York, The Free Press.
- ROTHENBERG, J. 1974. Cost-Benefit Analysis. A Methodological Exposition. In: M. Guttentag and E. L. Struening (eds.), Handbook of Evaluation Research, Vol. 2. Beverly Hills, Calif., Sage Publications.
- SCHILLER, H. 1971. Mass Communication and American Empire. New York, Beacon Press.
- -----. 1976. Communication and Cultural Domination. White Plains, New York, International Arts and Sciences Press.
- SCHULTZ, T. W. 1963. The Economic Value of Education. New York, Columbia University Press.
- ----. 1964. Transforming Traditional Agriculture. New Haven, Yale University Press.
- ----. 1965. Economic Crises in World Agriculture. Ann Arbor, University of Michigan Press.
- ----. 1971. Investment in Human Capital: The Role of Education and of Research. New York, Free Press.
- ----. 1975. The Value of the Ability to Deal with Disequilibria. Journal of Economic Literature, Vol. 13, No. 3, p. 827-45.
- SCHUMACHER, E. F. 1973. Small is Beautiful: A Study of Economics as if People Mattered. New York, Harper & Row.
- SMELSER, N. 1968. Toward a Theory of Social Change. In: N. Smelser, Essays in Sociological Explanation. Englewood, N.J., Prentice-Hall.
- STIGLER, G. 1961. The Economics of Information. *Journal of Political Economy*, June, Vol. 69, p. 213-25.
- VANEK, J. 1970. The General Theory of Labor Managed Market Economies. Ithaca, N.Y., Cornell University Press.
- WELCH, F. 1970. Education in Production. Journal of Political Economy, Vol. 78, p. 35-59.

# Chapter 8

# Financing and control of capital-intensive broadcasting systems

Stuart Wells Steven Klees

How can economists contribute to communication planning? There are several ways, but one of the most important is through the economic analysis of communication alternatives. In this chapter we will examine the implications of alternative financing and control patterns for broadcasting systems. In doing so, we will rely on two general criteria from markettheory economics: social efficiency and social equity. The control patterns that we discuss are based on different forms of ownership of broadcast media.

We will begin by justifying our focus on broadcast media and describing key characteristics of such media for economic analysis. We will then, in turn, define and discuss three forms of media control and three forms of financing. These elements, arranged in different combinations, make up the alternative control and financing patterns we intend to examine. At this point our discussion turns to a definition of the criteria of social efficiency and equity that will be the tools for our analysis. In the final section we present our analysis of three alternative forms of ownership of broadcast media. Within the analysis of each ownership/control pattern we will consider financing patterns as well.

# Capital-intensive broadcast media

The major media for communications in society have been print (journals, newspapers, pamphlets, books), mail, telegraph, radio, telephone, film and television. One may classify these in a variety of ways, such as the means of communication (audio, visual, or a combination of the two) or the relationship among the persons communicating (over time or over distance). A more useful classification for an economist is the relative amount of capital invested in the communication processes. Thinking of communication systems this way helps us understand how such systems operate within society. For example, there are two interesting effects of increasing the amount of capital investment (which is the monetary value of equipment and facilities used) in a communication system. First, the number of people using the system must grow to make the communication process economical: the more people that use a system with a fixed capital investment cost, the lower the cost per single user or message becomes. Secondly, the communication system tends to become controlled by the large-scale enterprises that are able to raise the money needed for large capital investments.

It is quite useful to distinguish between production, transmission and reception components of communication processes, as policy decisions on control and financing alternatives may be very different for each component. The primary factor contributing to the possibility of different decisions is the amount of capital invested. Virtually all of the communication methods mentioned earlier have relatively low capital investment for reception. Radios are usually inexpensive; print media have a price, but no capital investment is required for their use.<sup>1</sup> In fact, the cost of receiving a message is often heavily subsidized. Advertising (or tax-supported government ownership in the case of radio and television) is a major form of subsidization for message receivers.<sup>2</sup>

Message production can also have relatively low capital investment requirements. In telephone systems,<sup>3</sup> the cost of producing a message is negligible, and relatively low capital investment is possible for message production for virtually all mass media (radio, television, print).

Much higher capital investment is required for many communication transmission systems. This makes it more difficult for widespread access to the transmission 'market', making this market less than the competitive ideal. Thus capital-intensive transmission systems are the most costly, and to an economist the most interesting, components of communication systems. In this chapter we will focus our attention on capital-intensive transmission systems. These systems are typically the type one has in mind when one thinks of mass broadcast communication, especially radio and television—which will be the media of interest in this chapter.

<sup>1.</sup> The price of a newspaper is what you pay at the news-stand. Capital investment for reception of newspaper messages is not necessary, although one might count the cost of an armchair if it were essential for full message comprehension.

<sup>2.</sup> Many would argue that in some cases the advertisement is the major message, that it is in the interest of the producer of the message to pay costs of reception, and that any other message is superfluous. We will return to this point in later discussion.

<sup>3.</sup> The telephone has been labelled a 'convivial' technology because it is available when people need it and is controlled by users.

#### **BROADCAST FREQUENCIES**

Open-circuit radio and television broadcasting is affected by a special resource scarcity in communication processes. Most resources are scarce, of course. There are alternative uses for resources in the economy, and decisions must be made on their use. In a market economy a price is established to aid this decision process. This simple notion applies to all resources used in communications. However, broadcast frequencies are a special type of resource: there is a fixed supply. Once an allocation of existing frequencies is made, the supply is exhausted. As a resource, frequencies are very similar to land. Once land has been sold it is possible to resell it, but not to create new land and expand the supply. The analogy with land goes one step further. There are some frequencies that are cheaper to use (require lower investment in capital equipment), and other frequencies that are much more expensive. The whole question of frequency allocation is of international importance, since those who obtain the technology first are in a better position to take the 'best' frequencies, just as those who settled areas took the best pieces of land, leaving latecomers to fill in swamps and irrigate deserts. In some sense, the 'fixed' supply of broadcast frequencies may be expanded with technologies that allow for broadcasts using narrower frequencies.

The fact that frequencies are in limited supply has provided a rationale for public regulation of communication in virtually every country of the world, although the degree of control has varied considerably. The airwaves have come to be regarded as public property. However, one must point out that while there is a limited supply of open-air frequencies, there is no limit to the number of cables one may put in the ground or on poles; and these cables have similar capabilities for transmitting a wide range of 'information' for alternative communication methods. Such systems will be referred to as 'closed systems', as opposed to 'open' broadcasting. Technological advances are leading to smaller cables with higher capacities. While the existence of cables greatly expands the supply of transmission facilities, costs are currently lower for open-air frequencies for communications to widely dispersed populations. Thus 'open' and 'closed' systems both exist, and the issue of frequency allocation remains important.

#### THREE FORMS OF SYSTEM CONTROL

There are three major types of communication system control: private non-regulated ownership, publicly regulated private ownership, and public ownership. Private non-regulated ownership probably does not exist anywhere in the world on a large scale.<sup>1</sup> The other two forms of ownership have been most common: private with public regulation, and public ownership.

Regulation is undertaken in several ways, including price or profit limitations, ensured access to non-profitable users, and content regulation. Public ownership may be realized through government purchase of private stock, an independent public company, government-appointed directors, or publicly voted directors. While each public alternative may have different impacts on efficiency and equity, space precludes an in-depth analysis and we will only discuss government-controlled budget decisions in our analysis of this control alternative.

#### THREE FORMS OF SYSTEMS FINANCING

Communication systems may be financed in three primary ways: user charges, tax revenues and advertising. To some extent the financing alternative may be related to the choice of ownership. For example, the use of tax revenues would be impossible with non-regulated private ownership. Financing is also related to certain characteristics inherent in the system. If communications broadcasting occurs on frequencies for which individuals may purchase relatively inexpensive receivers it would be difficult to exclude these people from benefiting from the broadcasts (and virtually impossible to charge them). User fees under most circumstances would be a particularly cumbersome means of financing.

# Economic efficiency and equity criteria

A general statement of an economic efficiency criterion is that the social value placed on the last unit of any output is compared with the social value of the resources used to produce that last unit of output. If the social value is higher for the output than for the resources used, society would be better off with more production. If the value for the output is lower than that of resources used, production should be curtailed. Economists view prices in a market system as the measure of social value. Thus, the social value of a unit of output is its price, and the social value of resources is the sum of the prices paid for the resources used to produce the last unit of output.

Prices are determined in a market system through the forces of supply

<sup>1.</sup> Although analyses of regulation in the United States have led to conclusions that regulation operates to the benefit of the private owners and that one has a condition to some extent worse than non-regulated private ownership—a cartel.

and demand. Clearly those individuals with more income or wealth have more monetary 'votes' and more control in a market economy. This is regarded by many economists as legitimate, since wealth or income is accrued through the operation of a market system as a payment to the social value placed on the individual's contribution (a resource in the production of some good or service).

Even if the market system did function well and prices were a measure of social value, one would not necessarily desire the most efficient solution, because resources move to meet the demands of those with the power to spend. Instead, one might desire a distribution of resources other than the efficient solution. This would lead to some type of equity criterion for decisions. However, there are numerous imperfections in the market system that cause a deviation from efficiency, such as 'externalities' (positive or negative effects imposed on persons or groups not participating directly in the transaction), costs imposed on the buyer to enter into a transaction through imperfect information or time, large market shares held by relatively few firms, and free-rider problems (the inability to exclude non-payers from receiving benefits from market transactions). Key externalities to be considered in efficiency analysis include use of communications for education and development programmes, diversity of programming, and unbiased news sources.

Equity criteria are usually related to a distribution of resources, opportunities, or outputs. Equity clearly involves value judgements. Is an income distribution with every individual having the same money income more equitable than one with varying amounts of income? It would depend on one's opinions regarding efforts and needs. For communications one may discuss equity with respect to payments in relation to resources and incomes received by individuals. However, a more central aspect of communication-system equity is opportunity—the access of individuals to communication systems both as producers and receivers of messages.

# Patterns of communication-system control: an efficiency and equity analysis

In the three sections below we examine three forms of ownership: nonregulated private ownership, publicly regulated private ownership, and public ownership. Within the context of these types of ownership we discuss the efficiency and equity implications of alternative forms of financing the system. As the method of transmission of communications (open or closed) has potential impacts on efficiency and equity criteria we will separate our comments where appropriate.

#### NON-REGULATED PRIVATE OWNERSHIP

As we have said previously, there is probably no existing major nonregulated privately owned communication system.<sup>1</sup> Yet large segments of many economies are privately owned, and free-market economists argue that private ownership with a profit motivation is preferable because that system most efficiently satisfies consumer desires. It is therefore useful to examine non-regulated private ownership in theory and discuss potential impacts. Additionally, an exposition of non-regulated private ownership provides a useful base point for comparing the other two ownership alternatives.

The form of transmission (open or closed) has significant impact for this form of control, and we will therefore divide our discussion accordingly.

#### Open transmission systems

## Financing and social efficiency

Strong believers in non-regulated private ownership have argued that the government's only role in communication should be in the initial allocation of frequencies. The government would allocate frequencies to the highest bidders. According to free-market theory, this would be the most efficient allocation scheme as the highest bidder would be the individual or group with the highest expectation of profit. In a market system, those firms that accrue the most profit are undertaking activities deemed by consumers to be most socially valuable, and the system would operate efficiently.

Because it is impossible to monitor access to open-air broadcasting, a 'free rider' problem is created: it is impossible for private broadcasters to charge directly recipients of the programme. This leads to only one financing alternative—advertising.<sup>2</sup>

Advertisers usually wish to attract the largest audiences possible, and will design or purchase programmes with mass appeal, thus apparently satisfying consumer (of programmes) demand. The owners of the communication system also have the incentive to satisfy their consumers (the advertisers) by designing programmes with mass appeal, thus increasing their profits (by raising communication time rates) and the profits of the advertisers. As profit maximizers they also have the incentive to produce programmes efficiently (a given product for the lowest cost).

<sup>1.</sup> One of the main reasons for some form of government regulation has been that a public interest is involved in communications and that this interest is magnified by the limited frequencies available for open-air broadcasting. Such public interest is justified in part because the availability of information is central to the efficient operation of a market system.

