



THE UNIVERSITY OF
SYDNEY

COPYRIGHT AND USE OF THIS THESIS

This thesis must be used in accordance with the provisions of the Copyright Act 1968.

Reproduction of material protected by copyright may be an infringement of copyright and copyright owners may be entitled to take legal action against persons who infringe their copyright.

Section 51 (2) of the Copyright Act permits an authorized officer of a university library or archives to provide a copy (by communication or otherwise) of an unpublished thesis kept in the library or archives, to a person who satisfies the authorized officer that he or she requires the reproduction for the purposes of research or study.

The Copyright Act grants the creator of a work a number of moral rights, specifically the right of attribution, the right against false attribution and the right of integrity.

You may infringe the author's moral rights if you:

- fail to acknowledge the author of this thesis if you quote sections from the work
- attribute this thesis to another author
- subject this thesis to derogatory treatment which may prejudice the author's reputation

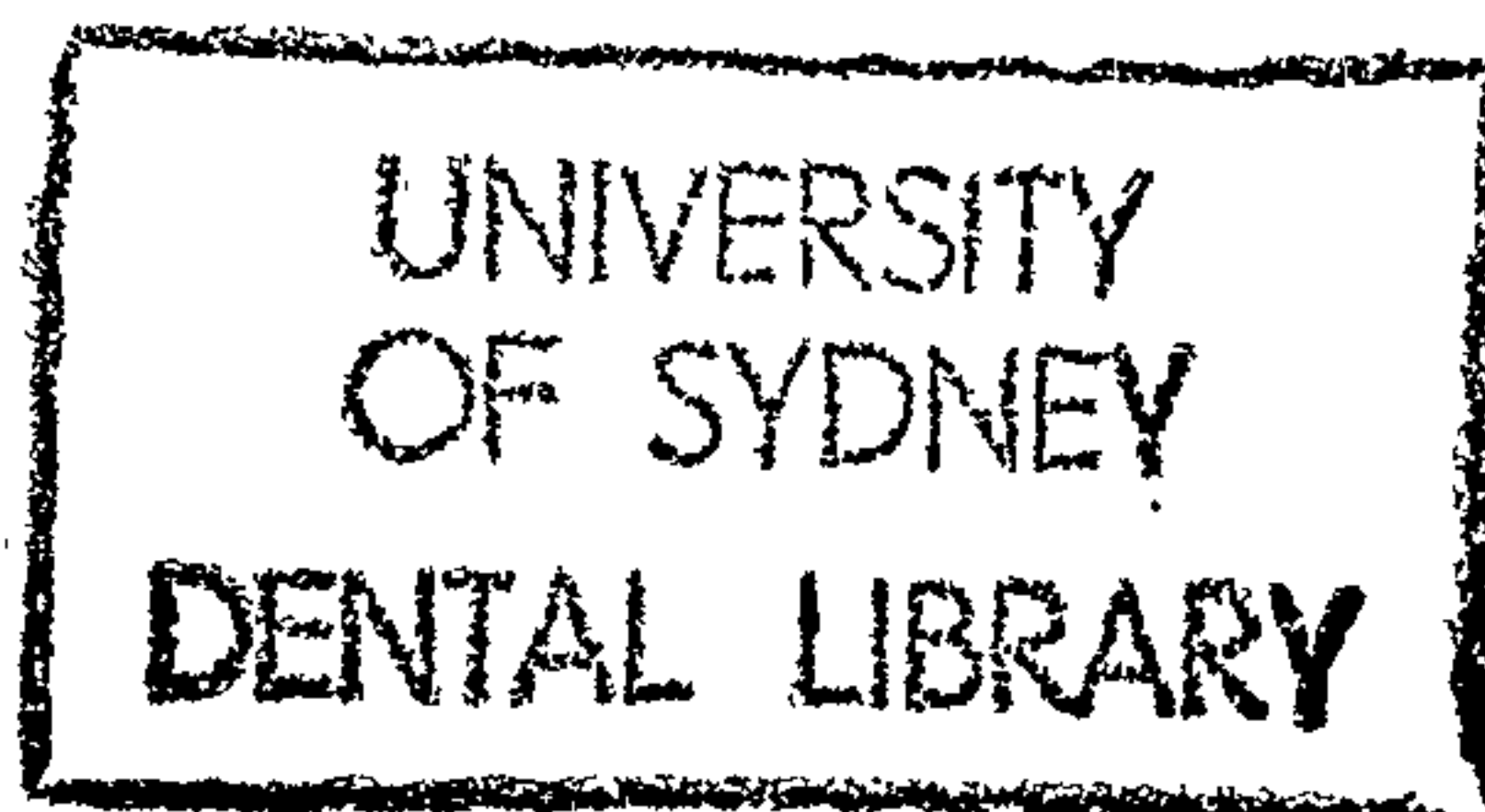
For further information contact the University's Copyright Service.

sydney.edu.au/copyright

AN EXAMINATION OF CONCEPTS OF DEFINITION AND
PURPOSE, AND OF CERTAIN ASPECTS OF METHODOLOGY
FOR THE EVALUATION OF HEALTH CARE SYSTEMS

A.J. Stockwell, B.D.Sc.(W.A.)

A thesis submitted in partial
requirement for the
DIPLOMA IN PUBLIC HEALTH DENTISTRY



Department of Preventive Dentistry
Faculty of Dentistry
The University of Sydney

1979

The writer gratefully acknowledges
the assistance of others during the
research and preparation of this
work:

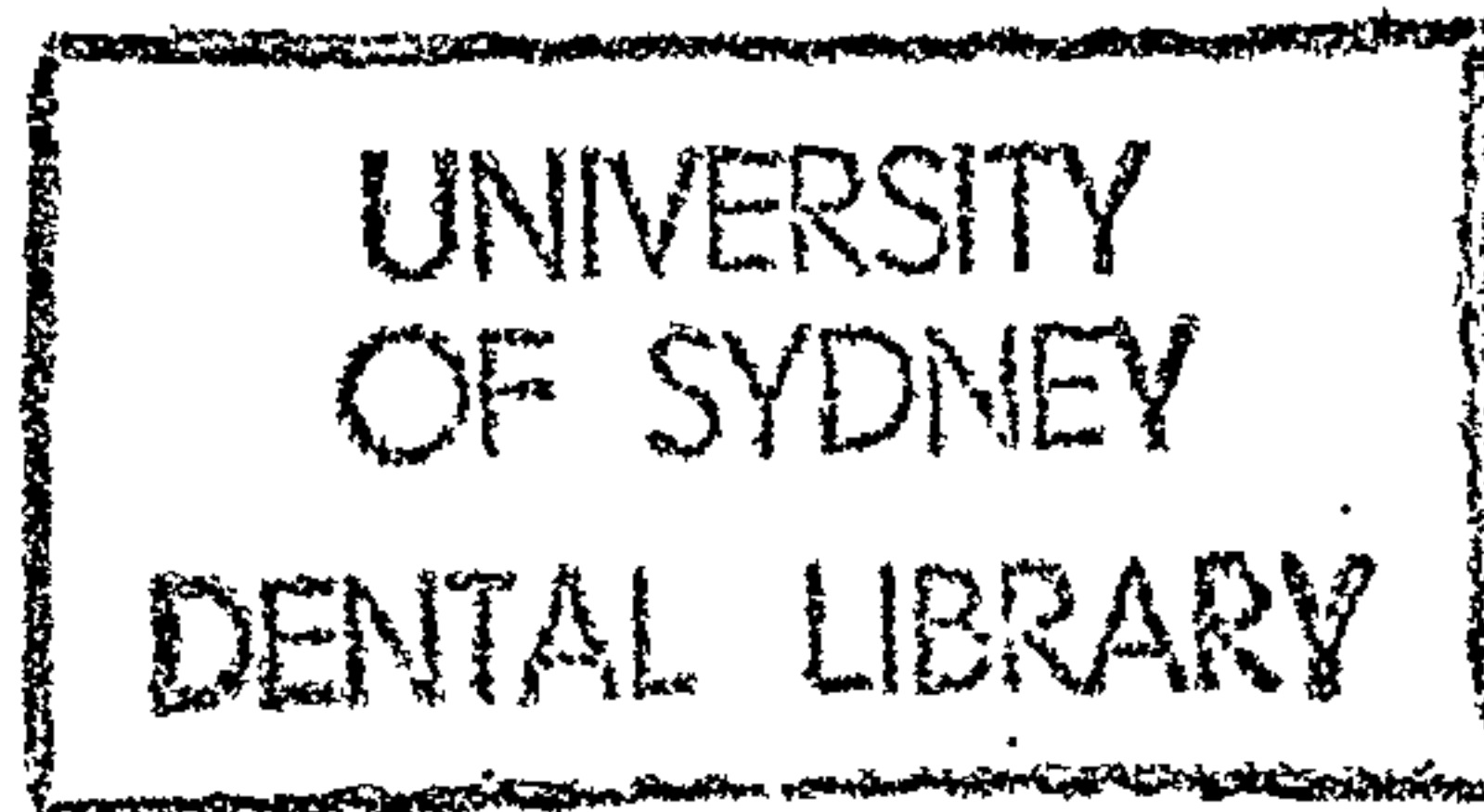
Associate Professor P.D. Barnard
Department of Preventive Dentistry
Faculty of Dentistry
The University of Sydney

Mr. S. Duckett and Mr. G. Rawson
School of Health Administration
The University of New South Wales

Dr. R.F. Southby
School of Public Health and
Tropical Medicine
Commonwealth Department of Health
and The University of Sydney

CONTENTS

		Page
Acknowledgements		ii
List of figures		vi
List of abbreviations		vii
1. INTRODUCTION		1
2. CONCEPTS OF DEFINITION OF EVALUATION		6
2.1 Conclusions		13
3. CONCEPTS OF PURPOSE OF THE EVALUATION OF HEALTH CARE SYSTEMS		14
3.1 Conclusions		21
4. A CLASSIFICATION OF EVALUABLE OBJECTS IN HEALTH CARE SYSTEMS		23
4.1 Introduction		23
4.2 Concepts of classification		26
4.3 Structure dimension		31
4.4 Process dimension		32
4.5 Outcome dimension		33
4.6 Systems/resource utilization dimension		34
4.7 Conclusions		40
5. A CLASSIFICATION OF OPERATIONAL FRAMEWORKS OF EVALUATION		42
5.1 Introduction		42
5.2 Goal-attainment framework		46
5.3 Systems/resource analysis framework		61
5.3.1 Systems analysis		62
5.3.2 Economic analysis		66
5.4 Conclusions		67



6.	EVALUATIVE CONSTRUCTS, CRITERIA, MEASURES AND STANDARDS	69
6.1	Introduction	69
6.2	Evaluative constructs	70
6.2.1	Constructs of goal attainment	70
6.2.2	Constructs of resource utilisation	72
6.2.3	Constructs of suitability	73
6.2.4	Constructs of social responsiveness	75
6.2.5	Constructs of flexibility	76
6.2.6	Constructs of performance	77
6.2.7	Miscellaneous constructs	78
6.3	Evaluative criteria	78
6.4	Evaluative measures	81
6.5	Evaluative standards	82
6.6	Conclusions	82
7.	COMPARATIVE TECHNIQUES IN EVALUATION	84
7.1	Introduction	84
7.2	Goal-attainment techniques	84
7.2.1	Medical/dental audit	84
7.2.2	Goal attainment scaling	89
7.2.3	Quality control	90
7.3	Systems analysis	91
7.3.1	Cost-benefit analysis	91
7.3.2	Cost-effectiveness analysis	93
7.3.3	Program planning and budgeting system	97

7.4	Economic analysis techniques	98
7.4.1	Introduction	98
7.4.2	Cost-benefit analysis	98
7.4.3	Cost-effectiveness analysis	102
7.5	Conclusions	104
8.	SUMMARY	106
9.	CONCLUSIONS	118
10.	BIBLIOGRAPHY	121

List of Figures

	Page
1. Degree to which goals are achieved.	18
2. Planning - action - evaluation cycle.	19
3. Evaluation - A schematic overview.	28
4. A model of a dental health system.	39
5. A management information system.	46
6. A model of the evaluation process.	50
7. Decision making and the evaluation process.	54
8. A medical care evaluation model.	57
9. Outline of the algorithmic approach to outcome-based health planning evaluation.	59

List of Abbreviations

CBA	cost-benefit analysis
CEA	cost-effectiveness analysis
GST	general system theory
HSA	health systems agency
MIS	management information system
OVA	output value analysis
PCU	primary care unit
PPBS	program planning and budgeting system
UK	United Kingdom
US	United States (of America)
WHO	World Health Organisation

1. INTRODUCTION

The writer has produced this work in an attempt, firstly, to examine, and hopefully clarify, the conceptualisation of the meaning and purpose of evaluation, and secondly, to examine and clarify certain aspects of evaluation methodology for health care systems. The writer has found no published literature which meets these aims.