<sup>2.</sup> Contributions are possible, but unlikely on a large scale.

#### Financing and control

However, there are several factors that can upset the achievement of social efficiency. The primary factor is that no market, in the true sense of the word, exists. The ultimate consumers of the broadcast do not pay. While costs of communication are likely to be transferred by the advertisers to the consumers of their products through higher prices, the two groups of consumers (products and communication) may not be the same. Furthermore, as there is a limited supply of channel frequency, communications producers can easily find themselves in a monopoly or oligopoly position where the price they charge for communication time may exceed the cost of creating that time (transmission and production costs). For social efficiency in a market system one must equate price of an output with marginal cost of creating that output.

A fundamental problem in achieving efficiency exists when there are externalities, that is, divergences between private and social benefits and private and social costs. The profit-maximizing owner is only interested in the private benefits she/he receives and costs she/he incurs. Any other benefits or costs are ignored. However, communication in developing countries has an important potential role in social and economic development. There are a multitude of examples of communications media used for education and information in formal schooling, rural development, literacy programmes, health and nutrition programmes, etc. Under nonregulated private ownership these programmes would be broadcast only if a profit could be made. Advertisers of consumer products would find it beneficial to sponsor or subsidize production of these programmes if sufficiently large audiences of potential consumers could be attracted. This mechanism is consistent with the free-market viewpoint that producers follow the demand of consumers. If sufficient demand were not forthcoming for programmes considered socially desirable, it would always be possible for the government to purchase commercial air time at normal market rates (to buy time at a lower rate would imply regulatory power). Again, one must recognize that the purchase of time may not be at socially efficient prices, since transmission owners are in a monopoly position: they can charge prices higher than costs.

As the above discussion illustrates, true social efficiency is difficult to achieve under non-regulated private control of communication systems. No true market exists because of the difficulty in obtaining payment from ultimate communications consumers. Moreover, the existence of externalities, notably the divergence between private and public costs and benefits, makes it difficult to incorporate broader social benefits in the efficiency calculation. This latter point brings our analysis to further consideration of important externalities (social benefits and costs).

We have discussed the problem of use of communication for education

and development above as an example of the difficulty of incorporating social benefits into efficiency analysis. Other often discussed social benefits of open broadcast radio and television systems include diversity of programming and an unbiased news source. Under private ownership programming diversity will exist only if it is profitable to the owners of communications networks. For example, suppose that the preference of an audience in a given time period for programmes of type A, B, and C is 70, 20 and 10 per cent respectively. If there are only two or three channels all will carry programmes of type A, as this would bring the highest expected audience share for each. If audiences paid directly for programming a very different pattern of programming might exist.

A reasonably unbiased news source might also be difficult to achieve. In seeking audiences, news producers might at a minimum reflect biases of the audience. With advertisers paying for time it would equally be unlikely to find news broadcasts adverse to the interests of the advertisers. Some degree of bias would inevitably enter any news broadcast. Limited time is available and a selection of 'relevant' and 'important' news items must be made. The manner of presentation must also be chosen. Through these selection procedures news stories may be systematically distorted. However, the most likely outcome is the construction of news to attract large segments of the audience, given a financing pattern in which large audiences are the sign of success. Study of national news of one of the major American television networks reveals that virtually all criteria for news selection have to do with audience attraction: a heavy use of news stories by recognized reporters; time limitations on any news story; avoidance of stories with too much regional emphasis; avoidance of stories from other sources if the story was not covered by network correspondents; and a humorous story for closing. All of these criteria arise because an important function of the news is to keep people viewing the news in order to observe the commercials-and to maintain the financing pattern.

## Equity

The two key equity criteria in our analysis are access to message producers and access to message receivers:

Access to message producers: Those firms that control transmission channels have control over production and hence over the content of communication. Communication time could, of course, be sold to others at the discretion of the owners. The potential for social control is enormous. Owners of communication networks could broadcast any political information or support any political candidate they chose. The only moderating influence would be the profit motive. If opposition parties could generate sufficient funds to purchase broadcast time the owner would be expected to sell to obtain profits. However, this is a short-run profit. If the political message or candidate being supported was detrimental to the long-run interests of the communication owners the time would not be sold. Therefore, under a system of non-regulated private ownership it is clear that access to communications time is available to any individual or group who can pay the charges and whose production will not conflict with values of the owners of the communications system. Whether this is equitable or not need hardly be argued. Equity is of little concern within the non-regulated private economy, the important principle being to protect the right of individuals to own and utilize private property.

Access to message receivers: Access to reception would be available to any individual who could purchase a receiver. However, it is not necessarily true that all individuals will live in areas where broadcasts exist. There may be little financial incentive for private firms to service remote, sparsely populated areas, creating an access problem. Highest profits are made in densely populated areas in virtually all parts of the communications industry (average transmission costs per customer or message served are reduced). In the United States economy one may observe the greater number of broadcast stations and the increased competition to mail and telephone services in urban areas.

#### Closed transmission systems

#### Financing and efficiency

Up to this point we have discussed private ownership of open broadcast frequencies for which relatively inexpensive receivers may be purchased. Some of the problems are partially diminished if the final reception is accomplished with a closed system such as cables. For example, although satellites can be used for open-air broadcasting of telephone calls, computer information, radio, or television, the cost of a satellite receiving station is extremely high and distribution to customers can take place over closed lines. Once closed lines are used, advertising is no longer the only financing scheme available. One may directly charge the users of the communication system, as it is possible directly to monitor use of the system. Telephone usage is the most simple example of this as it is easy to measure portions of the system used and the time for which the system was used. For cable-delivered radio or television it would also be possible to monitor the channel viewed or listened to by the receiver, the duration of the use, and the time of day of the reception. There are several financing possibilities for the cable-system owners for radio and television broadcasting, including consumers viewing specific programmes, lease of communications time to programme producers and collection of payment from customers for producers, lease of communications time and payment by advertisers, and purchase of programmes or lease of communication time with monthly charges to customers. In the first three alternatives the incentive is to create programmes that will attract the largest audience. Only the first two alternatives have direct payment by customers and the establishment of a market for communications. In the fourth alternative the incentive is to attract cable subscribers but not necessarily to attract audiences for specific programmes. In any event the incentive of the cable-system owner is to choose the alternative or combination of alternatives that leads to maximum profits.

The effects on efficiency are interesting. The possibility exists that more than one cable system will be installed in one location. If all cable systems are unable to operate profitably then those that fail will have wasted resources in the investment of the transmission systems because the cable and much of the equipment may serve no useful function. It is likely that cable systems will yield local monopolies if there is insufficient demand to support multiple cable systems. This is certainly the case for telephone services, which are regarded by many to be a natural monopoly (as the output, the number of telephone calls, increases, the average cost decreases).

From the standpoint of efficiency the preferable financing arrangement for cable radio or television is the standard market system in which the viewer is charged directly for the programmes watched, much as the telephone user is charged for individual phone calls. In this manner there is no subsidization of programming by one group of viewers for another. More importantly, one does not have subsidization of programming by consumers of products, as is the case with financing through advertising. It should also be realized that with direct charges to viewers it is possible that selling high-priced programmes to small audiences would be just as profitable as selling low-priced programmes to large audiences.

The purposive use of broadcasting to increase social benefits through education and development programming is an important externality consideration. Governments would still have to purchase broadcast time, possibly at socially inefficient prices, particularly when cable-system owners had achieved a monopoly position. On the other hand, the fact that closed systems make it possible to charge consumers of programmes directly raises the interesting possibility of consumer control over the content of development/education programming. Those development/ education programmes that were in highest demand among consumers (in competition with other forms of programming, such as entertainment) could be profitable, in terms either of small audiences and high-priced programmes, or large audiences and lower-priced programmes. In this sense, at least, public participation in the determination of development/ education programme content is facilitated.

This, of course, relates closely to the issue of programme diversity. It is clear then that diversity of programmes could increase as compared to open-air broadcasting because the cable-system operator is in a position to act in a manner similar to the discriminating monopolist. However, in this case, rather than selling the same product to different people for different prices, the system owner attempts to have programmes that meet the demands of several groups and then charges a price for each that will maximize profits.

If programme prices are charged directly to consumers under a closed system, bias on news sources flows from the preferences and biases of consumers. Profit-maximizing cable owners will develop news sources and formats that attract the largest paying audiences. The possibility of diverse news programmes developed to meet demand from smaller audiences exists, of course, provided such audiences are willing to pay relatively higher prices. It seems likely that this would tend to place control of news in the hands of those interest groups with largest financial resources, with the attendant social consequences.

# Equity

Tied to the potential improvement in diversity is a potential improvement in access to the communication system. With multichannel capacity the incentive is for the owner of the cable system to sell excess capacity at low rates. The cable is already installed and the marginal cost of utilizing the system for additional broadcast time is probably low.

There is likely to be less access for reception of broadcasts for individuals living in remote areas. Cable systems are most economical to install in areas of concentrated population. Individuals in remote areas may be able to receive open-air broadcasts and use radios for outside communication but are unlikely to have telephone service or other cable systems.

## Summary

In summary, privately owned, non-regulated communication systems will be operated in the interests of the owners of the system. To some extent the demand for services will be met. However, problems exist with non-market financing under open systems. Social efficiency is theoretically possible but difficult to achieve, primarily because of the absence of a true market of ultimate communication consumers. Social equity considerations fare poorly in open systems.
In closed systems, social efficiency can be achieved under the financing alternative through which consumers pay directly for programmes used. Potential exists for increased programme diversity, including increased consumer participation in determining content of development and education programmes. Unbiased news fares poorly in both open and closed systems, though for somewhat different reasons.

There is potential lack of access to communications under both open and closed systems for persons with diverse interests (though this is potentially less of a problem with closed systems) and for persons living in remote areas.

#### PUBLICLY REGULATED PRIVATE OWNERSHIP

A major rationale for introducing public regulation into a private-ownership market economy is to redress some failing in the operation of that economy. We will examine the impact of a variety of economic tools for regulation on the potential shortcomings of the free-market system that were discussed in the previous section. The main economic tools available are taxation and subsidization, and price regulation. Other types of regulation fall under the general classification of regulation by fiat-rules or policies private companies must follow. These are important but outside the scope of our analysis.

The main failings of the private market system in communications are a failure to account for broader social benefits such as education or development programmes, programme diversity, and unbiased news or political information; an ability to charge a price in excess of the socially efficient price; and a potential lack of access to production or reception for segments of the population.

#### Taxation and subsidization

Taxation is a technique for creating economic incentives to discourage certain kinds of behaviour through the addition of costs to actions that the regulatory agency does not deem desirable. Subsidization is in some sense the other side of the coin. Funds are provided by the government to individuals or corporations to induce desirable forms of behaviour. Taxation and subsidization are usually the most preferred tools of free-market economists, because these are thought to interfere least with the market mechanism. Firms are still 'free' to undertake any decision. However, social costs and benefits which would often be ignored by the private firm are forced into the decision process through taxation and subsidization. These tools are usually quite complex in actual application. We will only discuss the broad implications for the problems we identified earlier. We will not place particular emphasis on the form of communication transmission system (open vs. closed) because the form has greater implications for the extent of the problem than for the type of solution. We first examine four efficiency impacts of taxation and subsidization (education or development programmes, programme diversity, unbiased news or political information and socially efficient prices) and then examine equity considerations (access to message producers and message receivers).