"The subject of evaluation is very fashionable at present",²¹³ and "... the many ways in which it may be and has been approached are large, complex and dynamic."⁷¹ The resulting literature on health systems evaluation can justly be described as chaotic, with a remarkable lack of consistency in terminological and methodological conceptualisation. This inconsistency runs through almost every aspect of evaluation, from its broadest evaluable 'dimensions', to narrow evaluative 'constructs'. Weisbrod found that "a striking feature of current health programme evaluation theory is its profusion of conceptual formulations from many disciplines and sources with which is coupled an almost complete anarchy in terms of terminology."¹⁷⁹ This confused state is said to be due to three causes: firstly, the instruments of evaluation are still rudimentary; secondly, the facts to be observed are often abstract and multidimensional, and it is not easy to operationalise them; and thirdly, health itself is

not satisfactorily defined, and the variables that affect it are not all controllable during the period of evaluation.¹¹⁶ The writer has found, like Glass, that "the methodological and philosophical literature on evaluation is distressingly repetitious. Writers in one field exhume the same dead horses that writers in a second field interred years before. One who reads widely is in jeopardy of gagging from repeated swallowing of dry banalities about how evaluation can be distinguished as either 'process' or 'outcome'..."⁶⁴

The writer has approached his task in the following manner. Firstly, the conceptualisation of the meaning of evaluation itself has been investigated to ascertain the degree of agreement or discord in such ideas. Secondly, the rationale for evaluation has been examined, because, as Thompson observed, "an understanding of the motives of evaluation is essential to the understanding of its processes."¹⁶⁸ The examination of these two items has proved to be straightforward, in spite of the diversity of opinions reviewed.

The remainder of this work, which examines and classifies certain aspects of the methodologies of the evaluation process, does not so readily lend itself to straightforward examination, since the compartmentalisation of concepts is diffuse and often overlapping. Thus, for clarification, the writer has attempted to create his own discrete classifications of evaluation. One of the major problems has been the lack of published material directly related to the development of such a concept. In essence, the writer's schema pro-

poses two major facets of evaluation: (1) the classification of the objects to be evaluated into four 'evaluable dimensions', and (2) the organisation of evaluation methods into two 'operational frameworks'.

The four evaluable dimensions of health care systems are those of: structure, process, outcome, and systems/resource utilisation. The first three of these stem from the classical work of Donabedian in the evaluation of the quality of medical care⁴⁴, while the fourth is derived from the disciplines of systems analysis, economics and accounting.

In the selection and description of operational frameworks, the writer was guided by the fact that although published evaluation studies have taken place within a wide variety of fields (such as health, education and social welfare), in a broad selection of settings, have focussed on many phenomena, and have employed diverse methodologies, nonetheless, overall, these methodologies have been said to fall into two basic frameworks:^{131,179} (1) the goal-attainment framework, which measures the degree of success or failure of a program as compared with its predetermined goals, and (2) the systems (or control) framework, which examines the distribution and utilisation of resources within a system as compared with the optimal distribution.¹⁰² The writer considers that two other disciplines are related to this latter framework, these are the economics discipline and the accounting discipline. They have been included within a systems/

resource analysis framework.

The inter-relationships between the evaluable dimensions and the operational frameworks appear to be as follows. The resource utilisation dimension may be evaluated by means of the systems, economics or accounting disciplines. The outcome dimension may be evaluated through the goal-attainment framework. The process dimension has been evaluated using the goal-attainment framework on the basis that the implicit or explicit goal is a specified standard of performance of the process being evaluated. The writer has not pursued the evaluation of structure, which, it is felt lies outside the scope of this work.

The systems framework has been examined in some detail because its advocates propound that it alone is the correct method of evaluation. The economics approach has been followed in so far as it relates to the use of cost-benefit analysis and cost-effectiveness analysis, which techniques have also been strongly advocated within the systems framework. The accounting framework has only been briefly mentioned.

The goal-attainment approach has also been pursued in some depth because the writer considers that it is the most useful single methodology for evaluating health care systems.

The remainder of this work is devoted to an examination of a part of an operational sequence of eval-

uation, and envisages the following stages. Within each operational framework, an evaluation 'technique' may be applied. Each technique involves the selection of one or more evaluative 'constructs', to which one or more evaluative 'criteria' may refer. Then, either (1) in the case of the goal-attainment framework, evaluative 'measures' are selected for, and assigned to, these criteria, and finally, an evaluative 'standard' is used as a reference for the selected measure to ascertain the extent that actual performance relates to desired performance, or (2) in the case of the systems/resource analysis framework, a ratio or index is computed, from which judgements about potential choices may be made.

The terminology used by the writer conforms with that of the literature wherever a relatively consistent use of a word has been found. Where this has not occurred, an appropriate word has been selected.

The writer has attempted to summarise his findings, and to draw pertinent conclusions about this very confused subject, in the final sections of this work.

2. CONCEPTS OF DEFINITION OF EVALUATION

The Oxford English Dictionary¹⁰⁴ defines the word evaluate as: "a. Math. To work out the value of (a quantitative expression); to find a numerical expression for (any quantitative fact or relation). b. Gen. To 'reckon up', ascertain the amount of; to express in terms of something already known". It is therefore evident that four distinct meanings may be assigned to the word by this definition. Webster's New International Dictionary¹⁷⁷ defines evaluate as: "1a: to set down or express the mathematical value of: express numerically, b: to estimate or ascertain the monetary worth of. 2. to examine and judge concerning the worth, quality, significance, amount, degree or condition of". This definition extends the meaning of the word by adding two other dimensions involving economic and subjective qualitative judgements.

Thompson, in a critical review of evaluation projects within United States government departments, found that the original definitional equivalence to appraisal or quantification had been expanded to include concepts outside its strict denotations, and at the same time had become rigid and ritualistic. "Just as writers add to, subtract from, or otherwise embellish the concept of evaluation, they tend dogmatically to exclude views alternative to their own and to lay down inflexible formulae for their performance. Were evaluation permanently to be defined according to these unbending dicta, it might be constrained from ever contributing valuably to government."¹⁶³

Thompson concluded from his analysis that ambiguity about the meaning of evaluation is ineradicable, and that one of the main consequences of this is an imprecise methodology.¹⁶⁴

Suchman, in a review of the evaluative literature, whose purpose was to formulate a 'theory and method' of evaluation, lamented the increasing imprecision of evaluation, finding it "poorly defined and often improperly used."¹⁵⁴ He outlined the overall problem in these terms: "Different results obtained for different purposes by different methods and based on different criteria lead to a confusion which is doubly difficult to resolve in the frequent absence of any explicit statement of objectives or methods of procedure by the evaluator."¹⁵⁵ However, his own definition of the evaluation study was particularly diffuse: "In general, whenever one asks such questions as - 'How good is the program?' 'What effects are we having?' 'Is the program working as we expected?' - and uses such instruments as sheets, appraisal forms, evaluation guides or research designs which involve comparing accomplishments before or after or in the presence or absence of a particular action, one may be said to be conducting an evaluation study."¹⁵⁴ A similarly-diffuse treatment was given by Roemer in a paper published by WHO in which he defined evaluation as "determination of the value of a course of action."¹¹⁷ Another broad defin-

ition was that of Knutson who thought that "one might define evaluation as the process of determining the value or worth of something relative to a given purpose or standard. It is the process of making decisions, drawing together evidence, weighing the pros and cons of various suggestions, and selecting courses of action."¹⁶⁵

In contrast, Hanlon considered it "important to distinguish between evaluation and the mere development of indices of health or of health practice and the surveys necessary to obtain them." While praising the merit of such indices and surveys, he nevertheless concluded that they should not be confused with 'true evaluation'.⁷² Hutchinson explicitly distinguished between evaluations and 'demonstrations and program reviews', the latter being "similar in general purpose to the evaluation but less rigorous in design."⁸⁰

Zusman and Wurster noted that evaluation, like many of the key words of the health field, is a widely used term, which is apparently universally understood but rarely, if ever, defined precisely. They favored the definition by Gruenberg of 'evaluation research' (in the community health field): "analysis of program effectiveness, the use of scientific method in appraising the success of community health programs."²¹⁴ The ambiguity in the evaluation literature was specifically mentioned by Fleck, who commented that:

"Evaluation is a verbal symbol in vogue in public health administration circles. Its popularity is a result of its ambiguity. Like most sacred symbols, the term can be widely shared."⁵⁶ While to offset the imprecision which he felt resulted from so great a breadth of meaning for the term 'evaluation', Thompson created another broad definition: "the marshalling of information for the purpose of improving decisions."¹⁶⁷

More extensive claims for evaluation have been made by Anderson, who claimed that it "is also concerned with determining whether the goals themselves are valid"⁷, and by Bigman, who proposed as a main evaluative role: "To redefine the means to be used for attaining objectives and even to redefine subgoals, in the light of research findings."²²

It appears that in the economics discipline (and often in the systems framework) the term evaluation has a distinctly different meaning from that in other frameworks. It was defined by Drummond as "the appraisal of alternative courses of action. This appraisal may be in the form of two related questions: (i) what should the objectives be? (ii) are the chosen objectives being pursued in the 'best' possible way?"⁴⁷

In the health field, Hanlon emphasised that "the meaning of evaluation is of great importance in relation to any efforts to achieve it"⁷², and an expert committee of WHO reported in 1975 that "there is a need to clarify the nature and function of evaluation in public health

programmes, to show how it can be helpful in planning, and how it can be used to improve ... programme operations. To achieve that a common definition of the term evaluation is needed."¹⁸⁵ The committee broadened a previous definition given by a WHO scientific group to that of: "A process for making judgements about selected objects and events by comparing them with specified value standards for the purpose of deciding among alternative courses of action."¹⁸⁵

Weckworth defined evaluation as "the comparison of accomplishments with stated objectives"¹⁷⁸, while Weiss thought it to be "the degree of achieving some specified set of objectives"¹⁸³, and the American Public Health Association defined it, in relation to health, as "the process of determining the value or amount of success in achieving a predetermined objective. It includes at least the following steps: formulation of the objective, identification of the criteria to be used in measuring success, determination and explanation of the degree of success, recommendations for further activity."⁶⁴ This view of the evaluation process is similar to that of James, who believed that evaluation "differs from research primarily in that it does not seek new knowledge, but attempts to mark progress toward a prestated objective. While research can end with the presentation of results, evaluation is viewed as part of a circular process. Its findings are reincorporated into the specific program from

which they came."⁷⁹ A somewhat similar approach, which distinguished between research and evaluation, was that of Tripodi, Fellin and Epstein, who defined social program evaluation as "... the systematic accumulation of facts for providing information about achievement of program requisites and goals relative to effort, effectiveness and efficiency within any stage of the program. The facts of evaluation may be obtained through a variety of relatively systematic techniques, and they are incorporated into some designated system of values for making decisions about social programs."¹⁷²

The Subcommittee on Evaluation of Mental Health Activities of the (US) National Advisory Mental Health Council stated that "evaluation thus connotes scientific method, but has characteristics that distinguish it from that type of research whose objective is the accumulation and analysis of data in order to formulate hypotheses and theory for the sake of new knowledge itself, irrespective of judgement of the value of the knowledge."¹⁷³

Very recently, the (Australian) Senate Standing Committee on Social Welfare adopted the criterion-based definition of social program evaluation originally proposed by Lyons: "Social program evaluation is the process of thoroughly and critically reviewing the efficiency, effectiveness and appropriateness of any program or group of programs."¹⁷⁰ Deming expressed a somewhat similar view; to him, "evaluation is a pro-

nouncement concerning the effectiveness of some treatment or plan that has been tried or put into effect."⁴⁰ He noted, however, that "the problems of evaluation ... are compounded by failure to define terms operationally, as well as by failure to lay down criteria by which to weigh gains and advantages against losses and disadvantages."⁴¹

Of recent years there has been a trend to define evaluation in more operational terms. An example is that of Attkisson and Broskowski:

- "1. A process of making reasonable judgements about program effort, effectiveness, efficiency and adequacy,
2. Based on systematic data collection and analysis,
3. Designed for use in program management, external accountability, and future planning,
4. Focuses especially on accessibility, acceptability, awareness, availability, comprehensiveness, integration, and cost of services."¹¹

The most recent definition that the witer has encountered is that of Rawson, who has attempted to remove the connotation of research entirely from the concept of evaluation. His extended definition reads: "Evaluation is a comparative process which monitors performance or quality and/or examines the effects of programs. Normally carried out according to a goal fulfillment model (or a derivative) it utilises existing data or that obtained by recognisable forms of

research methods, i.e., objective, systematic ways of gathering information, on appropriate aspects of program input, process and/or output-outcome."¹¹¹

2.1 Conclusions

1. Evaluation means many things to many people. It is perceived within a range varying from narrow, specific definitions, to broad, vague ones.
2. Several of the more precise descriptions and definitions fall conceptually within the ambit of the major part of the definition of the APHA, which views evaluation as the assessment of the degree of success in reaching a predetermined objective.
3. In contrast, the systems/economics approach sees evaluation as the appraisal of alternative courses of action.
4. None of the definitions, other than those in the cited dictionaries, equate 'evaluate' with 'express numerically'. Thus quantification may be perceived to be a part of the evaluation process, but is not perceived to be equated with it.
5. None of the cited dictionary definitions correspond to any of the other definitions offered, with the exception of that of Roemer.
6. Recently there has been a trend to define evaluation in more operational terms.