# Efficiency

- Education or development programmes: The simplest means to insure the broadcast of education or development programmes is to purchase transmission time at normal commercial rates. This would not be subsidization. The combined tools of taxation and subsidization may be used by taxing the profits derived from commercial broadcasts, and using the funds both to purchase broadcast time and to produce education or development programmes. The normal decisions of the communication firm are unlikely to be affected, as the taxation merely reduces the profit of decisions in a neutral manner. That is, communication systems owners cannot modify a decision on programmes in order to lower taxes without reducing profits. Even with taxes, a profit is still possible. The only way to avoid taxes is not to broadcast a taxed programme, and this reduces total profit. To alter decisions through taxation, a high tax may be placed on revenues derived from programmes not deemed socially necessary and a low or zero tax on desirable programmes. (A negative tax or subsidy may also be used.)
- Programme diversity: The problem in programme diversity is the offering of similar programmes by different broadcasters at the same time or by a single broadcaster at different times. The regulatory agency need not try to induce the broadcast of any particular programme, but rather provide a wider range of programmes that respond to desires of larger segments of this population (e.g. more programmes of type B or C in the example given earlier). It is possible for the regulatory agency to subsidize programmes that individuals desire. However, as these programmes may merely represent alternative forms of entertainment, it may be difficult to justify the expenditure of government funds or the transfer of revenues from more popular programmes to subsidize less popular ones that would not be broadcast because of the search for mass audiences for higher profits.
- Unbiased news or political information: A bias may be introduced through the intentional disinterest of broadcasters in views they determine to be detrimental to their interests or the failure of persons with opposing

views to have sufficient funds to purchase broadcast time. While taxation and subsidization are potential tools to ensure a reduced bias in broadcasting news or political information, the effect of such government involvement may be to introduce its own bias. Subsidization is feasible if the government is providing funds for the presentation of alternative viewpoints. Shall the decision be in the hands of the broadcaster or the government agency? Both may use highly biased selection criteria. However, traditional economic analysis gives one little insight into this question.

Higher than socially efficient price: Taxation is a cumbersome tool to reduce prices. To reduce the price one would need to charge high taxes on excessive prices and then distribute the funds to the persons who had originally paid the high price. If funds were not redistributed then the high prices would still be in existence but the recipient of the funds would change. Price regulation is probably a preferable tool.

# Equity

- Access to message producers: The solution to the production access problem is not very different from the solution to the broadcast of education or development programmes or unbiased news sources. One either places different tax rates on different broadcasts or uses taxes from more profitable broadcasts to subsidize less profitable ones.
- Access to message receivers: There are two basic problems. The first is the inability of individuals to purchase receivers. In this case subsidies would flow directly to individuals to reduce the price of receivers. The second problem is that the population in a given area may be dispersed to such an extent that it would not be profitable to construct transmission facilities. In this case, subsidies would flow directly to the broadcasters to support construction of facilities.

# Price regulation

# Efficiency

Price regulation is a common tool for regulation of public utilities. The main goal in the use of price regulation is to place a limit on the amount of profit a firm can realize when it operates in a near monopoly situation. The computation of rates to allow for a 'reasonable' profit is often quite complex and involves a variety of schemes for depreciation of capital equipment and calculation of allowable expenses for the price base. A considerable amount of economic literature has been devoted to the topic of the most appropriate way to calculate prices. While recognizing the importance of

this topic we will not discuss it but rather concentrate on the types of price that may be set.

There are two broad categories of prices charged by owners of transmission systems: prices for access to broadcast and, in closed transmission, prices for access to receive. The prices set by the owner to maximize profits may not be the prices that would lead to the most socially desirable outcomes. However, price regulation is not a sufficient tool to ensure the accomplishment of any social goals. Price regulation merely reduces profits. The profit-making business person will undertake any action that is profitable and will charge as high a price as possible for a given demand. A reduction of prices will certainly not cause previously unprofitable ventures to become profitable. Any programmes that were not broadcast or individuals who did not receive access to production or reception, owing to an inability to generate profits from these actions, would still be excluded if prices were lower. One must tie regulation by fiat to price regulation. That is, one must not only set the prices for certain activities but also rule that the activities be undertaken. Otherwise, the only impact of price regulation will be to reduce profits.

By reducing rates and hence profits in communications, one may expect that social efficiency is being increased as prices are reduced to levels one estimates to be consistent with the prices that would be obtained if communications were operated on a more competitive basis. The more likely situation, however, is a transfer of wealth from communication system owners to communication system users. It is a simple point in microeconomics theory to demonstrate that the discriminating monopolist, a firm that sells goods or services in a market where it is not possible for one consumer to resell to another, will sell the socially efficient amount of goods or services. In the sale of transmission time it is fairly simple to prevent resale as the differences in rates will be based on time of day. It would not be possible to purchase broadcast time in a relatively inexpensive time and resell to be used at another time of the day. The problem is a bit more complex in preventing the resale of reception, since one may tape broadcasts or share receivers with large groups of people.

While the discriminating monopolist is selling the socially efficient amount of goods or services he/she is able to extract all of the consumer surplus. Consumer surplus exists when one is able to purchase a good at a price lower than the highest price one is willing to pay. For example, if desks are selling for \$15 and an individual is willing to spend \$20, then that individual would obtain a consumer surplus of \$5. The discriminating monopolist would be able to charge this individual \$20. Obviously, the more a buyer spends for one good the less money he has available to purchase others. The real question for social policy and the extent of rate regulation is determination of the most appropriate recipient of this surplus: communication system owners, corporate advertisers, or individual consumers.

#### Equity: access to message producers and receivers

To increase access to production, whether one has general goals or more specific goals such as increased broadcast of education or development programmes, increased programme diversity, or increased access for alternative news or political broadcasts, one can legislate lower rates for previously excluded groups. However, one must also force the sale of time to these groups at the lower rates to ensure access. Similar arguments may be made with respect to increasing access to reception where some individuals or areas were excluded.

The decision to increase access to production or reception may be based on some broad definition of social benefits and hence result in an improvement in social efficiency. One is forcing a subsidization through diminished profits of communication owners to other individuals. The *quid pro quo* often given by governments in these cases is to allow for significantly higher profits on sales to one class of user (in effect forcing a subsidization of one group of consumers by another).

Another point of co-operation by government regulatory agencies is to guarantee the retention of a monopoly position through the restriction of licences. The interesting impact of simultaneously restricting the number of licences and regulating existing licence holders is that the regulatory agency may, in fact, be contributing to the creation of monopolies and the resulting higher prices.

#### PUBLIC OWNERSHIP

The final alternative ownership pattern for communication systems is public ownership. Several specific organizational structures are possible, including government appointment of directors, direct public election of directors, and an independent public corporation. As each of these is a form of public ownership one may expect the same basic impact on economic criteria, because similar powers may be granted to any of the organizational forms. However, if the alternative structures lead to control by different interest groups the effects could be different.

# Efficiency and financing

We have identified many of the problems that are encountered in unrestricted private ownership of communication systems. However, the mere transfer of power to public authorities does not guarantee the resolution of these problems. If public ownership were to be established in a manner similar to private ownership—in which maximizing profits is the goal—one should expect no difference in impacts. In fact, it is unlikely that incentives would be as strong for a publicly owned system to create profits as for a privately owned system. One would still be financing in a manner similar to the private system: advertising for open-air transmission and a variety of other financing schemes for closed transmission systems. A user tax in the form of annual receiver licences or excise taxes on receiver sales are alternative financing schemes which are possible with public ownership.

It is when the public system is not operated to increase profits that one may expect differences from the private system. One theoretically assumes that a public agency will consider all benefits and costs of any decisions that it makes. However, the underlying incentive often existing in public agencies is to expand the total budget over which one has discretion and to shift the burden of expenses to other agencies. Public officials usually cannot increase personal wealth through increased profitability of their agencies, but they can raise their own status through budget increases. One loses the benefits of profit-seeking behaviour (an attempt to operate more efficiently) without simultaneously ensuring the potential benefits of public control (a wider consideration of social benefits and costs).

Financing mechanisms that have no relation to performance may be used in publicly run systems. To demonstrate the 'value' of a particular programme one does not necessarily match its benefits (as measured by revenues) with its costs. The major mechanism for control of programming by public agencies is budgets. Through allocation of budget funds one is allocating relative emphasis to be placed on different types of programme.

# Concluding remarks

The major problem with communications facilities remaining completely in private hands with no government control is that programme content will correspond to private profit-making and access to communication facilities will be limited to those with the funds to pay for the access. Even if one believes in a non-regulated private economy it is important to realize that its correct functioning is based on the availability of information. Confining the production of that information to those groups with financial resources may result in biased information reflecting a single set of values. Giving complete control of communication facilities to a government agency may be appealing, since governments should be representative of the population. Nevertheless, substituting a public monopoly for private oligopoly may merely shift the type of biased information and single set of values. A middle course may be to adopt some form of public regulation of privately owned communication facilities to correct some of the deficiencies that may occur through the private market system if this gives rise to a lack of socially desirable programmes, the charging of prices that are not socially efficient, and limited access to message producers and receivers. An alternative to public regulation is the allocation of some frequencies to public control and other frequencies to non-regulated private control. Whichever alternative is chosen it would seem important to ensure a decentralization of communication system control to allow for widespread access to producers and receivers.

# V

The evaluation approach

# Chapter 9

# Evaluation strategies for communication planning

John Mayo

Evaluation, the exercise of insight and informed judgement to assess programme effectiveness, is an activity that is assuming an increasingly important role in national communication planning. Rare is the development plan in any sector, be it education, agriculture or health, that does not now propose one or more communication strategies as well as an evaluation unit and some feedback mechanism for monitoring the enactment of such strategies in the field. The need for reliable information, gathered through the systematic application of evaluation research methods, is rarely, if ever, disputed. With the diffusion of media-assisted distance-learning strategies and the encouragement of communication planning at the national level, persons with evaluation research experience have become honoured guests at international forums of all kinds. Their presence is also sought on planning missions and feasibility study teams whose mission it is to help governments analyse their communication needs and formulate innovative policies for the expanded use of the mass media in development.

Yet, in spite of the recognized need for evaluation as a tool for effective system planning and management, uneasiness and suspicion frequently characterize the relationship between communication planners and administrators on the one hand, and evaluators on the other. The groups are in fact often quite sceptical about the purpose and efficacy of their dealings with one another. The scepticism is based on a number of factors. Planners and administrators, whatever their backgrounds or institutional affiliations, are rarely immune from political pressure and the concomitant need to justify their actions in a political milieu. To this group, research that does not legitimize or strengthen favoured policies is often viewed as irrelevant or, worse, as outright threatening. In turn, evaluators are frequently dismayed by the seeming indifference of planners and administrators to the results of evaluation studies. So, too, are they troubled by some of the methodological inadequacies of their field, particularly the lingering pre-eminence of experimental and psychometric research techniques that are believed to be scientifically sound but whose results often seem to have little relevance to the political needs or day-to-day concerns of programme leaders. Increasingly, evaluators are attempting to resolve these problems by entering more directly into project planning and implementation, realizing that 'instead of running alongside the train making notes through the windows, (they) can board the train and influence the engineer, the conductor and the passengers' (Stanford University Consortium, in Glass, 1976, p. 212). Such an activist posture certainly promises more exciting roles for evaluators, and it may, by analogy, even be the most effective way to keep a particular train on track. However, its long-range value to the running of a railroad poses more difficult questions.

The remainder of this essay will examine the range of issues that planners and evaluators must resolve together if they are to share responsibility for the development of communication policies and plans. It is the author's contention that techniques of evaluation cannot be properly understood, applied or improved apart from the varying professional perspectives and political realities that have heretofore tended to separate planners and evaluators from one another and from the people whose needs and aspirations they ostensibly serve. The potential bridges for merging the expectations and needs of evaluators, clients and planners will also be examined in the sections that follow. Appropriate evaluation strategies and techniques will be discussed in a sequence of activities that encompasses the formulation of communication policies and the development of communication systems. The activities include policy definition, system design, system implementation and maintenance, and the assessment of system impact.

# Policy definition

Only recently have governments begun to consider communication policy in a holistic and integrated fashion. The trend towards multisectoral communication needs assessments and planning has been strengthened by the realization from many quarters that development programmes incorporating one or more communication media must be better targeted and be made more mutually supportive if they are to influence their audiences significantly and, at the same time, make the most cost-effective use of material and human resources. Accordingly, in many countries primary decision-making power over the allocation of communication resources is being shifted away from established public agencies such as education, agriculture and health and put under the aegis of newly formed planning commissions and interministerial planning groups. The members of such entities, along with their staffs, fit perhaps better than any other group the emerging profile of what a communication planner is supposed to be.

The individuals acting as communication planners are customarily drawn from a variety of social sectors and represent a broad range of academic backgrounds. Some have substantial administrative experience and expertise in particular areas such as education, health, family planning or agriculture. Others lack content specialization but possess special skills and training in economics or management. Relatively few, if any, have had any direct managerial experience in communications *per se*, although most have been drawn from sectors that may have used the mass media (usually print and/or radio) in the past. Planners' impressions of what can be achieved with the media are often coloured by what they buy at the newsstand or receive at home in their leisure time.

The cross-sectional perspective demanded of high-level decision-makers presents some clear advantages as well as some disadvantages. On the positive side, planning-commission members are usually not expected to champion the interests of one particular ministry. Their allegiance is more often to the highest executive authority—king, president or prime minister. Furthermore, the analyses and recommendations made by planners are expected to supersede traditional ministerial lines. Only in such a manner, it is argued, can national needs be objectively reviewed and some set of priorities established among them. In ideal terms, such a strategy implies that a fresh look will be taken at a range of social issues or problems simultaneously and that resulting communication policies and projects will not be restricted by the desires or experience of a particular sector.

Despite the breadth of their responsibilities and status, members of national planning groups or interministerial commissions rarely have the mandate to execute or follow through on their recommendations. For this they are wholly dependent on the leaders they serve and, ultimately, on government bureaucracies that must either individually or together implement whatever policies the leaders endorse. This fact constitutes a critical check on both the assertiveness and potential innovativeness of planners.

Another problem is time. Because policy planning is an activity of potentially high political cost and benefit, the procedures involved must frequently be compressed in order to guarantee the most visible and immediate impact possible. Background analyses may be cut short and the number of alternatives under consideration reduced, all in the interest of announcing a new policy or reform at some critical political juncture. The need to legitimize and enhance existing political authority is therefore an operating condition that constrains communication planning by sharply imiting the policy options which can be identified and evaluated in any detail.

To cope with such pressures and to compensate for their relative inexperience with communication issues, planning groups characteristically turn to outside agencies both to help assess their needs and to propose the appropriate communication techniques and technologies for meeting them. All too often, however, the envisioned collaborative basis of such studies breaks down, either because the advisers involved are regarded as representatives of outside interests or, what is more likely the case, the planners themselves simply do not have the time to devote to in-depth, analytical work. Unfortunately, what results under such circumstances is the appointment of a token, lower-level counterpart team that has neither the competency nor the authority of the planning group. This, in turn, causes the advisers to work in isolation, returning eventually to present policy and project recommendations that embody reasonable state-of-the-art judgements of what a particular country should consider doing. However, such recommendations may or may not reflect the needs and political realities of the host country.

Another problem, and one often related to the participation of outsiders in the preliminary stages of communication planning, is that the whole process assumes a technological bent. The very definition of social needs or problems may be prejudiced by the availability of some communication technology and the eagerness of a bilateral or multilateral agency to see that technology installed. Under such circumstances, planners may feel pressured either to adopt an extremely attractive loan and technical assistance package for a need they have only partially recognized, or risk losing what appears to be an extremely attractive investment opportunity. The frequent rush to accept a new technology and its trimmings is closely related to the time and political pressures cited above.

Given this set of pressures and constraints, what function can evaluation play in the preliminary stage of communication planning? What kinds of information can best serve the individuals identified as high-level communication policy-makers?

Because senior planners generally spend more time overseeing major sectoral assessments and adjudicating competing sectoral demands than they do actually designing new systems, their primary research and evaluation needs during the policy definition stage centre on the criteria for selecting among competing priorities. A heightened awareness of alternative goals and strategies and their relationship to communication resources, both existing and potential, is the bedrock upon which realistic future policies can be formulated. Therefore, to optimize communication planning at this preliminary but pivotal stage, evaluation activities must concentrate on enhancing awareness of the relationship between development objectives and (a) the alternative communication strategies for reaching such objectives, (b) the performance characteristics of alternative technologyassisted systems and (c) the variety of resources that have been mobilized in support of alternative communication strategies in the past. A thorough awareness of such alternatives, their relative strengths and weaknesses, can help planners avoid the repetition of senseless errors that have marked this field in the past.

#### DATA SOURCES AND RESEARCH TECHNIQUES APPLICABLE TO POLICY DEFINITION

At the preliminary stage of communication planning, information is required that: (a) summarizes the state of existing communication and development efforts within different sectors, and (b) broadens the planners' horizons by making them aware of what is going on elsewhere. Existing national experience is often not adequately recognized or appreciated, and new policies are formed without regard to it. To remedy this situation and to provide an objective assessment of what is possible for the future, existing communication projects should be reviewed. Even if formal written evaluations are unavailable, testimony may be taken from project administrators and funders to determine what policies seem to have worked out best over time. Lessons learned in one sector, education for example, may prove to be instructive to the planning of communication policies for agriculture, health, and family planning.

Although such investigations may employ many of the same techniques commonly associated with needs assessment, at this stage they will not approach the depth of analysis or specificity characteristic of single sector planning or project development. Assuming that communication planning remains multisectoral at this stage, the level of the analyses will also remain superficial, relying on secondary sources instead of detailed performance data from the field. However, if the focus of effort for whatever reason has been previously narrowed to a single sector, the depth of such diagnostic studies may be increased proportionately.

Often existing government priorities will limit or expand the number of sectors in which communication planning activities can be instituted. During the 1960s, for example, many Third World nations were interested primarily in expanding school enrolments and in upgrading the quality of instruction at both primary and secondary levels. This trend was encouraged by a variety of international agencies, which in turn provided capital and technical assistance to launch educational television projects in a number of places, including Colombia, Niger, American Samoa, El Salvador and the Ivory Coast. In more recent years communication planning has shifted away from school-related applications of media towards educational strategies focusing on the information needs of rural people. The logic of this shift has derived as much from changing political winds as from careful analyses of social needs.

The goal of identifying the best possible policy options in the preliminary stage of communication planning is served also by a careful evaluation of experiences in other parts of the world. This can be accomplished through a variety of means, including a review of other nations' feasibility studies and project reports, personal visits and syntheses of existing summative evaluations and cost-effectiveness studies. Together such sources can help planners articulate goals and guidelines based not only on national considerations but also on international experiences.

Case-studies of existing communication projects from a decisionmaker's perspective may be particularly useful to planners. The essence of a case-study is its attempt to illuminate a decision or set of decisions: why it was taken, how it was implemented, and with what results. By marshalling evidence from existing projects, planners can gain valuable insights into how to frame their own investigations and, at the same time, avoid duplicating the mistakes of others. Of particular value to planners are casestudies that document the manner in which projects responded to problems. Such concerns highlight the fact that implementation and problemsolving abilities of programme administrators will likely be as responsible for success as the thoroughness of the policy-planning effort. In John Montgomery's (1974) words:

Planners still take much less notice of administrative and behavioral dimensions of change than of the expected advantages of the economic and technological systems they wish to introduce. Given the limited capacity of the social sciences to appraise the means and predict the consequences of government intervention, it may seem unwise to impose on planners the additional burden of analysing program implementation. But one explanation of the more celebrated failures of planning units is inadequate consideration of precisely this element (Montgomery, 1974, p. 96).

The product of this preliminary stage of planning and the research incorporated within it will be increased understanding of needs and resources in a variety of sectors and some ordering of priority among them. Upon such a base, more detailed feasibility studies can be mounted with the aim of actually instituting one or more new communication systems.

# System design

Once broad sectoral needs and resources have been reviewed and some priority established among them, high-level decision-makers customarily delegate authority for the actual design of new communication systems and projects to individual ministries or interministerial planning groups. The logic of this shift of planning responsibility back to established government ministries is based on the fact that only within such established centres of authority can enough trained and experienced people be found to design, staff and carry out new programmes. The degree to which career officials in various ministries are sincerely committed to existing policies and procedures and therefore somewhat resistant to innovation is often not appreciated by the high-level planner identified in the previous section.

The allegiance and active participation of various levels of established government agencies in the enactment of new communication policies are among the most difficult challenges facing a nation's political leadership. The difficulty of the task is compounded if the communication policy to be implemented requires interministerial co-operation. Only through the measured application of incentives and sanctions, of carrots and sticks, can the parochial loyalties and the natural inertia and scepticism of existing power centres be overcome. The formation of interministerial planning groups is one method customarily adopted to deal with this problem. The purpose here is to enhance the representativeness of the planning effort and to preserve the organic nature of the system envisioned by the leadership.

The challenge of winning the allegiance of disparate government bureaucracies during this stage of communication planning is relatively simple and straightforward compared with that of eliciting the participation of local people in whose name such planning is begun in the first place. To expand the framework for planning beyond the safe haven of government officialdom and 'experts', both native and foreign, is to introduce problems and delays which most planners find intolerable. First, there is the problem of representativeness. Who is to speak for the eventual recipients of the envisioned communication service? Secondly, there is the problem of consensus. Even if representatives of the potential client group can be identified, do they know what they want or really need? Thirdly, there is the problem of efficiency. Will not costs and delays be increased too much by trying to clarify wants and needs? While such questions raise difficult ethical as well as methodological issues that must be of concern to planners and decision-makers, avoidance of them has been far too common. In practice, it has proven easier to plan for people than to wrestle directly with their grievances, needs and aspirations.

The vulnerability of poor people to natural forces and to various economic and political institutions over which they have very little, if any, control is perhaps the fundamental justification for planning communication systems in their behalf. Yet, for a variety of reasons, the communication systems that have been instituted in most countries have not appreciably alleviated the vulnerability of oppressed or disenfranchised groups. Neither have they enhanced people's control over the forces that determine major aspects of their lives. They have failed to do this because they have too frequently ignored a fundamental tenet of good planning, namely that participation in the decision-making process enhances one's commitment to the outcomes of that process. If planners are intent on designing communication systems that envision a broadly based public commitment to social action, they must pay as much attention to the means by which their designs are developed as to the components of those designs. A participatory approach to communication planning is particularly essential when decision-makers hope to raise the consciousness of the audience members and thereby 'animate' them to undertake development activities at the local level (cf. Hall, 1975).

Participatory development strategies cannot simply be decreed from above, they must grow incrementally and be based on the raising of understanding and the clarification of needs at the local level. Correspondingly, evaluation activities that are undertaken to expand the planners' knowledge of local conditions and needs prior to the definition of some new communication system must also stimulate local interest and initiative if they are to lay the groundwork for subsequent citizen involvement. In other words, the links between evaluation and action must be forged early in the formation of a communication plan if it is to generate the local energies necessary to make it effective.

# DATA SOURCES AND RESEARCH TECHNIQUES APPLICABLE TO SYSTEM DESIGN

Whereas high-level communication policy-making and sectoral assessments depend to some extent on census data, school enrolment and attrition figures, health clinic utilization rates and comparable statistics from other sectors, the decision to proceed or not with system design and the development is rarely based on commissioned field studies. Political pressures surrounding the initial planning process often dictate that shortcuts be taken in order that policy initiatives may be announced and implemented as quickly as possible. The urgency for decision is not necessarily a bad thing at the pre-investment stage; it focuses the attention of high-level planners, however briefly, and it establishes a momentum that is frequently lost in detailed multisectoral analyses that invariably seem to drag on for years and frustrate all participants in the process. A much greater danger exists when the gathering of detailed information from the field to guide system design is also telescoped or ignored to put the plan into operation. For sound planning to proceed at the system-design stage, a wide range of needs assessment and diagnostic activities is usually required.

Although planners may intuitively know a great deal about their countrymen's needs and expectations, field research conducted with the potential audience or client groups can reduce paternalism and provide useful information to guide the design of any new communication system. Needed at this stage is information on a broad range of topics related to potential clients: their physical location and patterns of movement, the homogeneity of their culture and language, their educational and literacy levels, and their experiences with existing development programmes and communication media. Sociometric studies can help provide information about the potential client communities: the position of different individuals and classes in the community, the social subgroups, as well as their communication and leadership patterns. Such data are particularly vital if interpersonal communication channels are to be relied upon to reinforce broadcast messages.