3. CONCEPTS OF PURPOSE OF THE EVALUATION OF HEALTH CARE SYSTEMS

A recent report by the (Australian) Senate Standing Committee on Social Welfare commenced with the sentence: "This report asserts that, in order to achieve an efficient, effective, **rational** and equitable health and welfare system, it is necessary to conduct on-going evaluation."¹⁶⁹ The provision of health services is becoming increasingly costly in all Western countries, in terms of money, manpower and other resources.¹⁹⁶ "All over the world more is being spent on health services and a higher proportion of the cost is being paid by taxation or social insurance. Increasingly access to health services has come to be regarded as a right: the modern state is seen to have a duty to ensure that all can use them, whether they can afford to purchase them or not.

"But how useful are health services in promoting health? On what principles should health services be organized and financed? How can the services be improved? How is it possible to ensure that a full range of services is available whenever they are needed? How is it possible to ensure that value for money is obtained for health services?"¹

The British Secretary for State emphasised the place of effective management and the importance of managerial

skill in his introduction to the consultative document outlining plans for the reorganization of the National Health Service.⁴³ A popular comment in America has been "hospitals need management more than money",⁹⁶ Australian surveys have concluded that management of the health services should be given close attention,¹¹³ and Sax has pointed out that efficiency and economy in the use of resources are valued by society.¹²⁴

A noted economist, Feldstein, drew attention to the fact that health care is an important area for the application of economic analysis. "The measurement and comparison of plant productivity and costs, widely practised in private industry, is of particular importance in public enterprise [which is] sheltered from the competitive forces that impel cost reduction and lack ... the incentive to efficiency provided by the profit motive ... Used judiciously, cost and productivity indices for individual public plants could provide a basis for better budgetary control, could identify particular plants that diverge from national averages, and could stimulate locally responsive managers to seek methods of improving their own efficiency."⁵⁴ Abel-Smith noted that the health market is virtually unique in the lack of mechanisms for self-regulation.²

It was noted in the Introduction that "An understanding of the motives of evaluation is essential to the understanding of its processes."¹⁶⁸ These motives have

been succinctly summarised by Roberts, in the context of health education, but may be extended to all health service functions: "We evaluate to aid future planning and to improve programs, to increase our understanding of health education practice, to add to the body of knowledge upon which our work is based. We evaluate to help achieve operational efficiency and, related to this, to obtain data that permit interpretation of program effectiveness so as to obtain administrative support, community support, even financial support. We evaluate for reasons associated with motivation - to give staff and volunteers satisfaction, and a sense of success. To give priority to these purposes ... we evaluate primarily to study the effects of practice so that we can turn our findings back into practice and improve it and, at the same time strengthen the scientific basis of practice in health education."¹¹⁵

WHO has enumerated the managerial process, whether in the health services or elsewhere, as:

1. Scanning the object being managed, and the influences that surround it, and diagnosing present and future problems;
2. Formulating those problems so as to assess their significance, and to define aims and objectives;
3. Generating alternative means of meeting the objectives, examining them, and choosing between them;
4. Obtaining (always subject to constraints) the resources (material, human, and financial) necessary to

implement the chosen means;

5. Defining tasks (of organisation as well as of individuals or groups) in such a way as to make effective use of available skills;

6. Developing and enlarging skills and capabilities;

7. Motivating people to accept the objectives and to work towards them by the chosen means;

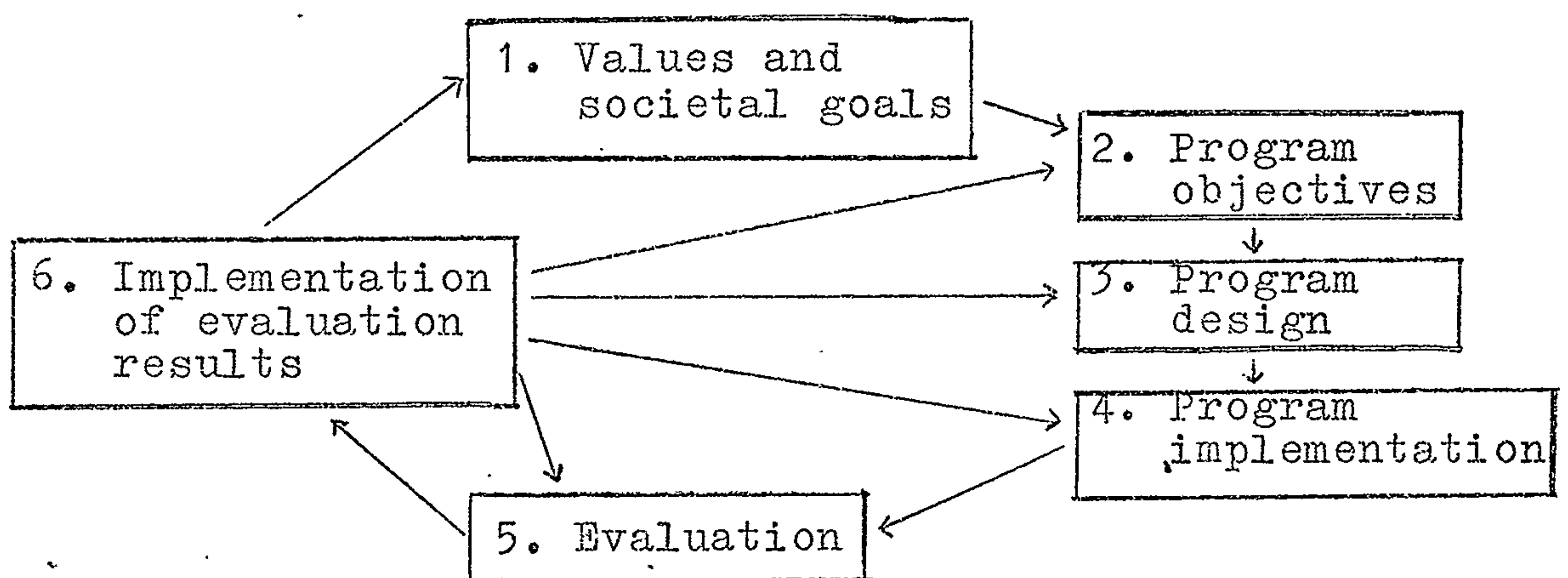
8. Monitoring, controlling, and evaluating so as to adapt the chosen means in accordance with experience.¹⁹²

Program evaluation should be an integral aspect of organization design and development,⁹ and is fundamental to administrative control and indispensable for ensuring that continuous planning is systematically based on experience gained.¹⁹³ "Evaluation may be usefully viewed in the context of a general procedure recognisable in any sophisticated management process, this is, determination of need, development of plans, monitoring of operations, measurement of impact and consequently policy level decisions."¹¹¹

Hogness suggested that in health service terms the core of the evaluation process is its "identifying and isolating the kinds of data policy makers need to reshape existing programmes, initiate new programmes and allocate resources."⁷⁵ Arnold clearly expressed the relationship between evaluation, planning and policy making as "that systematic feedback of information needed for guiding future actions."⁸

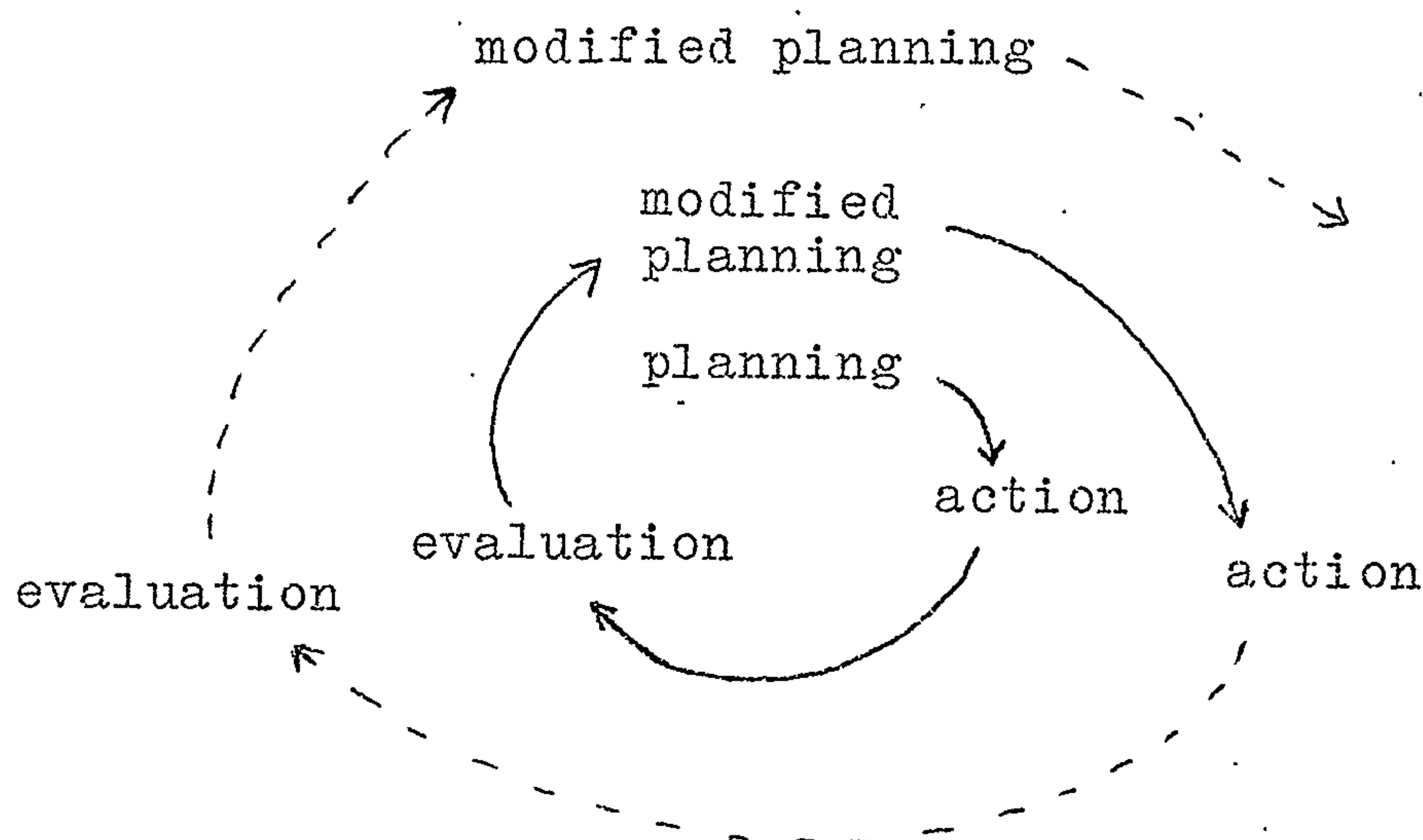
"Evaluation, it has been said, is the mirror image of planning, being intended to determine the value or worth of the programme, to see if it is in fact being carried out as prescribed, and to determine whether the required outcomes and the planned output are actually occurring."¹⁹⁷ The Regional Office for Europe of WHO has shown that evaluation is both the concluding stage of the planning cycle and the starting stage of the next cycle.²⁰⁴ This is shown in Figure 1.

Figure 1. Degree to which Goals are Achieved



This cyclic nature of management was seen by Reinke as comprising a "unified dynamic feedback mechanism"¹¹² rather than independent discrete stages of a process. Caro³⁰ emphasised that the planning - action - evaluation cycle may have to be repeated many times before the stated objectives of a program are realised. His cycle is shown in Figure 2.

Figure 2. Planning - Action - Evaluation Cycle



Haggerty⁶⁸ noted that some workers in the evaluative field have made the distinction between 'formative' evaluation, being that carried out during the course of developing a program to assist in the improvement of the input or independent variable, and 'summative', being the evaluation of the outcome or effectiveness of the program. Edwards, Guttentag and Snapper⁴⁹ thought that formative evaluation described program feedback either to correct errors or to adapt to changing circumstances, while summative evaluation was supposed to be the final verdict on the program. In their view, this distinction was unwarranted, since the ideal evaluation technique would be continuous, i.e., similar to that of Reinke.

Recognition of the need for systematic evaluation of the public sector has led to the introduction in a number of countries of systematic routines of analysis,

in the hope that irrational social decision-making would be reduced and that governmental processes would become more efficiently responsive to the needs of the citizenry.¹⁶² In the UK the principal manifestations of this trend were the increasing reliance on cost-benefit analysis and microeconomic justification for governmental decisions,¹⁰¹ the perfusion of techniques for 'la rationalisation des choix budgétaires' in the French government,⁸⁹ and the adoption in the US of the Planning Programming Budgeting System (PPBS) in 1966.¹⁶²

For public organisations, at least, the products of the evaluation enterprise must meet external as well as internal management and planning needs. That is, program evaluation functions must address the organisation's responsibility to be accountable to its external environments. This external accountability requires that evaluative information be available to meet the administrative needs of other public sector agencies.¹⁰

Although evaluation studies are still in a state of development, their wider use should lead to a fuller appreciation of their potential contribution to the continuous improvement in the efficiency of management. "Better methods of evaluation are urgently needed, from complex systems analysis to simple field studies performed by local health administrators. This kind of research should be built into every national health

plan so that it can be periodically revitalised, in keeping with medical progress and advances in organizational theory and practice. Local and regional health programmes should serve as a permanent laboratory for evaluation studies in public health practice."¹⁹⁴

3.1 Conclusions

1. The evaluation of the organization of health services and their activities pertains to the basic responsibilities and activities of the public health administration;^{65,84} in fact, evaluative research of health programs is not an option but a necessity.¹⁷⁹

2. As a consequence of social pressures, health administrations are now showing an increasing awareness of the need to plan and evaluate their programs in a more systematic and orderly manner than in the past.