A variety of research methods exist for gathering this kind of background information. The most widely used is the field survey, which enables the decision-maker to discern population characteristics and social trends from statistical analyses of data gathered from a sample of the potential client group. The level of care taken in the design and administration of such surveys generally determines the reliability of the information obtained and its validity *vis-à-vis* the planning of new communication systems. The costs involved in mounting field surveys, particularly those that enhance the participation of potential client groups, may seem high, but they are actually quite low when compared with the costs likely to be incurred in having to redesign a system should an initial plan prove unworkable.

Whether conducted by means of written questionnaires or personal interviews, the field survey is an important diagnostic tool that can improve communication planning by providing decision-makers with a wealth of critical background information as well as useful baseline data to facilitate subsequent evaluation of system effectiveness. However, surveys present serious methodological and ethical problems. In inexperienced or insensitive hands, they can undermine the very communication links that planners need to forge with their prospective clients in the field.

Critics of standard survey techniques question both the purpose and methods of such research. In the pursuit of some 'objective' reality, they argue, researchers often impose a reality of their own. In their appeal to researchers to develop less obtrusive measurement methods, Webb et al. (1966, p. 1) argue convincingly that:

Interviews and questionnaires intrude as a foreign element into the social setting they would describe, they create as well as measure attitudes, they elicit atypical roles and responses, they are limited to those who are accessible and will cooperate; and the responses obtained are produced in part by dimensions of individual differences irrelevant to the topic at hand. To offset the biasing and potentially exploitative nature of survey research, steps can be taken to involve elements of the community in the design and conduct of field work. In this way, local people as well as planners can gain some insight into local needs and the obstacles to change that must be overcome. This does not require that local people oversee the details of the entire process, or that they cease being the 'object' of such research under all circumstances, but rather that they, too, benefit directly from the process in terms of increased self-understanding and control. If increased self-awareness can be retained as a genuine objective of all planning research activities, it is far less likely that the results of such research will be distorted or used to manipulate people later on.

A specific strategy for eliminating the forced choices and response sets often associated with standardized survey instruments is for the planner to remain open-minded and to try, through whatever research approaches are adopted, to illuminate the people's own perceptions about what kind of communication needs should be met. A greater reliance on unstructured interviews, participant observation and other methods commonly associated with anthropological research can help to bring out the personal context and significance of people's attitudes on given social issues. When such techniques are applied by well trained field research staffs, they can elicit responses that are 'spontaneous rather than forced ... highly specific and concrete rather than diffuse and general . . . and selfrevealing and personal rather than superficial' (Selltiz et al., 1976, p. 318). They can thereby increase the planners' sensitivity to the experiences, values and hopes of the client communities as well as the client communities' sensitivity to the organizations that are planning new communication services in their behalf.

Needs assessment evaluations can provide only part of the information planners require to design effective communication systems. Also needed, but frequently overlooked or underplayed, are reliable data on the resources available to initiate and maintain new systems. Frequently inventories must be made of the material and human resources currently being devoted to communication activities of one kind or another. It is common to find no centralized records of the amount of local effort or foreign assistance involved in the operation of existing communication projects. This hinders the design of new systems because the planner is forced to make rapid and rather haphazard estimates about what resources are available. The diagnosis of existing resources may also be impeded by the tendency of administrators to withhold vital information in the belief that to release it may lead to the removal of resources from established programmes and services. The more innovative a new programme is perceived to be, the greater will be its threat to the status quo and the greater may be the tendency of existing programme administrators to withhold information pertaining to ongoing operations. Although there is no sure counterbalance to the problem of obtaining reliable estimates of resources available to launch a new communication system, the genuine participation of all sectors, along with strong political support from above, can minimize the obstructionism of existing power centres.

Effective means for evaluating the quantity and quality of existing communication resources are also required during the system design stage. On the hardware side, planners need to know what kinds of equipment (production, transmission and reception) are already in use. How old is such equipment? What is its maintenance record? How compatible is it across different installations? Equally important are assessments of the facilities in which such equipment is being used. What are existing use patterns in schools, local community centres or wherever the programmes are received? Are there marked differences between urban and rural utilization rates? How reliable is electrical power and how fast is it likely to become available in areas currently lacking service? The answers to such questions can provide planners with detailed information to guide their choices on where, when and how fast to implement new systems.

In addition to the hardware elements of existing and envisioned systems, planners also need detailed knowledge of the programming and managerial talents available to them. More often than not the administrators of new programmes will have to be recruited from existing projects. Estimates of such talents are hard to derive, for they are invariably based on intangible impressions and intuitions. If a system is judged to be operating poorly, the tendency is to blame its leaders. In some cases, however, poor performance may be more attributable to some design characteristic of the system itself, to inadequate fundings or to extraneous pressures. Evaluation strategies at this stage have to be sensitive to such alternative explanations of project performance.

Two other concerns are likely to play a major role in the diagnostic research outlined above. The first is training. Planners must not only be aware of what resources are available in different sectors, they must also prepare estimates of what has to be done to upgrade and/or transform existing resources in line with new system requirements. Virtually all new communication systems incorporate training strategies of one kind or another. Therefore, in addition to assessing performance levels, planners are customarily called upon to judge how much training will be needed and in what areas. The need for this information may expand the frontiers of the research considerably, requiring planners to make assessments of the training capacities of both native and foreign institutions.

Costs constitute a second overarching concern of planners at the system

design stage. Each component of the design must be evaluated in terms of its potential cost as well as its potential contribution to overall system effectiveness. This is no simple task. Too often, cost projections are minimized or deflated in the effort to justify a new communication system. Under such circumstances cost accounting is left to project managers and is rarely scrutinized until the system itself is well under way. Such post hoc analysis may provide relatively little information on what is the most economical way of reaching programme objectives. As Block et al. (1977, p. 53) noted in their case study of India's Satellite Instructional Television Experiment (SITE):

If different policy options for an operational system are articulated before the experiment starts, and designed into the project, separate costs can be tallied for each alternative . . . research may also help establish what the likely relationship between the costs of running a small-scale experiment and the costs of a much larger project will be.

Given the political exigencies, traditions, and bureaucratic pressures under which communication planners customarily operate in the system design stage, it is little wonder that they are often unable to evaluate fully a range of alternative strategies for achieving programme objectives, that they often must cut short their assessment of available resources, and that they must often forgo the opportunity to engage in participatory planning methods with potential clients. Yet, it is the attention paid to these areas that may ultimately determine the success of their planning efforts. If not forced to bend unduly in the direction of paths already chosen, researchers can facilitate sound communication planning by identifying a range of alternative strategies and by providing estimates of the costs likely to be incurred in implementing them. If researchers are expected to provide such information, planners must resist taking shortcuts that would speed up system implementation, but might risk the adoption of an inappropriate and/or unworkable system. In turn, researchers must agree to focus their efforts on the planning process and emphasize those issues over which planners actually have some control. When effective communication and a shared sense of purpose can be established between planners, administrators and researchers before operations begin, there is a much better chance that their relationship will endure and be productive as the system itself grows.

# System implementation and maintenance

Once a decision has been taken to implement a new communication system, the information needs of programme planners and administrators customarily become much more immediate and specific. The centre of activity now shifts away from high-level planning groups to the professional project personnel who bear direct day-to-day responsibility for the installation and control of all system-related activities. Such individuals frequently have not participated in the formation of system goals, yet they are the ones who ultimately will be held accountable for the achievement of such goals.

The pressures on programme leaders and creative people during the start-up period are intense; time is short and many of the activities called for in the studio and in the field are innovative and therefore unfamiliar. Furthermore, the environment in which practitioners work is often charged with high expectations and the necessity of living up to someone else's inflated promises. For these reasons, research and evaluation strategies based on anything but the urgent needs and timetables of practitioners are likely to be ignored.

Pragmatism and ingenuity are perhaps the two most important qualities that programme leaders must demonstrate during the initial operational phases of a system. No matter how well thought out the original communication plan and no matter how elaborate the objectives spelled out during the policy definition and system design stages discussed above, administrators are frequently left in the dark with regard to how much emphasis they should put on a specific component of a communication system or what proportion of their resources they should allocate to the attainment of a given objective. Such dilemmas can only be resolved over time, and successful communication project leaders seem to be those who are able to retain an experimental attitude during the preliminary operational stage of a project. The abilities to diagnose problems quickly and accurately, to keep options open, and to engage in a certain degree of trial and error policy-making are also required if administrators are to determine which strategies are likely to work in the field and which need to be reformulated or abandoned altogether.

Unfortunately, evaluation activities undertaken within most communication projects to date have not adequately taken into account the constraints under which programme leaders operate. Evaluators have examined conditions before and after a project in the attempt to make judgements on the extent of change. Rarely have they succeeded in discriminating between those elements of a communication system that functioned well and those that did not. Nor have they worked out reliable methods for relating individual system components to system outcomes. Although extremely difficult to design and carry out, such evaluations are potentially of immense value to administrators who are responsible for allocating resources across a variety of media and among a range of supplemental services and materials at the local level.

If evaluators are to exert a positive influence on the implementation and maintenance of a communication system, they must be willing to enter into the daily operations of a system and share responsibility for its successes and failures. Their allegiance to the system must be beyond doubt. Above all, the results of evaluation efforts at this stage must be immediately translatable into programme alternatives and concrete policy recommendations. In short, evaluators must be willing to become action-oriented and responsive to emerging programme needs.

The lovalty and orientation of the evaluator to an emerging communication system is a necessary but by no means a sufficient condition to guarantee that preliminary data from the field will be taken seriously either by the creators or the managers of that system. There is almost always a range of professional and interpersonal obstacles to be overcome, as the experiences of Sesame Street and India's Satellite Instructional Television Experiment (SITE) illustrate. In each of these innovative programmes, working relationships between the producers and evaluators were strained initially and had to be revised over time. Part of the problem stemmed from different notions about what the programmes were supposed to achieve. Without an established frame of reference and an agreed set of objectives, it was almost impossible at the outset for the producers of either system to agree among themselves, much less with their colleagues in the respective evaluation sections. From the evaluator's standpoint, the programme producers often seemed preoccupied with the aesthetic quality of their work without reference to objectives or to the connections between their work and desired changes in the intended audience. The producers, on the other hand, initially felt that the notion of accountability and the emphasis on cognitive, affective and behavioural objectives were undermining their artistic creativity and freedom. Such concepts also flew in the face of much of their production training.

In the vast majority of communication projects installed to date, production and evaluation staffs have gone their own ways, thereby minimizing the impact of field data on operations.

The Sesame Street and SITE projects are examples of projects that overcame the initial friction between programme planners and evaluators. Over time there developed within the Children's Television Workshop and the Indian Space Research Organization, the patron organizations of Sesame Street and SITE, common reference points and eventual agreement over what specific programme series were supposed to accomplish. Such agreements may have been the most valuable attribute of both systems from an administrative standpoint because they enhanced performance as well as the cohesiveness and self-confidence of the personnel involved.

# DATA SOURCES AND RESEARCH TECHNIQUES APPLICABLE TO SYSTEM IMPLEMENTATION AND MAINTENANCE

A wide variety of evaluation strategies can be carried out to provide the kinds of information practitioners find useful to the development of their programmes. How such strategies are categorized is not as important as the rationale behind them, that is, the desire to provide information that is relevant to the allocation decisions administrators must continually make in implementing and maintaining new communication systems. Anticipating such decisions and being able to provide the right kind of information at the right time are the mark of a responsive formative evaluation unit. Listed below are some of the main formative research strategies implemented by evaluators of communication projects in recent years. Although the strategies are discussed in the order they are customarily pursued, once initiated such approaches are frequently carried out simultaneously. This trend reflects the organic nature of communication planning and the continuing need to replenish and, where necessary, redesign messages as well as other components of a particular system.

# Pretesting and simulation

Rare is the communication plan that does not call for some increase to be made in the amount of information available to some sector of a society or to the society at large. Often, the information called for is of a fairly detailed kind and is related to the development activities of one or more government ministries such as education, health, or agriculture. Accordingly, the production of content and of discrete messages is stepped up. The messages may be of various kinds, verbal as well as non-verbal, and they may be organized or 'packaged' in a variety of ways. No producer, no matter how expert or experienced, can be sure that the messages he or she produces will have the effect intended on an audience unless they have been tried out on members of that audience. This is the justification for pretesting.

Whether disseminated via face-to-face meetings, training sessions, or the mass media, messages must fulfill certain criteria if they are to influence an audience in a desired way. First, they must attract an audience's attention and then hold it long enough to communicate something. Although

utilization strategies tied to reception of messages at the local level can enhance attention, at least for a short period, the vast majority of development messages must also be intrinsically appealing if they are to hold an audience's attention over time. This is particularly true in information campaigns directed at scattered rural audiences for whom organized reception is not practical. Secondly, the messages must be comprehensible. If they use vocabulary, references or communication conventions that are unfamiliar to the audience, risks are increased that the audience will either fail to understand or misperceive major points of the message. Under such circumstances, very little if any learning or behaviour change will result. Similarly, if a series of messages or programmes is not internally consistent or reinforcing, comprehension will be diminished. Finally, the messages must move the audience to action if they are to achieve given development objectives. Such action is frequently very difficult to achieve because of various environmental, economic or social constraints. Failure to anticipate and deal with such constraints in the design of messages can undermine the effectiveness of the communication system as a whole. For these and many other reasons programmers must determine as much as they can about the quality and impact of the messages they are producing before such messages can be disseminated with confidence.

Various techniques have been used to pretest messages before distributing them to an audience. Peer ratings are a natural and effective way to judge whether or not messages fulfill the expectations of fellow producers and are comprehensible to them. If the messages are not comprehensible to this group, the odds are slight that they will be comprehensible to a wider audience. The attractiveness and motivating potential of messages can also be rated by outside experts, but here judgements must be considered only partially valid. The levels of understanding and the aesthetic sensibilities of a programme producer or media critic are likely to be far more acute than the average audience member's, rendering the former's assessment of message appeal or potential impact only partially useful as a guide to future programme development.

One way programme formats have been pretested without incurring excessive expense is by presenting prototypes to representatives of the potential audience, either at production headquarters or in the field. The evaluation unit of the Children's Television Workshop, for example, developed a variety of studio-based techniques to pretest programme formats for both *Sesame Street* and *The Electric Company*. Such measures included physiological as well as cognitive and attitudinal indices of message appeal and comprehension. The information gathered allowed television producers to reformulate programme formats and modules more closely attuned to the tastes and viewing habits of their young audiences. The evaluators of SITE tried to carry the *Sesame Street* model one step further by pretesting their programmes in rural Indian villages. To do this, they transported video-tape recorders to the field to reach potential audience members in as natural a setting as possible. Regrettably, the logistical requirements of moving cumbersome and expensive studio equipment into the field undermined the naturalness of the test and, one suspects, the validity of the results as well. However, the strategy was worthwhile, for it did call attention to the need for simpler pretest strategies that could simulate as much as possible the viewing conditions that SITE broadcasts were likely to encounter in the villages.

#### Pilot testing

Rare is the communication plan or system that has simply sprung into full-scale operation. When this has been attempted, as in the case of Samoa's educational television system in 1964, major frictions and disturbances have resulted. In the Samoan case, the strain on the traditional elements of the island's school system brought about by the inauguration of classroom television in twelve grades at one time was too great. Teachers resisted the innovation, forcing programme administrators to reconsider various elements of their plan as well as the rate at which that plan would be implemented. The Samoan example, coupled with others in different parts of the world, has resulted in much more gradual approaches being taken to the implementation of new communication systems. These days communication projects generally begin with a pilot phase during which major components of the system are checked out in the field. The logic and techniques of pilot testing are based on the premise that when experience with a new system is relatively low, the number of people exposed to that system must also remain low.

From an evaluation standpoint, what usually distinguishes the pilot testing phase of a communication system from the pretesting phase described above is the range of activities that are examined. Pretesting activities are primarily message-oriented, the intent being to improve programme planning and production. During the pilot testing stage, the arena of concern broadens to include not only the appropriateness and impact of individual programmes, but also the support structures and subsystems that eventually will determine their effectiveness in the field.

It is not uncommon for a wide range of evaluation studies to be inaugurated during the pilot testing phase of a new communication system. The research techniques involved are essentially the same as those used to gather feedback data and, eventually, to arrive at summative judgements about a system's overall performance and value. Such techniques (which may include cognitive tests of learning, surveys of participants' reactions, observations in the receiving communities or classrooms, as well as analyses of the overall cost-effectiveness of system components) are critically important to managers who must determine how fast to expand operations without sacrificing quality or straining their resources beyond acceptable limits.

The relationship between pilot testing, evaluation, and decision-making is a critical one given the propensity of communication plans to call for new and often complex administrative arrangements, involving a variety of programme elements and jurisdictions. Traditionally, evaluators have focused their attention on cognitive and attitudinal outcomes of communications projects, ignoring administrative or managerial issues. Yet, when projects or systems fail, it is rarely because they have been poorly designed. More often it is because their leaders have been incapable of either recognizing or dealing with crucial problems early enough. Small problems magnify and, in turn, disrupt operations. Although not all managerial problems can be identified during the pilot period, evaluation at this stage can identify deficiencies and, at the same time, provide useful information to guide both the redesign of system components and the manner of their implementation.

#### Feedback

All well designed communication systems try to maintain as close a contact as possible between the people producing messages and those receiving them in the field. Through frequent and purposeful monitoring of the receivers' learning, attitudes and behaviour, administrators are better able to detect and respond to problems. To a great extent, the strength of any communication system rests upon the sensitivity of its feedback mechanisms and anyone concerned with improving system performance must pay special attention to the means of monitoring an ongoing system and to the means of implementing changes once data from the field have become available.

A variety of feedback strategies have been developed to measure the effectiveness of communication systems once they have been implemented on a broad scale. Some have focused on the learning of programme content, others on audience ratings of specific programme series of courses, and still others on the mechanics of message transmission and reception. In El Salvador's Instructional Television System, for example, attitude surveys were administered to a random sample of teachers to check on their reactions to the teleteaching of individual subjects. The results of these surveys provided useful information about the strengths and weaknesses of the various production teams. In addition, members of the project's evaluation unit conducted an extensive programme of pupil testing to monitor the level of basic skills learning and achievement in the context of El Salvador's educational reform.

Despite the variety of feedback mechanisms implemented in the El Salvador system, the evaluators were not satisfied that they had provided the production teams with the kinds of information the latter needed to improve their performance on a short-term basis. The evaluator's frustration during the implementation stage of the ETV system was summarized in the following way (Merino de Manzano et al., 1971, p. 2):

As the project expanded, it became clear that the production teams needed more kinds of information from the schools, information that would better guide the day-to-day tasks of lesson development and presentation. Neither final exam grades nor year-end attitudinal data were found to be particularly helpful in pinpointing the concepts within specific content areas that were giving students the most difficulty or, on the other hand, the concepts that were being mastered with relatively little trouble. Similarly, more information seemed to be needed to guide the production teams in the most appropriate teaching techniques for their subject specialities.

The Salvadorean experience is instructive. Most communication systems implemented in the past ten years have established evaluation units to monitor system performance and to provide feedback of various kinds to guide decision-making. Yet, while the information provided has customarily permitted administrators to pinpoint problems, only rarely, if ever, has it provided enough data to illuminate concrete solutions to those problems.

A brief hypothetical example helps to illustrate the dilemma. Suppose that comprehension of a particular topic is found to be low or that an audience seems to be losing interest or failing to take recommended follow-up action to a series of development-related television broadcasts. Where does the problem lie? Is it in the design of the message, the adaptation of that message to a television format, the pace and clarity of the programmes, the management system, the availability of local resources required to take follow-up action at the local level, or other, pre-existing constraints of an economic, political or social nature? One or all of these factors could be to blame. To arrive at an answer, evaluators would probably have to probe the problem in a number of ways and with a more varied set of tools than they have been accustomed to use.

The formative evaluation strategy employed in Nicaragua's Radio Mathematics Project comes closer than any yet initiated to providing the kinds of feedback that can be used to solve as well as identify problems. It does so by combining a rigorous weekly testing programme with in-depth observations of teachers and students in the schools. The rationale behind such an intensive and, one must assume, expensive approach was explained by Jamesine Friend, the project's leader, in an essay entitled 'Closing The Feedback Loop' (Friend, 1976, p. 4):

Although we rely heavily on the weekly tests for information, there are many aspects of the learning process for which testing is not a suitable tool for investigation. The tests really only tell us about the cognitive outcomes of a sequence of instruction and nothing about reactions of the children and teachers to the instructional process. Do children understand the instructions? Do they participate in the drill? Do they have enough time to solve the written exercises? Do they sing the songs? Laugh at the jokes? Enjoy the riddles?

Regular classroom observations made by trained and interested observers are the only reliable source of information on these questions  $\dots$  We observe every lesson in several classrooms. Six to eight people observe in three or four classrooms every day. Each of these people carries an 'observation sheet' on which he makes comments as the lesson unfolds. To help the observer focus attention on the specific topics covered by a lesson, a different observation form is designed for each lesson.

The pretesting of prototype programmes, the pilot testing of system components in the field, and the regular monitoring of system performance through the observation and testing of programme users are the essential components of a diversified programme of formative evaluation. To be useful to decision-makers, each of the strategies must be adapted to local conditions and be tied to the range of resource allocation decisions that planners and administrators are required to make. However, in the final analysis, the utility of a particular communication system's formative evaluation strategies rests as much on the ability and inclination of its leaders to base their decisions on information from the field as it does on their ability to authorize the gathering of such information in the first place. On this point, the psychologist Donald T. Campbell, himself a practitioner as well as a critic of evaluative research techniques, has drawn some crucial distinctions between what he calls 'trapped administrators' and 'experimental administrators' (in Weiss, 1972, p. 220). 'Trapped administrators' are those who have so committed themselves in advance to a particular policy or programme that they cannot really afford or use honest evaluation. To such individuals, information that challenges in any way the efficacy of some aspect of a system is seen as threatening and potentially destructive. This group demands that evaluation activities reinforce decisions already taken, policies already enacted. In contrast, 'experimental administrators' strive to maintain a problem-orientation towards their work and constantly seek alternatives for improving the systems they direct. Above all, they rely on evaluative research to improve

their operations either by making existing practices more efficient or by illuminating new practices worthy of trial. It is the openness and flexibility of the 'experimental administrator' that permit evaluation to play a genuinely formative function in communication planning.

# Assessment of system impact

In addition to improving day-to-day operations, decision-makers must also from time to time make cumulative judgements about the long-range impact of their policies. They need to determine to what extent their investments in communication infrastructures and services make a difference to the users of such services, and ultimately, to society as a whole. They may also wish to determine whether the communication system they have developed represents the optimum expenditure of resources or whether some other policy might have accomplished the same objectives more effectively (either in less time or for less money or both). These kinds of comparative cost-effectiveness and cost-benefit issues are at the frontier of what has come to be known as summative evaluation-evaluation that purports to ascertain the impact and social utility of a particular programme. To the extent that future communication planning efforts are contemplated, reliable information on the cumulative impact and efficiency of a particular system can be extremely important to decisionmakers. To the extent that such information exerts some influence on future planning, either in the same country or elsewhere, it may also be considered formative in character. The realization of this fact has led planners and researchers alike to recognize 'the formative function' of summative evaluations.

Even if planners have acknowledged the need for and are willing to sponsor research on the long-range impact of their communication policies, evaluators are often left with the dilemma of what to evaluate. Their dilemma is magnified in many instances because the original objectives of the system were ambiguously defined. It is not uncommon for programmes directed at audiences in Third World nations to have multiple goals all subsumed under umbrella concepts such as 'creation of a new man' or 'integrated rural development'. Such concepts reflect the aspiration and idealism of planners who frequently must embellish their policies with highly rhetorical arguments. But such concepts, if not operationalized during the planning process, provide very little guidance to evaluators attempting to arrive at yardsticks for actually measuring programme effectiveness.

Another problem facing the summative evaluator is identifying the audience or audiences for whom the research is intended. High-level decision-makers, administrators, external aid providers and programme recipients all have a stake in programme results, yet each may have a different agenda and a different set of questions to be answered through summative evaluation. Inevitably, compromises are arrived at; a limited number of objectives are specified and the group actually conducting the research usually decides which questions shall ultimately be addressed.