3. The ultimate goal of evaluation is to improve the program or to develop another one.⁶⁹

4. The essence of such an evaluation consists of "the systematic and scientific process of determining the extent to which an action or set of actions were successful in the achievement of the predetermined objectives."¹⁹⁵

5. Planning, action and evaluation are the three component steps in the continuous, dynamic, and cyclic process of management.

6. Systematic planning and management of a health program with in-built evaluation processes offers the best hope, not only of wise allocation of resources and of preventing waste, but also of ensuring that programs are adhered to, that targets are reached, and that improvements are effected.²¹⁵ As Isman elegantly phrased it, "accountability through evaluation translates into getting the most 'bang for the buck'."⁷⁸

4. A CLASSIFICATION OF EVALUABLE OBJECTS IN HEALTH CARE SYSTEMS

4.1 Introduction

The question of what is to be evaluated is a pertinent one. It has been discussed succinctly by Bennett and Lumsdaine²¹ as follows:

"One important consideration centers around the question of what is evaluated and, secondarily, in what context and under what aegis. Evaluation can be applied to persons, programs, products, or institutions... The evaluation of persons has a long history... Program evaluation, at least as an object of conscious attention and with respect to study of its methodologies, is a newer arrival on the stage of scientific and public awareness..

"An institution or program may have as one of its goals, or even its sole purpose, the turning out of specific products... Alternatively, the product may be a human 'product' - a skill, capability or characteristic such as clerical competence, business judgement, honesty, or health. The latter kind of 'products' may also be viewed as persons possessing various attributes - educated citizens, trained pilots, cured patients, etc. - in which case 'product evaluation' can be seen as a form of person evaluation. The immediate point to be made here is that the evaluation of a program or institution is not coterminous with

the evaluation of the products - whether they are material or human products - that it produces. For one thing, a generally effective and viable institution may occasionally turn out an unfortunate product...

"Even if a program or institution has a single line of products, evaluation of the product, while a very necessary aspect, is not a sufficient basis for evaluating the institution - a company may fail, despite its turning out a good product, because it produces the product in an inefficient manner or because there is insufficient market for the product. Similarly, evaluation of a social program needs to consider not only its impact in terms of fulfilling its intended social goals, but also, among other things, its cost-efficiency in doing so. To accomplish this latter goal may well require comparative evaluation of alternative programs - that is, programs that have been designed to meet common objectives by alternative methods or procedures.

"We should note both similarities and differences between evaluation of programs and of institutions. Any program designed to achieve given objectives requires some apparatus, organization, facility, or institution to carry it out. This may be either an existing facility, or one set up ad hoc to implement a particular program. Conversely, an existing institution may adopt a new program in addition to other ongoing functions, either to meet its own internal needs or to accomplish program objectives in the service of some agency that retains its services for that purpose.

In either case, evaluation of the program per se overlaps but is not synonymous with evaluation of the agency or institution associated with the program. Methods for evaluation of organizations and institutions - e.g., corporations, foundations, universities, laboratories, or research and development centers - frequently differ greatly from those for evaluating programs. The differences may arise because the goals of an institution are usually more general and diverse than those of a limited, specific program; or because programs, being more like procedures than like structural entities, are likely to be more easily terminated, modified, or replicated in new locations; or because programs have specific impacts that can be more clearly isolated and measured...

"The single term program is commonly used to cover a considerable range of kinds or levels of programs. Within the purview of current parlance, there is a need to distinguish broad, comprehensive, far-flung, continuing, 'macro' programs (e.g. ... the social security program) from narrower, more focused programs of limited scope and duration. The latter represent specific program innovations, treatments or procedures whose outcome may be assessed in terms of specific impacts or effects. These specific impacts, in contrast to the cumulative and multiple outputs of 'macro' programs and of continuing institutions, normally can

more readily be gauged empirically over a definite period of time, in more circumscribed locales, and in comparison with the impacts of alternative treatments, including no-treatment control conditions."

4.2 Concepts of Classification

Donabedian (1967),⁴⁴ in a classical work on the evaluation of quality of health care delivery, classified the objects of such evaluation into three 'dimensions': 'structure', 'process' and 'outcome'. Since then, this classification has been almost universally adopted in the health care evaluation literature. The writer, therefore, has retained these terms for use throughout this work, but has expanded this conceptualisation by the inclusion of a fourth dimension of 'systems/resource utilization'. This dimension has been widely aired in the health service evaluation literature, particularly with respect to systems analysis.

The 'persons' and 'institutions' of Bennett and Lumsdaine fall into Donabedian's dimension of structure, the 'programs' into process, and the 'products' into outcome. WHO, in one of its Public Health Papers, in discussing the organization of health services, used the terms 'organization', 'structure' and 'functioning' to cover Donabedian's structure and process.¹⁹² While McLachlan added a further concept of 'social acceptability' which he subdivided into two facets: the

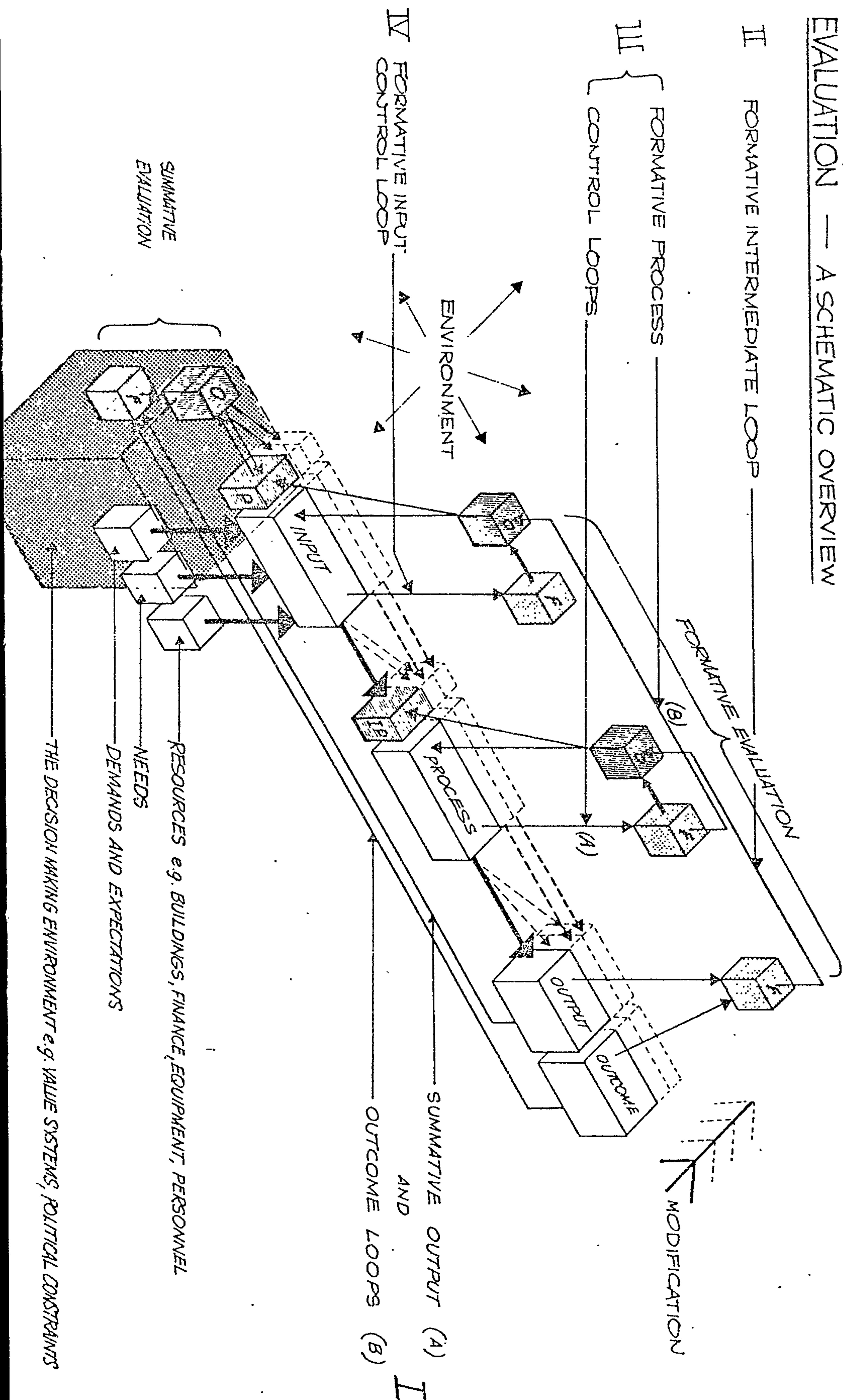
level of quality and standards accepted by the community, and the process of caring for the patient.⁹⁴

Rawson¹¹¹ categorised evaluative objects into four groups: input, process, output, and outcome. Input was perceived as being "an amalgam of all resources, needs, demands and constraints." Output and outcome were jointly seen as being the "products, consequences, effects or impacts in terms of input and process", but were differentiated as follows: "program output may be in conformity with objectives as defined yet outcome may yield no perceptible health benefit."

He developed a comprehensive overview of evaluation which he presented in schematic form. It is shown in Figure 3. The explanatory notes for this very detailed model are given in the succeeding pages. It appears that the model embraces some aspects of the systems view within its goal-attainment framework. His definition of evaluation, as was mentioned in Section 2, considers evaluation as a comparative process; data acquisition, and other steps are considered to be not a part of evaluation.

Figure 3.

EVALUATION — A SCHEMATIC OVERVIEW



<u>Formative evaluation</u>	: a continuous process of monitoring and modifying program design and operation through communication and feedback. Mainly concerned with regulation and control
<u>Summative evaluation</u>	: a technique used to measure effects of impacts of a program by assessing the degree to which goals and specific objectives have been fulfilled


 action alternatives: specific indicators of modification which may be initiated from or directed towards any position where appropriate formulating and response mechanisms exist, for example, process may warrant modification although input remains static, on the other hand, changed input may make process adjustment inevitable

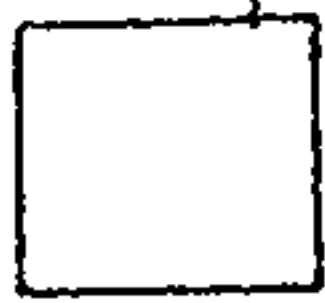
I Summative outcome/output loop: goals and specific objectives are transformed into action by planning using input resources subject to political and other constraints, the program is implemented and outcome commonly examined in terms of fulfillment of initial objectives on various defined criteria.

At this stage comparisons are made, alternatives considered and decisions as to continuation, modification or replacement made, again within an intense value system, political constraint milieu. There are two feedback arms identified: output IA and outcome IB, recognising in the main that program output may be in conformity with objectives as defined yet outcome may yield no perceptible health benefit.

II Formative intermediate loop: examines the relationship between program activities and consequent output. Feedback on fulfillment of process objectives through output can influence immediate planning or impinge directly upon process through inbuilt mechanisms set up for the purpose. The intermediate feedback loop should be ideally continuous and evolutionary in terms of Box's adaptive process control. Monitoring at this level may be either internal or external (say, a government agency) to a program or, optimally, both.

III Formative process control loop: provides immediate operational feedback on whether or not the program is running according to prearranged standards, norms and expectations: immediate planning can be modified while inputs (loop B) or service activities (loop A) are adjusted. There may also be feedback to the various resource inputs including the consumer and professional components.

IV Formative input control loop: ensures that all aspects of input are themselves subject to evaluation. For example, buildings and equipment must conform to recognised standards, qualifications of personnel and operating standards must be adequate, facilities accessible to consumers, realistic needs and expectations of all sectors satisfied, while planning should be flexible.



aspects of the system: agglomerates of resources or information

INPUT : an amalgam of all resources, needs, demands and constraints

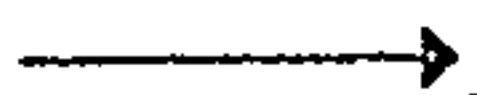
PROCESS : a combination of activities, the service, which is applied in order to accomplish stated goals

OUTPUT)
OUTCOME) : products, consequences, effects or impacts in terms of input and process. Division into output and outcome reflects a common differentiation in the literature, see Summative output/outcome loop

ARROWS : represent flows of resources or information



dark arrows link the various aspects or components of service provision, that is, input, process and output.



light arrows complete the various evaluative feedback loops



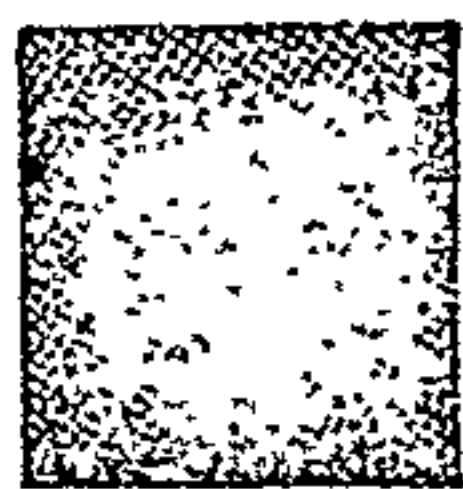
: P - planning, or IP - intermediate planning



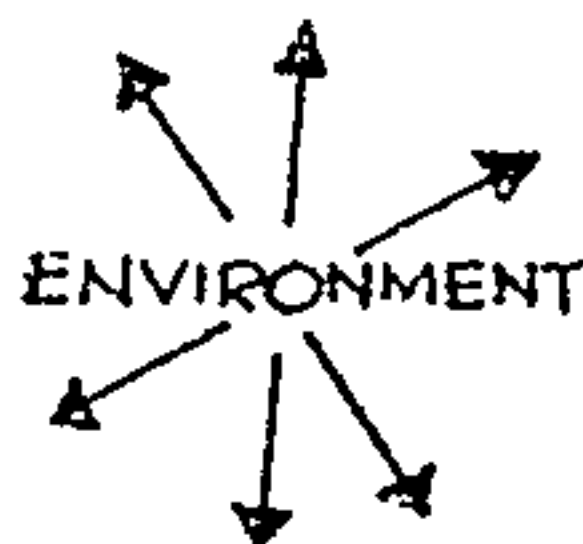
formulating devices: organisational mechanisms, depositories of information generating objectives, evaluative criteria, standards, operating techniques, etc. - in terms of cost constraints, viable program alternatives and overall goals



fulfillment monitors: mechanisms to examine the degree to which objectives of formulating devices have been accomplished: essentially a process of comparison



the planning, decision making environment - the square world: an indicator that overall goal formation is subject to social mores, value system and political constraints and planning. It is noted that the demands and expectations component at input, planning, and goal formulating device and fulfillment monitor are crucially influenced by these factors



overall context indicator: recognition that all health service planning, implementation and evaluation do not take place in a vacuum but are subject to similar constraints at all levels as outlined above. Discrimination of the two recognises that major pressures impinge upon starting, modifying and stopping programs



'dynamic modification' indicator: recognising that the evaluation process is a continuing phenomenon, that is, objectives are continually monitored and planning modified. On the schematic model this indicates a continual shift of aspects of the system or program, formulating devices and monitors, singly or in concert, on the horizontal axis

4.3 Structure Dimension

In this dimension is evaluated the health care organization in terms of its structural adequacy relative to the performance of the tasks for which it was created, and the settings in which health care takes place together with the instrumentalities of which it is the product. "It is concerned with such things as the adequacy of facilities and equipment; the qualifications of medical staff and their organisation; the administrative structure and operations of programs and institutions providing care..."⁴⁶

Martins listed the following aspects as being included in structure:

1. Physical structures, facilities and equipment;
2. General organizational features;
3. Administrative organisation;
4. Staff organisation;
5. Finance, costs and related aspects of organization;
6. Geographic factors.⁹¹

A broader definition of structure was that of an expert committee of WHO: "the human resources, knowledge, technology, organization, facilities, equipment, and finances that assist or constrain the expenditure of effort and the achievement of effects or end-results."²⁰⁰

The assumption is made, in evaluating this dimension, that given the proper settings and instrumentalities, good health care will follow. This approach

offers the advantage of dealing, at least in part, with fairly concrete and accessible information. It has the major limitation that the relationship between structure and health outcome is often not well established.

There has been very little evaluation of structure reported in the literature, and consequently the writer has not pursued this dimension any further.

4.4 Process Dimension

This dimension has been widely used in health systems evaluation, particularly with respect to the operation of hospitals. It is a very broad dimension which covers such aspects as:

1. The extent to which screening and case-finding activities are carried out;
2. Diagnostic activities;
3. Treatment;
4. Consultation and referral;
5. Coordination and continuity of care;
6. Staff turnover;
7. Staff absenteeism;
8. Use of health services by providers;
9. Client complaints;
10. Compliance or non-compliance with health and illness management program;

11. Knowledge about health and illness in general, and any current illness in particular;
12. Changes in knowledge or behavior expected after prior exposure to medical care;
13. Volume of care provided.⁹¹

4.5 Outcome Dimension

The evaluation of outcome looks at the final result of health care intervention. It measures the "ultimate result of the interaction between people and the services being provided expressed in terms of:

- i) Health outcomes, including changes in morbidity and mortality in general, and in relation to specific groups of people and diseases, life expectancy and restoration of physical and social functioning.
- ii) Satisfaction of the clients and providers with the structures, processes and health outcomes."⁹¹

Many advantages are said to be gained by using outcome as the criterion of quality of health care. Donabedian noted that validity of outcome as a dimension of quality is seldom questioned.⁴⁴ Nor does any doubt exist as to the stability and validity of the values of recovery, restoration and survival in most situations and in most cultures. Additionally, outcomes tend to be fairly concrete, and, as such, seemingly amenable to more precise measurement.

However, a number of considerations limit the use

of outcomes as measures of health care. Many factors other than health care may influence outcome, and precautions must be taken to hold all significant factors other than health care constant if valid conclusions are to be drawn. In some cases long periods of time must elapse before relevant outcomes are manifest. Although some outcomes are unmistakable and easily measured, others are not so clearly defined and can be difficult to measure. These include patient attitudes and satisfaction, social restoration and rehabilitation.

Provided measures of outcome are used with discrimination, they remain, by and large, "the ultimate validators of the effectiveness and quality of medical care."⁴⁵

The outcome dimension is evaluated through the goal-attainment framework.

4.6 Systems/Resource Utilization Dimension

This nebulous dimension has received attention from systems analysts and economists. That which is to be evaluated is either 'the system' or the utilization of resources within it. A somewhat detailed examination of this dimension follows.

Webster's Third New International Dictionary¹⁷⁷ defines system under thirteen sections. The relevant sections to this dimension are:

"1a A complex unity formed of many diverse parts subject to a common plan or serving a common purpose.

1b An aggregation or assemblage of objects joined in regular interaction or interdependence: a set of units combined by nature or art to form an integral, organic, or organized whole: an orderly working totality: a coherent unification...

3a The structure or whole formed by the essential principles or facts of a science or branch of knowledge or thought: an organized or methodically arranged set of ideas, theories, or speculations...

9 An organized or established procedure or method or the set of materials or appliances used to carry it out..."

The Oxford English Dictionary¹⁰⁴ devotes more than one page to its definitions of system. The relevant aspects are:

"I An organised or connected group of objects.

1. A set or assemblage of things connected, associated or interdependent, so as to form a complex unity; a whole composed of parts in orderly arrangement according to some scheme or plan...

4. In various scientific and technical uses: A group, set, or aggregate of things, natural or artificial, forming a connected or complex whole.

II A set of principles, etc., a scheme, method...

8. The set of correlated principles, ideas, or

statements belonging to some department of knowledge or belief; a department of knowledge or belief considered as an organised whole; a connected and regularly arranged scheme of the whole of some subject; a comprehensive body of doctrines, conclusions, speculations, or theses."

General System Theory (GST) is considered to have been founded by von Bertalanffy as a general science of 'wholeness'. Its major tenets have been stated to be:

- "1. There is a general tendency towards integration in the various sciences, natural and social.
2. Such integration seems to be centred in a general theory of systems.
3. Such theory may be an important means for aiming at exact theory in the nonphysical fields of science."¹⁷³

As a part of his work on GST, von Bertalanffy defined a system as "sets of elements standing in interrelation."¹⁷⁵ A system is closed if no material enters or leaves it; it is open if there is import and export, and therefore change of its components. Living systems and social systems are open systems.

Management principles, therefore, must begin with the premise that viable organisations are ever-changing internally because they are open systems with permeable boundaries. As open systems, organisations are continually in transaction with their surrounding

environments: that collection of all persons and other organisations that influence or can be influenced by the organisation.²⁷

A very important principle of systems theory is that systems will strive to maintain equilibrium, or a quasi-steady state, because such a condition is more conducive to efficient production. Efficient systems devote the largest share of their available energy to productive outputs or services, using only minimal energy for necessary maintenance and internal coordinating activities. However, in view of the ever changing environment, an ideal steady state is not possible. The adaptive system, therefore, must possess feedback to remain in harmony with its environment.²⁸ Evaluation is the information component of feedback within a system.

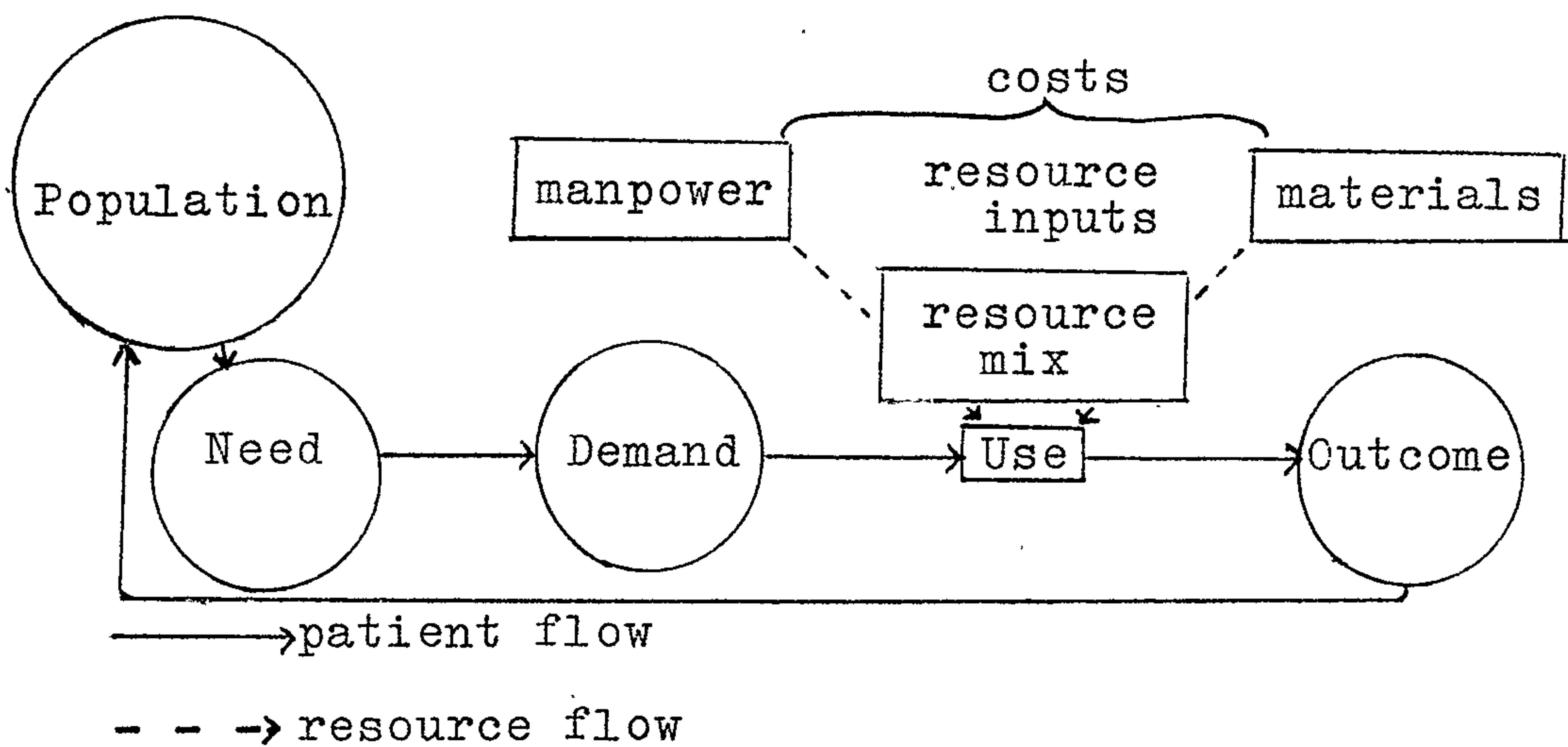
Viewing an organisation as an open system places it within the context of its environment and ensures that management's interest extends to both external and internal affairs. Whereas classical management theory is extensively devoted to internal organizational structures, general systems theory emphasises that a key task of management is to promote stabilized productivity while remaining responsive to changing external pressures and demands.

Afifi et al defined a system as "an aggregation of objects or components united by some form of interaction so as to perform a specific function"⁵, and noted that a system is usually a part of a larger system and

consists of molecular subsystems. Weisbrod¹⁸⁰ did not give a definition of his own, but he considered that central to any discussion oriented towards systems is the notion of 'wholeness'. An expert committee of WHO on the application of systems analysis to health management, stated that "a system is not merely the sum of its parts; it includes the interaction between the parts."¹⁸⁴ The committee exemplified this holistic view by that of an organization which "consists not merely of the boxes of an organogram but also the pattern of interrelationships between them and within them. It also involves the clientele served and the resources consumed (which are inputs into the system) and the services and products (outputs) resulting from the organization's activities." It considered that, however a system may be conceived, it is important to know how the interrelationships operate, how they are managed and how information flows through the system to facilitate management. "While all these properties may be observed in natural systems such as a flower or a forest, they are equally discernible in man-made systems such as those affecting human health and well-being."¹⁸⁴ Nelson and Machin considered that "systems thinking directs one to consider any unit of organization or responsibility centre, not in isolation, but in interaction with the larger organization or environment of which it is a part."¹⁰⁰ Emery thought

that "the analysis of human organizations as systems is the common practice of most of the disciplines concerned with them."⁵¹ The Regional Office for Europe of WHO produced a simple model of a dental health sytem. It is shown in Figure 4.²⁰¹

Figure 4. A Model of a Dental Health System



In contrast to the confidence of others in systems theory, Phillips¹⁰⁷ found it to be a discredited philosophy because it had no predictive value. To him it was, therefore, a pseudo-scientific theory.