The complexity and scale of most communication systems require that difficult choices be made at the outset of an evaluation with regard to the depth and breadth of the analyses to be undertaken. There are no hard and fast rules to guide evaluators in this area and the choices made inevitably reflect the budgetary as well as the political constraints outlined above. Summative evaluations often get off to a bad start because researchers have not gained an adequate understanding of what a particular system is expected to do. This lack of definition may be related to the problem of vague or conflicting objectives mentioned above, but it may also be traced to the tendency of outside evaluators to impose their own models or preconceptions about the nature of a given system. Such an imposition can be just as threatening to the derivation of useful generalizations as holding a project too rigidly accountable to its formally stated goals. To avoid oversimplifying reality or relying on too narrow or short-term a definition of success, evaluators must work with planners and administrators to understand the rationale as well as the substance of the latter's policies. At the same time, they must remain open-minded with regard to original objectives and be prepared to deal with the evolutionary character of most communication systems as well as their possible side effects and unintended consequences.

# DATA SOURCES AND RESEARCH TECHNIQUES FOR THE ASSESSMENT OF SYSTEM IMPACT

Various approaches have been proposed by this author and others to guide the collection of data in the evaluation of communication systems (McAnany, Hornik and Mayo, 1973). These encompass: (a) verification of system activities, (b) analysis of system effectiveness in absolute and relative terms, and (c) determination of system efficiency. Each of these approaches, along with the methodological problems they pose, is discussed briefly below.

# Verification of system activities

For communication systems that require efforts to be initiated at various levels (national, regional and local), an evaluator may first wish to verify

that all anticipated system elements and activities actually exist. An example is Mexico's *Telesecundaria* system, where secondary schooling is provided for rural children via television. Before assessing the quality of learning the evaluators had to confirm that the television lessons were actually being produced and transmitted on time and that local teachers were adhering to the broadcast schedule. The common experience of evaluators arriving unannounced at a school during a scheduled broadcast and finding the pupils still in recess or occupied with some completely extraneous activity underscores the necessity of this level of evaluation. In media systems purporting to serve adult populations, the audiences are characteristically less visible and less cohesive than the typical student group. The monitoring of local reception patterns in such instances is necessary to determine whether or not the programme has in fact been reaching those for whom it was originally intended.

# Effectiveness

Once the evaluator has established that a particular audience is being reached, subsequent questions may be raised. Drawing upon the objectives of the evaluation, studies may be initiated to test whether desired outcomes have been obtained. In most evaluations, success is still defined in cognitive terms: students mastering new concepts, adults learning to read, potential parents learning about new birth-control practices. Increasingly, however, other kinds of measures have been introduced to ascertain how audiences reacted to the programmes they received and whether or not their attitudes and expectations changed in line with system goals. In rare instances observational measures have been developed to determine if audience behaviour has been affected in any lasting way. Behavioural changes are the most difficult to evaluate because of methodological problems as well as the time limits imposed on any evaluation. In the case of media-assisted birth-control campaigns, for example, one can ascertain if essential information has been communicated effectively and if people have been inclined to use it, but a number of years may be required to determine whether or not such a campaign has had any lasting impact on the birth rate in the target area.

Although effectiveness of a communication system can be measured in numerous ways, most summative evaluations have relied on one or more of the following three criteria: impact on a target population, impact on a target population vs. a non-target population, or performance to some predetermined standard. Returning to the example of a project teaching secondary subjects by television, success on the first criterion might consist in students simply being able to improve their scores on a test administered at the beginning of the year and then again at the end of the year. Such a before-after design can record the level of achievement of pupils who have studied with television, but it cannot isolate the impact of televised instruction *per se.* To do this, evaluators often try to measure success in comparative terms. Did class A, which studied with television, outperform class B, which did not? The validity of such a comparison depends on the evaluator's ability to be on guard for the familiar novelty effect and to study classes that closely resemble one another on all variables save the use of television, i.e. preparation and experience of classroom teachers, background of students, community support of the school, availability of ancillary learning resources, etc. The random assignment of students to television and non-television classes is perhaps the most effective way to insure comparability between instructional conditions, but random assignment is often an impossible strategy to implement for administrative reasons.

The third and strictest definition of success is predicated on the existence of well-defined objectives and performance criteria. Under such circumstances, effectiveness may be defined as a target population's ability to obtain a certain level of knowledge, reach a prescribed level of consensus on some subject, or exhibit certain behaviours with sufficient frequency to fulfil project goals. While such an approach has proven useful in the construction of feedback instruments as well as other short-term evaluative tools, project administrators have been reluctant to be held accountable for longer-term social or even educational objectives, which are often unrealistic or vague and subject to considerable modification in the course of a project.

# Efficiency

Even with well-defined objectives, the evaluation of a project may still not provide planners with the kinds of information they need to determine if results were obtained in the most efficient way. In fact, most summative evaluations of media systems to date have been of limited practical value to planners because they have examined the effectiveness of a single communication strategy and have ignored possible alternative strategies for accomplishing the same ends. When only one approach is analysed, it is hard to be sure that it is the most effective one to pursue in the future. Furthermore, even when alternative approaches are compared on some criteria and one proves to be superior, it may still be impossible to decide which approach to adopt permanently without re-analysing costs. If a new system proves superior to some traditional alternative but costs substantially more, it may be necessary for the planner to consider how

the traditional alternative might perform were the same amount of money to be applied to improving it. While it may never be possible to examine thoroughly every possible use of project funds, it is likely that future evaluations will be required to do more in the way of comparative costeffectiveness analysis. Such comparisons can never be complete, and many will continue to be flawed by methodological problems. On this point, economists Carnoy and Levin have sharply challenged the validity of the cost-effectiveness analyses conducted on the Mexican Telesecundaria, the British Open University and El Salvador's educational television system (Carnoy and Levin, 1975). Such studies, they claim, contain too much 'benefit of the doubt' thinking in which overly optimistic, drawing-board estimates of technology-based systems are compared with the costs actually incurred by systems operating in traditional modes. Such practices, if widespread, tend to accentuate the cost-effectiveness of media-based systems, but only time will tell how accurate the enrolment and cost projections of the various systems were to begin with. The lesson here seems to be that planners should proceed cautiously in forecasting the cost-efficiency of new communication systems.

With or without realistic and reliable cost estimates, summative evaluations of communication systems have been of only limited utility to communication planners. Although most of the studies have provided some insights into what a given system or project achieved, in terms of either its original objectives or objectives attributed to the system by its evaluators, few evaluations have offered insights into how the system actually accomplished what it did. Planners require more than summaries of learning, attitude and behavioural changes associated with a given system if they are to learn from experience. They also need to understand how and why a particular system or project arrived at its results. This need is reflected in the attention now being paid to process evaluation and related efforts to understand in detail the variety of environmental, political and administrative conditions under which communication systems operate. The growing interest in process evaluation also reflects the view that success may be based more on factors related to system implementation, control and adaptability than on the theories and assumptions underlying system design.

The multiplicity of conditions and activities embodied in any communication system caution the evaluator against drawing finely tuned inferences regarding the precise causes of system success or failure. Success in most communication systems is related to community receptivity and support, field staff capability and motivation, and prior experience, as well as a host of variables associated with the appeal, comprehensibility and motivating potential of the messages produced by the system. To 'explain' how a system achieved its results, it is necessary to account for these and other factors. This is extremely hard to do. Consider for example the case of a rural radio system such as the one operated by Colombia's Acción Cultural Popular (ACPO) for the past thirty years. In this case, it is extremely hard to judge whether it is the cumulative impact of the radio school movement or the gradual improvement of economic conditions and opportunities in the Colombian countryside that has led to greater farmer productivity and well-being. Chances are that both forces played a part, but to what degree and in what relationship to one another? Unless conclusions can be reached on such issues, and unless subtle distinctions can be drawn between a particular communication system's influence versus the influence of other social programmes as well as various socioeconomic and environmental factors, little specific information will be forthcoming of great import to planners of new communication systems.

Controlled field experiments have been undertaken in some situations to resolve the kind of issues cited above. This technique involves a comparison of groups (individuals, classrooms, communities), some of which have participated in a particular communication system, some of which have not. The goal of this research technique is to maximize the comparability of groups on all factors save those directly related to the communication system itself. The sampling and assignment of recipients to different conditions (some with, others without radio, for example), is extremely difficult and subject to serious error and bias. Furthermore, such designs are difficult to maintain over time and to justify in that some potential beneficiaries may have to be denied services under the new system. This may be impossible to do on either political or ethical grounds. However, as Hornik (1976) has pointed out, historical circumstances and nature may, from time to time, provide researchers with unique opportunities for controlled field experimentation. Such opportunities spring up because of the anomalies of communication media themselves (one side of a mountain may receive a radio signal while the other side does not) or because the rate of a system's growth may permit communities yet unserved to act as controls for those for whom service has recently been established. Early identification of such possibilities can enhance the possibilities for genuine control groups and, ultimately, broaden the range of conclusions flowing from the evaluation.

In the final analysis, time has been the greatest handicap under which summative evaluators have laboured and perhaps the main reason why the results of their studies have not been sufficiently incorporated into the communication planning process. The anticipated social benefits of any new communication system customarily lie far in the future, forcing evaluators to judge success on a range of what Carol Weiss has termed

'surrogate measures'. Such measures may reveal significant short-range improvement or change, they may even highlight promising trends, but only rarely do they yield results that are unambiguous and therefore clearly instructive to the planners of new communication systems. Traditional evaluations have also been remiss in that they have not treated unanticipated consequences of communication systems or dealt adequately with the issue of how system benefits are distributed across a recipient population. For example, a number of projects such as Sesame Street that have been heralded as a means for uplifting previously underpriviledged or oppressed segments of the population have ended up yielding proportionately greater benefits to more priviledged segments of society (Cook et al., 1975). This phenomenon underscores the importance of (a) replicating measurements within as well as between projects, (b) conducting longer-range observational studies of project activities, and (c) paying more attention to the administrative behaviour of communication systems. In these ways, it may be possible to draw better generalizations for planning from different communication systems and to gain deeper understanding of the processes over which planners can actually exert some control.

#### Summary and conclusions

Table 1 summarizes the relationships between communication planning and evaluation discussed in this essay. As the chart illustrates, each of the stages is associated with a variety of research methods and data sources, although within most systems there is likely to be a considerable duplication of methods and sources across stages. Such duplication suggests that basic evaluation strategies and data sources do not vary markedly during the planning and implementation of a communication system. What does vary, however, is the nature of the decisions needed to guide the system and, therefore, the means of aggregating data to bear on those decisions. During the policy definitions stage, for example, communication planners are customarily preoccupied with normative questions concerning what new services and investments ought to be undertaken. During the system design stage, strategic issues concerning what can be done become paramount, while the implementation and maintenance stage is characterized by the need to determine how policies can be put into practice and then institutionalized most efficiently. In the final stage, assessment of system impact, answers are sought to questions concerning how well a system actually performed what it was designed to do. Ideally, the stages feed into and reinforce one another, although experience suggests that planning and evaluation are often disjointed activities, involving different actors and divergent expectations.