Systems science is a term which has evolved fairly recently; it is a broader-based system of enquiry than conventional systems research. It evolved, according to Banathy, because "during the past two decades we have been faced with increasingly more complex technological and social problems. It has not been possible to find solutions to many of these problems by applying the skills of a single analytically-oriented discip-

line. We have had to develop a new way of thinking and a new approach to disciplined inquiry. That new way of thinking is known as systems science."¹⁹

Systems science aims to examine isomorphism among concepts and principles that operate in different fields of discipline and in diverse applications. Specifically, it attempts to design models which can be used across disciplinary boundaries. In the process, it is developing a metalanguage which should allow interdisciplinary communication. This "would be extremely useful for communication among health professionals on the one hand, and economists or mathematicians on the other." Systems science also aims to develop specific generic models within the framework of which people can interact and solve complex problems. The writer has not come across any case of application of this worthy ideal within the health system evaluation literature.

4.7 Conclusions

1. It is pertinent to consider what aspect of health care is to be evaluated.
2. Structure, process, outcome, and systems/resource utilization are considered to be the dimensions that can be evaluated.
3. Some disparity in conceptualisation and definition of some of these dimensions exists.

4. The structure dimension is not a common target of evaluation studies in the health literature.
5. Process is the most commonly evaluated dimension.
6. The outcome dimension is the ultimate validator of the effectiveness of health care, but it may pose operational problems.
7. The systems/resource utilization dimension is nebulous, and much has been written about it from a theoretical angle.

5. A CLASSIFICATION OF OPERATIONAL FRAMEWORKS OF EVALUATION

5.1 Introduction

It is generally agreed in the literature, that one of the most critical and often one of the most difficult phases in program evaluation is the process of clarification of program objectives. This emphasis results from the conception of evaluation as the process of measuring the degree of success or failure of the program in realising its predetermined objectives. Related to this conception, is the assumption that if specific program objectives can be defined, then the appropriate methodology and criteria for assessing the program may be selected correctly. This is the foundation of the goal-attainment framework.

It has been said that the weakness of this approach, inter alia, is that it does not assure that its findings are implemented for improving the success of the program. For example, Schulberg and Baker discovered that "nowhere is any indication found, however, of the manner in which the evaluator can insure closing the circle of the evaluation process in the goal-attainment model. More often than not, the previously linked series of cooperative processes between evaluator and administrator break down at the point of goal modification."¹³³ A comment in similar vein was made by Thompson.¹⁶⁶

Etzioni considered that the characteristics of the goal-attainment framework that render it relatively ineffective at the point of implementation of findings, include: (1) Disagreement between the administrator and the researcher as to the goals of the program; (2) The relatively circumscribed perspective with which this framework views an organization. Since it assumes that specific goals can be evaluated and modified in isolation from other goals being sought by the organisation, it constitutes an artificial, if not fallacious, approach. Organisational theory emphasises the interrelated nature of goals and the manner in which modification of any one is constrained by the characteristics of the other.¹³³ Additionally, Etzioni, in a critique of the goal-attainment framework, considered that one of the major shortcomings of this approach was that "it frequently makes the studies' findings stereotyped as well as dependent on the model's assumptions. Many of these studies show (a) that the organisation does not realize its goals effectively, and/or (b) that the organization has different goals from those it claims to have."⁵³ He classed goals as cultural entities, whereas organisations, as systems of coordinated activity of more than one 'actor' are social systems; and as Parsons¹⁰⁵ noted, there is a general tendency for cultural systems to be more consistent than social systems. This

discrepancy in consistency results from two causes. Firstly, cultural images, to be realised, require investment of means, and since the means needed are always larger than the means available, organisations are always less perfect than their cultural anticipations. Secondly, organisations are multifunctional units which only devote a part of their means directly to goal activities, and devote another part to other functions, such as the maintenance of units performing goal activities and service functions.

Consequently, some writers have concluded that the goal-attainment framework may not be the best one for effective evaluation, and have espoused either a systems analysis or a resource analysis. Notwithstanding these comments, the majority of published articles relating to the evaluation of health care systems still come within the ambit of the goal-attainment framework.

Additionally, confusion has arisen in the literature as a result of the use of terms with different connotations in different fields. By way of example, Rushing considered the use of the terms efficiency and effectiveness within the frameworks of goal-attainment and systems theory. "The concepts of organizational efficiency and effectiveness date to Barnard's²⁰ (1938) analysis. For Barnard, effectiveness was defined in terms of organizational goal attainment and efficiency in terms of satisfaction and cooperation of organizational participants. Thompson (1967:4-6)¹⁶¹

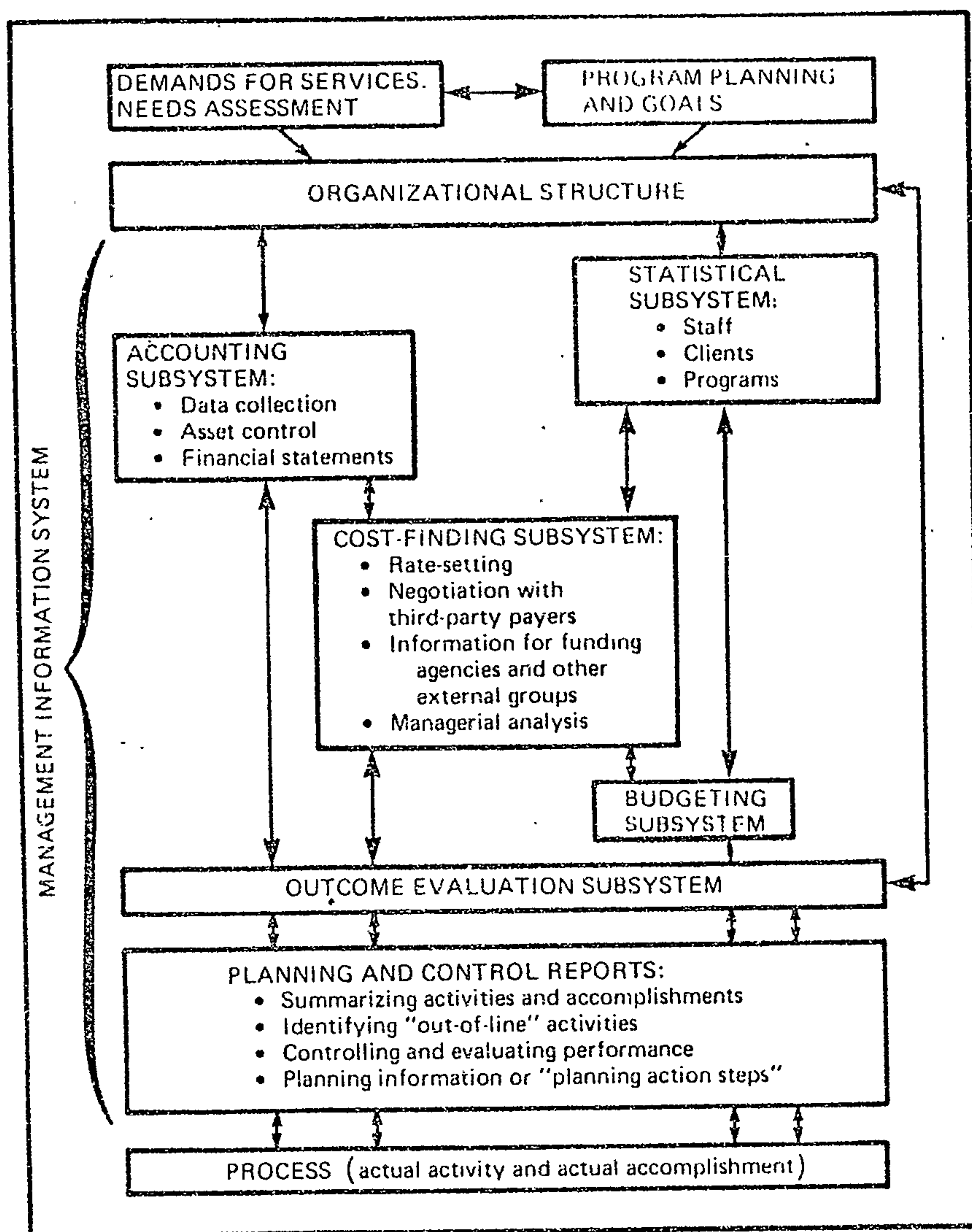
however noted that in scientific management, administrative science, and bureaucratic theory, it is efficiency that is viewed in terms of goal attainment.

Etzioni (1964:8)⁵² spoke of effectiveness in terms of attainment, with efficiency being defined in economic terms."¹²¹

The Management Information System (MIS) is a concept that has emerged only in the last two decades. It is used to describe the system that generates data on all phases of operations, including financial information - both actual and budgetary - and all types of statistical information.¹⁴⁴ The role of the MIS in connecting the organisation's structure to its process is shown in Figure 5.¹⁴⁶ The MIS is seen to embrace both the applied systems approach (through the cost-finding, or cost-analytic, subsystem) and the goal-attainment approach (through the outcome evaluation subsystem). It therefore links both concepts through the measurement of outcome, which is equated with benefit and effect. The practical value of this approach appears not to have yet been explored in the health care evaluation literature.

The operational frameworks which the writer has adopted are: (1) the goal-attainment framework, which already exists in the literature, and (2) a systems/resource analysis framework which encompasses both systems analysis and economic analysis.

Figure 5. A Management Information System



5.2 Goal-Attainment Framework

The overall processes that have been pursued within the goal-attainment framework have been basically similar. The generalised process has been to decide what information is needed, to gather the information, and then to make the information available to those who need it for comparison with pre-established goals.

Stufflebeam et al¹⁵³ defined evaluation in terms of these three stages, which they characterised as: delineating, obtaining, and providing information for decision making. A similar approach was used by Stake¹⁴⁹ who recommended that the program intents be specified and used to determine what observations to make to describe the antecedents, transactions, and outcomes of the program and whether the intents had been realised. This information would then be made available to the program audience (managers and planners) for use in judgements and decisions. Rippey¹¹⁴ proposed a continuous cycle of delineating, obtaining, and providing stages, as the evaluation consultant strives to respond to the developing and changing information needs of program personnel.

Greenberg⁶⁷ discussed components of program evaluation which included: measurement of need through community diagnosis, program design and setting of goals, measurement of service, evaluation of goal fulfillment, and cost benefit analyses and other input-output studies. Hyman and Wright⁷⁶ enumerated five main aspects of evaluation research, but confined their approach to looking at program effectiveness:

1. Conceptualisation and measurement of the objectives of the action program and of unanticipated relevant outcomes.

2. Formulation of a research design and of the crit-

eria for proof of effectiveness of a program.

3. Research procedures themselves, including provisions for estimating and reducing errors in measurement.

4. Problems of index construction and the proper evaluation of effectiveness.

5. Procedures for understanding the findings of effectiveness or ineffectiveness.

A more detailed stepwise list of evaluation procedures has been produced by Pearson¹⁰⁶:

1. Background reading.

2. Delineation of the general problem.

3. Refinement of the problem and development of hypotheses.

4. Formulation of concepts.

5. Description of variables to be measured.

6. Methods of measurement to be used.

7. Development of criteria.

8. Identification of the study population, including controls.

9. Sampling methods to be used.

10. Methods of identification and contact of study population.

11. Preparation of study population.

12. Development and testing of data collection instruments.

13. Organization of available resources for the study.

14. Assessment of resources needed.
15. Acquisition of resources needed.
16. Hiring and training of staff.
17. Testing the feasibility of data collection in the specified setting.
18. Development and testing of the administrative mechanisms.
19. Finalisation of the data collection method and procedures.
20. Development and testing of the methods of data analysis.
21. Interim reports.
22. Analysis, review of results and additional analysis.
23. Final report.

Pearson noted that these items fell into seven categories of: reading, study, population, resource, data collection, administration and reporting, and data analysis.

Straton¹⁵² has produced an evaluation process model (see Figure 6) which he developed from a model suggested by Ten Brink¹⁶⁰, and which is related to a very detailed model by Stufflebeam et al¹⁵³. This model, which he developed in the educational context, appears suitable for the health care field. It is comprised of the following stages:

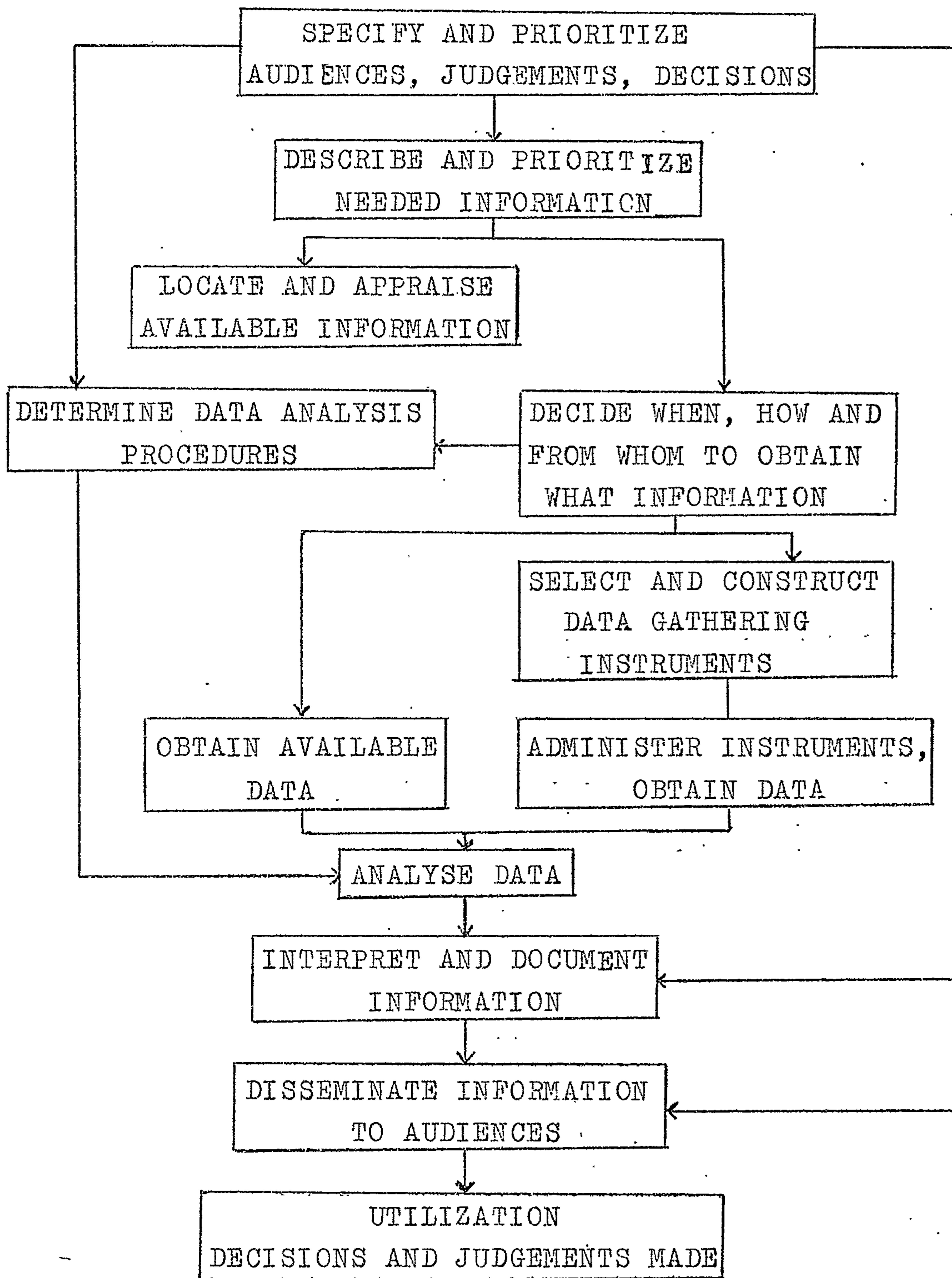


Figure 6. A Model of the Evaluation Process

1. The Delineation Process

This stage involves determining what information is needed to meet the purposes of the study as seen by its recipients (audiences).

2. The Obtaining Process

i. Determination of what information is already available and what are its characteristics, i.e. data accessibility, completeness, reliability, validity, cost of recovery, and comparability across sites and time.

ii. Study Design. At this step, decisions are made of what, when, how, and from whom information should be obtained. Alternative design strategies are considered in terms of constraints and payoffs.

iii. Variable Operationalization. This step considers alternative data acquisition methods within the confines already established during the previous step.

iv. Data Acquisition. The concern of this step is with actually obtaining the samples, administering the instruments, making the observations and recording the basic data specified in the study design.

v. Data Analysis. The choice of particular data analysis methods reflects the choice of indices and methods of data presentation.

3. The Providing Process

This stage is the interface between the evaluator and the recipient of the study, i.e. the communication

of the information which the obtaining process has obtained. The decisions which have to be made at this stage are those relating to the documentation and dissemination of the information.

4. Information Utilization Process

This process is the responsibility of the recipients of the study. It involves the making of judgements and decisions based on the information received from the study.

A six stage sequence of the evaluation process was developed by Levey and Loomba⁸⁷ which was comprised of:

1. Formulation of the objectives.
2. Specification of measures of performance.
3. Development of the model, plans, and programs.
4. Measurement of results.
5. Determination and explanation of the degree of success.
6. Recommendation for appropriate action.

Martins⁹⁰ modified this model essentially by expanding the second step into (i) a statement of values and criteria for the evaluation process, and (ii) the specification of measures of performance.

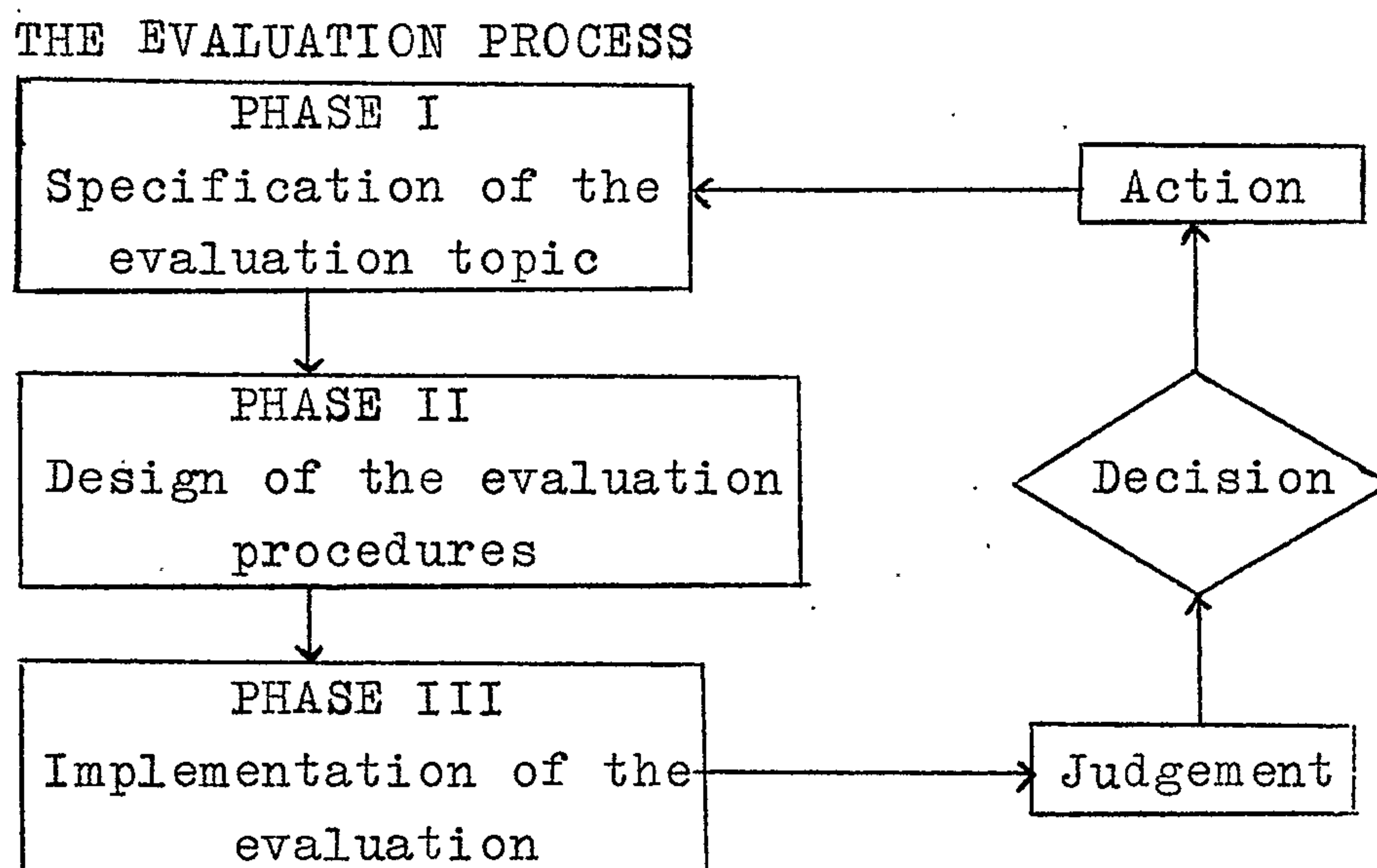
Recently, Haggerty⁶⁸ has listed six stages in evaluation of health programs:

1. Specification of the program goals, this includes that of subgoals.

2. Description of the program goals, and definition of the independent variable.
3. Definition and measurement of dependent variable or outcome. The difficulties encountered with measurement of the dependent variable is not only the reliability and validity of the measures used, but even more important is the issue of what to measure. This necessitates the steps of (i) defining the results of an optimum-functioning program, and then (ii) developing reliable and valid measures to ascertain whether these objectives have been achieved.
4. Assessment of intervening variables, since all evaluation of health services goes on in a politico-social climate where many confounding variables are operating.
5. Analysis of data.
6. Implementation of findings.

A comprehensive model for the evaluation process has been formulated by a WHO expert committee on evaluation of family planning in health services.¹⁸⁶ This model divides the evaluation process into three phases, and places it in the context of decision-making, as shown in Figure 7.

Figure 7. Decision Making and the Evaluation Process



Phase I is the specification of the evaluation topic, i.e., the delineation of precisely what is being evaluated and why. Phase II is the design of procedures for the evaluation of these topics. This phase consists of specifying procedures for making judgements about the evaluation criteria that have been selected, as well as selecting procedures for collecting and processing the required data. Phase III is the implementation of the evaluation; it deals particularly with the organization and administration of evaluation. Each of these phases can be viewed as having a product that is a requirement for commencing the next phase.

The committee subdivided each phase into a number of steps, and noted that the core of the evaluation process is the series of steps that defines the manner in which judgements will be made, i.e., Phase II(A), below:

Phase I

- i. Specification of the subject of the evaluation.
- ii. Specification of the type of evaluation (of need, plans, performance, or effects).
- iii. Specification of the purpose of the evaluation (for general knowledge, to make judgements, to make decisions).
- iv. Specification of the decision options (who is the decision maker, when must the decision be made, what are the constraints on the decision options?).
- v. Finalisation of the evaluation topic.

Phase II

A. Procedures for Making Judgements

- i. Selection of the study design (extensive, intensive, experimental, non-experimental).
- ii. Selection of the criteria that will be used to make the judgements (inputs, outputs, effects, impacts).
- iii. Selection of the comparative standards for determining effectiveness and efficiency (need, demand, objectives).
- iv. Selection of the evaluation measures of the criteria and standards.
- v. Specification of the data analysis procedures (data classification, type of statistical analysis).

B. Procedures for Collecting and Processing Data

- i. Selection of the sampling procedures.
- ii. Specification of the data collection procedures.
- iii. Specification of the reporting procedures.

iv. Finalization of the evaluation design.

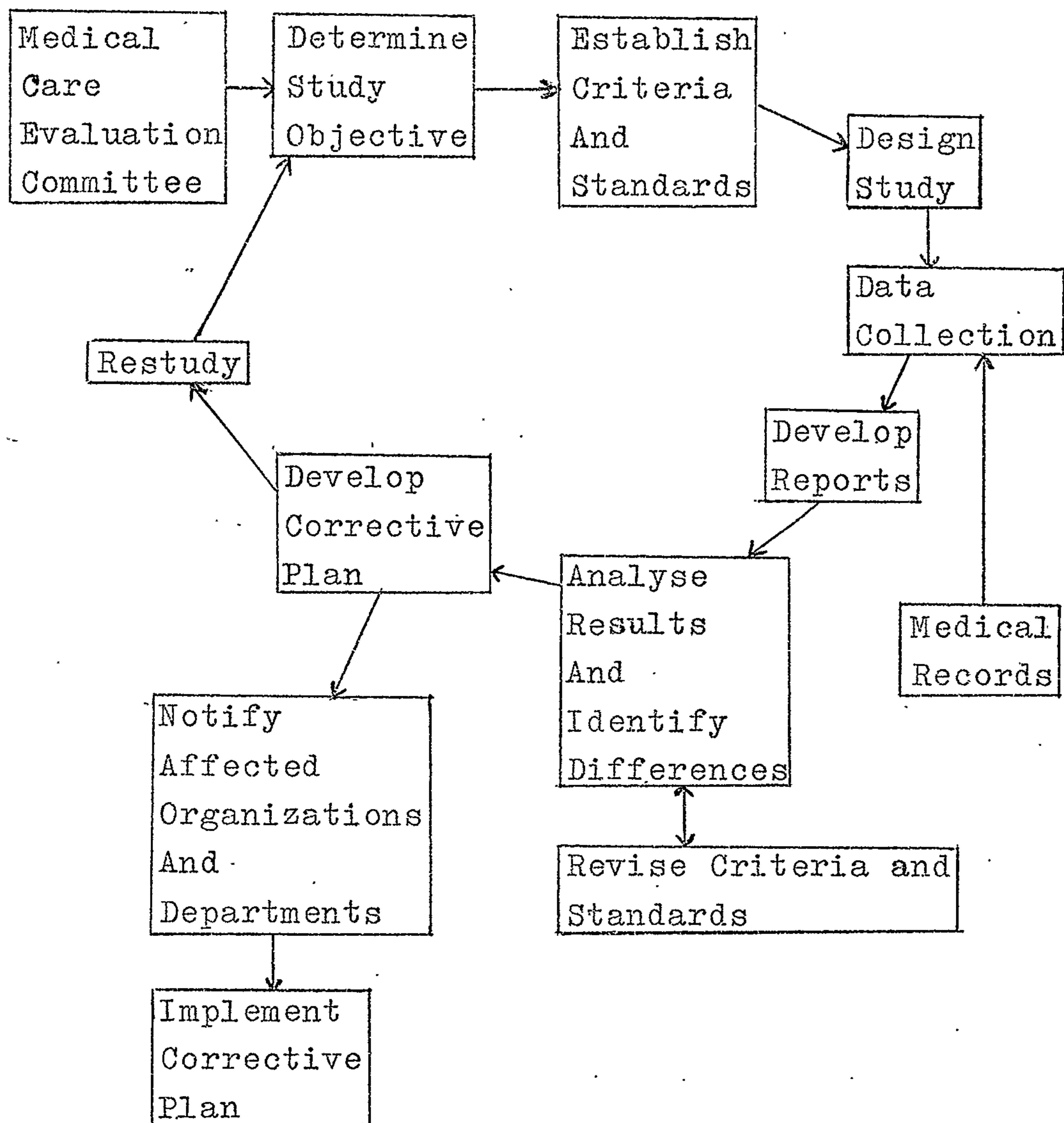
Phase III

- i. Checking the feasibility of the evaluation.
- ii. Designation of the organization of the evaluation.
- iii. Testing and refining of the evaluation procedures.
- iv. Collection, analysis and reporting the results.
- v. Evaluation of the evaluation.
- vi. Completion of the findings and recommendations.