Stages	Methods	Data sources	Examples
Policy definition	Sectoral reviews and needs assessment Analysis/interpretation of in-country experience Analysis/interpretation of experiences elsewhere	Census figures, school records, health clinic utilization rates, etc. Reports of various government agencies Interviews with appropriate officials Performance information in the form of annual reports, administrative memoranda, formal and informal accounts of project activities Published evaluations, including case-studies	Unesco-directed pre-investment communi- cation needs assessment studies in Thailand, Indonesia and Afganistan USAID/IIEP-sponsored case studies of 23 communication projects ( <i>The New Media:</i> <i>Memo to Educational Planners</i> , by Wilbur Schramm et al., 1967) Published evaluations of communication projects in Samoa, Niger, El Salvador, Ivorv Coast, Mexico, U.S. ( <i>Sesame Street</i> )
System design	Analysis of potential audience's needs, preferences and habits Diagnosis of existing communication infrastructures and resources, hardware and software	Survey of potential audience(s)—approximate numbers to be reached, location, language patterns, media use, interpersonal communication networks Content analysis of existing programmes Inventories of existing production, transmission and reception equipment Current enrollments and projections of teacher training and other high-level institutions	Planning studies for instructional television in Samoa, Ivory Coast, El Salvador and Korea; for rural television in India (SITE), and for rural radio in Canada, Tanzania and Nepal, among others
System implemen- tation and maintenance	Pretesting of materials and methods at production headquarters Simulation of listening/viewing conditions Pilot testing of system components in the field Feedback (e.g. continuous monitoring of the system via tests, questionnaires and other instruments designed to vield quick results)	Reactions (cognitive, affective and behavioral) of sample audiences—students, teachers, farmers, mothers' club members, etc. Cost and effectiveness information pertaining to the production, distribution and use of all system components	Sesame Street and The Electric Company, SITE (India) Basic Village Education Project (Guatemala), Radio Mathematics Project (Nicaragua)
Assessment of system impact	Panel studies (e.g. repeated measures over time of the same subjects to determine extent of change) Comparative cost-effectiveness analyses Field experiments Case studies	Information related to knowledge, attitudes and behavior of all groups involved in the system All of the above, but coupled to cost estimates of existing procedures and those of traditional and/or viable alternative systems Sites where essential system elements can be varied through administrative manipulation	ITV evaluations in El Salvador and Ivory Coast Comparative cost-effectiveness of Mexican Telesecundaria Radio vs. multi-media approach in Iran Radio campaigns in Tanzania, Cuba and Honduras: ITV in El Salvador, instructional radio in San Luis Potosí, Mexico

# TABLE 1. Evaluation and communication planning: a summary
Unfortunately, the issue of how communication planning and evaluation can be better integrated has not been adequately addressed within most systems and what has resulted is haphazard and highly improvisational communication plans on the one hand and rather dull, irrelevant evaluations on the other. Programme administrators and producers continue to view evaluation, especially summative evaluation, as a potential threat. In such a climate, evaluators have tended to isolate themselves, deciding on their own what aspects of a system are worthy of study. Upon completion of their investigations, they have reported their conclusions in lengthy, jargon-filled reports which neither planners nor administrators have had the time or inclination to read. By sidestepping the critical issue of who will use their research and for what reasons, evaluators forfeit the opportunity both for making useful contributions to the development of the systems they are studying and for influencing in any significant way the development of new communication systems elsewhere.

## WHO WILL EVALUATE THE EVALUATORS?

What can be done to improve the relationship between communication planning and evaluation? As a first step, planners must become better informed and more demanding critics of the evaluation activities conducted throughout their systems. In so doing, they must recognize that evaluation is potentially as much a tool for guiding allocation decisions as it is for legitimizing policies enacted in the past. Secondly, they must distinguish between research designed to help a system work better, and research aimed at discovering how well a particular system worked. As argued above, such concerns are not incompatible, but they do require different schedules and, frequently, different personnel. In the first instance, helping a system to work better, the need is for reliable information on how a particular system is performing on a week-to-week, even day-to-day, basis. Such research necessitates data collection and analysis strategies keyed to the immediate concerns of programme producers and administrators. To gain the attention of both groups, the identification of the evaluation unit with the system and allegiance to its goals are probably required. In this way, the evaluator's motives will not become an issue and the probabilities for immediate incorporation of research results into the planning process will be enhanced.

The utility and control of longer-range evaluations of system effects pose more troubling issues for the planner. Overly close supervision may bias the evaluation. As James Q. Wilson (1973) warns in his first law of social science evaluations: 'All policy interventions in social problems produce the intended effect if the research is carried out by those implementing the policy or their friends' (1973, quoted by Carnoy and Levin, 1975, p. 387). On the other hand, a *laissez-faire* attitude towards externally-conducted research may run the risk of bias in the other direction and also reduce the value of the results to the system itself. Here a compromise would seem in order, perhaps involving decisionmakers and programme administrators in the planning of the longer-range evaluations but leaving the execution of such studies and the interpretation of their results solely in the hands of the evaluation team, be it internally or externally directed.

Finally, what can evaluators do to improve communication planning? First, they must become more cognizant of the political milieu in which virtually all communication planning occurs. Planning inevitably involves vested interests and, therefore, perpetual negotiation and compromise. To claim that communication planning should be based only on objective criteria is to blind oneself to the personal ambitions and institutional pressures that fuel the entire process. By clarifying policy alternatives, evaluators can unquestionably perform a much needed service to planners, but such alternatives will inevitably be framed by the latters' existing commitments and loyalties. By the same token, useful evaluations, at whatever level, can rarely avoid judgements about personal as well as institutional relationships. Indeed, such relationships may be far more relevant to the success or failure of a particular system than anything else. However, research in this area may leave individuals extremely vulnerable. This is especially so if the behaviour or attitude to be reported is in any way threatening or damaging to an administrator's authority or self-respect. Here the evaluator must tread carefully, balancing the need to explain system performance fairly and accurately with the desire to improve decision-making within that system.

A second consideration that must remain uppermost in the evaluator's mind is the vital, organic nature of communication systems. Rarely will it be possible to base judgements of system impact or effectiveness on a single criterion. Rather, communication systems are more appropriately viewed as constantly changing phenomena that influence and respond to their environment in multiple, and often contradictory, ways. This quality calls for multidimensional evaluation designs, longitudinal analyses of effects, and enough time to consider side effects as well as promising policy alternatives to guide future planning efforts.

Recent experience from a variety of communication projects suggests that evaluation can exert a positive influence on communication planning when and if decision-makers and researchers genuinely agree to collaborate. This requires evaluators who are decision-oriented and willing to sacrifice some methodological rigour for the opportunity of providing useful guidance to planners, as well as planners who not only value information from the field but who are also willing to insist that such information be provided before major policy decisions are taken.

## References

- CARNOY, M.; LEVIN, H. M. 1975. Evaluation of Education Media: Some Issues. Instructional Science, Vol. 4, p. 385-406.
- COOK, T. D.; APPLETON, H.; CONNOR, R. F.; SHAFEL, A.; TAMKIN, G.; WEBER, S. J. 1975. Sesame Street Revised. New York, Russell Sage Foundation.
- FRIEND, J. 1976. Closing The Feedback Loop. Paper presented at the Conference on the Use of Radio as an Instructional Tool in Primary Schools, Managua, Nicaragua, 6-9 September.
- HALL, B. L. 1975. Participatory Research: An Approach for Change. Convergence, Vol. VIII, No. 2, p. 24-32.
- HORNIK, R. C. 1976. Useful Evaluation Designs for Evaluating the Impact of Distance Learning Systems: Methodology. Educational Broadcasting International, March.
- MCANANY, E. G.; HORNIK, R. C.; MAYO, J. K. 1973. Studying Instructional Television: What Should be Evaluated. Stanford, Calif., Institute for Communication Research.
- MERINO DE MANZANO, A. M.; HORNIK, R. C.; MAYO, J. K. 1971. Feedback on Student Learning for Instructional Television in El Salvador. Stanford, Calif., Institute for Communication Research.
- MONTGOMERY, J. D. 1974. Technology and Civic Life: Making and Implementing Development Decisions. Cambridge, Mass., The MIT Press.
- SCHRAMM, W. 1969. Feedback for Instructional Television. Stanford, Calif., Institute for Communication Research.
- -. (ed.). 1972. Quality in Instructional Television. Honolulu, University of Hawaii Press.
- SELLTIZ, C.; WRIGHTSMAN, L. S.; COOK, S. W. 1976. Research Methods in Social Relations. New York, Holt, Rinehart & Winston.
- STANFORD UNIVERSITY CONSORTIUM. 1976. Review Essay: Evaluating the Handbook of Evaluation Research. G. V. Glass (ed.), Evaluation Studies Review Annual. Beverly Hills, Calif., Sage Publications.
- SWANTZ, M. L. 1975. Research as an Educational Tool for Development. Convergence, Vol. VIII, No. 2. WEBB, E. J.; CAMPBELL, D. T.; SCHWARTZ, R. D.; SECHREST, L. 1966. Unobtrusive
- Measures: Nonreactive Research in the Social Sciences. Chicago, Rand McNally.
- WEISS, C. H. (ed.). 1972a. Evaluating Action Programs: Readings in Social Action and Education. Boston, Allyn and Bacon.
- . 1972b. Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs, N.J., Prentice-Hall.
- WILSON, J. Q. 1973. On Pettigrew and Armor: An Afterword. The Public Interest, Vol. 30.

## Notes on contributors

John Middleton is Assistant Director for Planning and Evaluation and Research Associate, East-West Communication Institute, Honolulu, Hawaii. He received his doctorate in educational planning from Harvard University. His research and professional activities have focused on the use of communication for development, with an emphasis on planning. Recent publications include *Perspectives on Communication Policy and Planning* (with S. A. Rahim); *Educational Planning: Lessons for Communication Planners* (with Barclay Hudson); and *The Communication Planning Workshop: A Simulation Game* (with Gus Root, Meow-Khim Lim, George Beal and Meheroo Jussawalla).

James Lyons is Director of Computing Facilities and Associate Professor and Director of Academic Services at INSEAD (the European Institute of Business Administration), Fontainebleau, France. He received his doctorate in physics from the State University of New York, and worked for six years in the area of theoretical atomic physics; he has also worked on land-use allocation models, occupied a variety of managerial positions, and taught in both the United States and Europe. His published work is mostly in the area of atomic physics, related to research with a strong computational emphasis.

Gus Root was until recently Associate Professor of Education at Syracuse University, United States, with a particular interest in instructional technology and educational systems research. Originally trained at MIT as an engineer, after considerable industrial experience he received, in 1967, a doctorate from Syracuse in the field of psychology, and has concentrated since then on theories of learning, systems analysis and modelling, and research into decision-making. He has also worked in Iran, and in 1977 organized an Asian communication planning workshop in Malaysia.

Bernard Webster bridges the two worlds of education and technology; after a period in industry spent mainly in the communication research and development laboratories of the Marconi Company he moved into higher education. He is Head of the School of Electrical Engineering, Plymouth Polytechnic, United Kingdom, and Director of the Devon Educational Television Service. The author of numerous articles in professional and technical journals, he is a fellow of the Institution of Electrical Engineers, and the Institution of Electronic and Radio Engineers, and a Member of the Royal Television Society, the International Institute of Communications, and a founder member of the Educational Television Association. Upon its formation in 1967, he was also appointed to membership of the National Council for Educational Technology of the United Kingdom.

John Spence is a Director of W. D. Scott and Company Pty Ltd, International Management Consultants. In more than twenty years with the company, he has been particularly associated with the development of new approaches in long-range planning, information systems and resource evaluation. He has directed many Australian and international large-scale projects associated with planning the expansion and operation of communications systems, for such clients as the Australian Postal and Telecommunications Commissions, the Australian Department of Supply, the International Telecommunications Satellite Organization and the Asian Development Bank. He is on the Australian National Commission for Unesco and has also served on government advisory boards in resource planning and in setting up advanced education programmes for management sciences.

Steven Klees was an Assistant Professor in the Economics of Education at Cornell University following the completion of his doctorate at Stanford University. He has recently served on a visiting appointment at Stanford and is currently teaching at the University in Natal, Brazil. His research activities have led to many publications, including *The Costs of Educational Media* (with Dean Jamison and Stuart Wells).

Stuart Wells received his doctorate in economics and business from Stanford University. His research has specialized in the economics of education with an emphasis on communication and education systems in developing countries. He has authored numerous articles and several books, including *Instructional Technology in Developing Countries* and *Costs of Educational Media* (with Dean Jamison and Steven Klees). He is currently an Associate Professor of management at San Jose State University.

John Mayo has been since 1977 Associate Director of the Learning Systems Institute and Associate Professor of Communication at Florida State University. He received his doctorate from Stanford University, where from 1973 to 1977 he was Assistant Professor in the Communication Department and a Research Associate of the Institute for Communication Research. His publications are mainly in the field of evaluation, including co-authorship of *Educational Reform* with Television: The El Salvador Experience (with R. Hornik and E. McAnany). He has worked at various times in Colombia, Mexico, Chile and Nepal.

[A.58] CC.79/XXXIII.1/A