Goran⁶⁶ has formulated a Medical Care Evaluation Model, on lines similar to the WHO model, for the (US) Professional Standards Review Organization (PSRO) Hospital Review System. The model, which is shown in a modified form in Figure 8, lists the following steps:

- 1. Determine the study objective.
- 2. Establish criteria and standards.
- 3. Design the study.
 - i. Study site (location or magnitude of study).
 - ii. Type of study (retrospective, prospective).
 - iii. Sample size.
 - iv. Data collection method.
 - v. Data source.
 - vi. Data process.
 - vii. Manpower and resource requirement.
- 4. Data collection.
- 5. Develop reports.
- 6. Analyse results, and identify deficiencies.

Figure 8. A Medical Care Evaluation Model



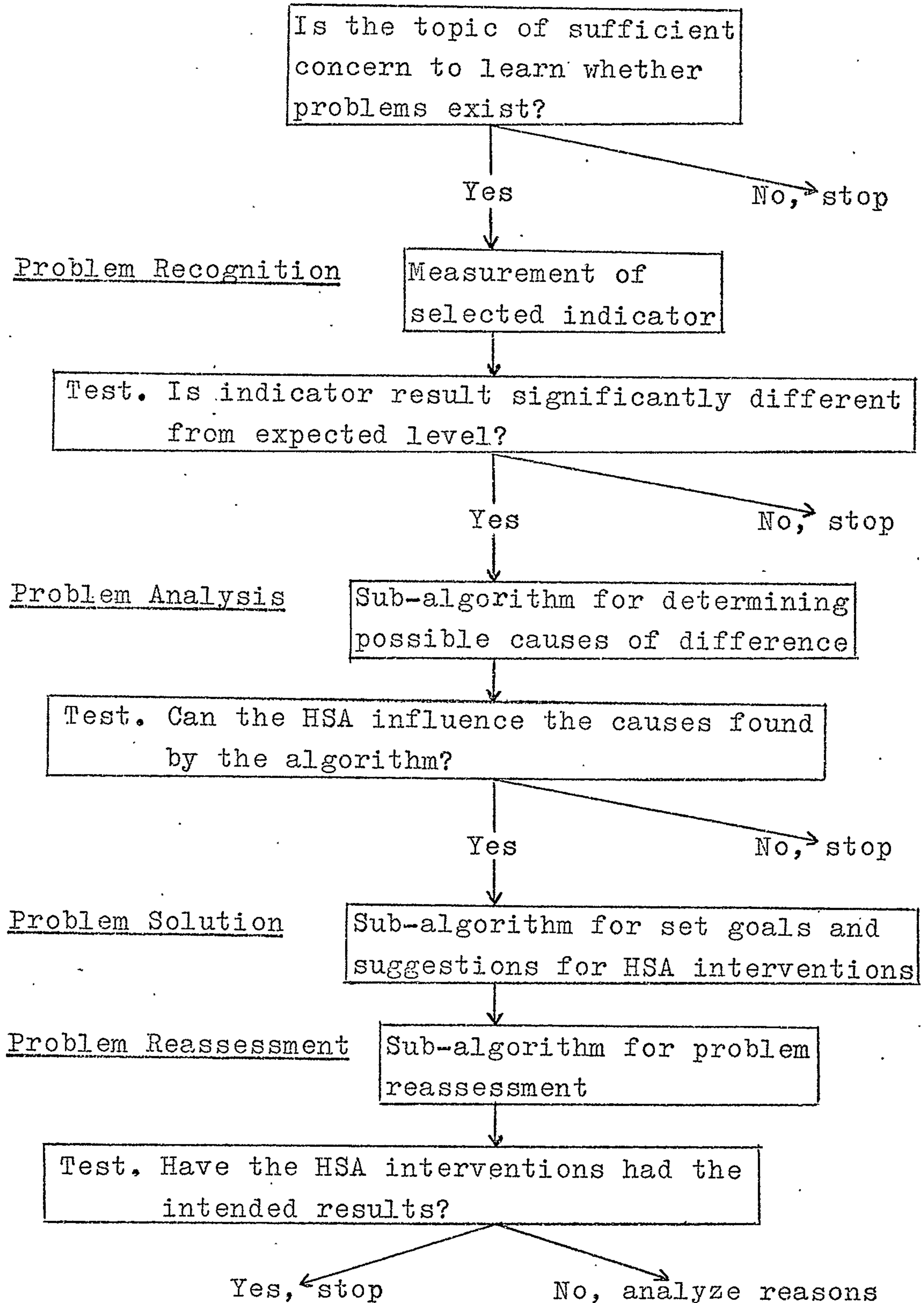
A new outcome methodology has recently been developed by the Rand Corporation⁸² which is said to permit health planners to assess the impact of a health care system on the health status of the population. The methodology is in algorithmic form, and is designed to assist in the development of objectives and actions related to the occurrence of selected health status indicators. It has been designed to utilise readily obtainable data and is said to require no detailed mathematical analysis. It was developed specifically for Health Systems Agencies (HSAs) in the USA. The sections in this evaluation model are:

1. Selection of the evaluation topic.
2. Problem recognition.
3. Problem analysis, in which the HSA is given a detailed list of questions to consider in performing the problem analysis as well as a set of sample standards to be used in answering the questions.
4. Problem solution is based on an evaluation of the preceding section.
5. Problem reassessment.

An outline of the algorithmic approach is shown in Figure 9.

Figure 9. Outline of the Algorithmic Approach
to Outcome-based Health Planning Evaluation

Selection of Evaluation Topic



Stern et al¹⁵¹ have developed a model for specifically evaluating quality of dental care within a framework of quality assurance. The model itself comprises the phases of:

1. Topic selection.
2. Case selection.
3. Utilization of criteria and standards.
4. Selection of data sources.
5. Data collection.
6. Analysis and peer review.

Deming⁴² has summarised the four requirements for an effective system of goal-attainment evaluation:

1. A meaningful operational measure of success or failure, satisfactory to experts in the subject matter, of some proposed treatment applied to specific material, under specified conditions...
2. Some satisfactory design of experiments, tests, surveys, or examination of data already recorded. The design of a new study will include selection of samples of the specified material; a record, for the duration of each phase of the study, of certain specified environmental conditions that appear to be important; procedures for carrying out the investigation; and statistical control to aid supervision of the investigation.
3. Methods for presentation and interpretation of the results of the experiments, tests, survey, or other

investigation, that will not lead to action different from the action that would be taken on the basis of the original data. The data must include a record of the environmental conditions, including test method...

4. Some official or some group of people authorised to take action.

5.3 Systems/Resource Analysis Framework

There are two aspects of analysis which are covered within this framework; firstly, the potential payoff from alternative strategies of investment or procedure, and secondly, the actual payoff received by a system as it operates. Both aspects, it appears, may be examined by either the systems, economics, or accounting disciplines, the choice being dependent on the training of the protagonist.

The systems and economics disciplines are examined later in terms of cost-benefit analysis and cost-effectiveness analysis. These terms are widely used by both systems analysts and welfare economists to denote a methodology which has a somewhat different approach in each discipline. The accounting discipline is not pursued in this work. It is rarely mentioned in the health care evaluation literature.

5.3.1 Systems Analysis

The utility of systems theory to evaluation is said to lie in the application of systems analysis. Systems analysis was broadly defined by a WHO committee as methods of making practical use of the systems view of the world. "In general, these methods seek to define the relationships existing in a system (and between it and other systems) and to calculate the effects of altering either the elements of the system or the ways in which they interact."¹⁸⁴

Systems analysis was succinctly defined by Afifi et al as "the analysis of systems."⁵ A more detailed definition is that of the McGraw-Hill Dictionary of Scientific and Technical Terms: "The analysis of an activity, procedure, method, technique, or business to determine what must be accomplished and how the necessary operations may best be accomplished."⁹³

Couger stated that "systems analysis consists of collecting, organizing, and evaluating facts about a system and the environment in which it operates."³²

Schleip, in his description of the German RPS System of planning and control in management, described system and systems analysis in the following terms: "The system means literally 'that which is put together'... By the system one therefore understands an entity put together according to principles. The modern concept of systems research, often misunderstood as an examination of the optimal system of data

processing, consists of a most exacting analysis of a system, i.e. a construction or a complex of procedures. And the purpose of systems research is almost always optimisation of the system."¹²⁵

A note of caution was introduced by Raiffa who stated that "systems analysis has different meanings for observers in different contexts. For the managers of complex health systems, the discipline focuses upon ways of optimizing system operation... [for others it is] the application of a rational way of thinking and of scientific methodology to resolve complex decision problems."¹¹⁰

As to the application of the systems approach to real world problems, Thompson commented that "though the systems approach has achieved salience as the alternative to the goal-oriented model of evaluation, it puts forward no well-delineated methodology."¹⁶⁶ He considered that the value of the systems approach lay in recognising the basic organisational truth that the pursuit of a unitary goal for output often requires a number of subgoals at the level of input allocation, and that effective organisational goals often differ from the objectives for which the organisation was created. "Arguments that these organisational characteristics require an entirely new approach towards evaluation are not persuasive. The insights are valuable yet appear best applied within the goal model framework for evaluative research."¹⁶⁶

Akman and Gordon were equally critical of many of the claims made for systems analysis. "In recent years, systems analysis has become quite popular and its role as a panacea for many of the world's ills is legendary even if in fact quite untrue."⁶

A somewhat similar view was held by Mechanic, who felt that one of the most pervasive techniques, cost-benefit analysis, had been inappropriately used by administrations that work with problems and outcomes that tend to be intangible and which cannot be quantified in terms of cost or benefit. He considered that administrators should employ paradigms of program planning and evaluation that help make salient those issues which are important.⁹⁵

Walter considered that systems science approaches work best with well defined systems, yet many important components of the health care system, particularly in relation to life style and environmental factors are ill defined. Contrary to Thompson, he thought that the systems approach "intelligently applied at all decision-making levels in the people-intensive health field, seems the best hope for improved understanding of what is happening, could happen, and should happen to control [health care systems] ..."¹⁷⁶

A Public Health Paper of WHO¹⁹² considered that systems analysis, which draws on all available techniques to design a plan of action for solving complex

problems, may be regarded as a procedural and organizational device incorporating many techniques, rather than a technique in itself. Etzioni considered the systems paradigm to begin, not with goals but with a "working model of a social unit which is capable of achieving a goal. It is assumed a priori that some means have to be devoted to such nongoal functions as service and custodial activities, including means employed for the maintenance of the unit itself."⁵³

Robins thought the systems paradigm to be concerned with such issues as intraorganizational strains and flexibility. A similar approach was adopted by Schulberg and Baker who noted that, whereas the goal-attainment approach requires a relatively constant environment, avoids the question of adaptation to change, and ignores the important problem of perpetuation of the program itself, the systems approach addresses such things as organisation, allocation of resources, adaptability, and feedback mechanisms. It is thus concerned with establishing a "social unit which is capable of achieving a goal" and not simply with the achievement of goals themselves.¹³² A practical extension of this view was that of Raeburn and Seymour who categorised the needs and goals being met by the system as Type A, and those required to organise and manage the system as Type B.¹⁰⁹

5.3.2 Economic Analysis

Economic analysis within the resource analysis framework attempts to measure the 'contribution' of a given service to the health of a given population relative to the resources spent on such a service. The use of the term 'contribution' carries the implication that the existence of health services makes something, or someone, better off. But the term is vague, and the study of who is made better off, in what ways, and by how much, by the provision of health services in particular magnitudes and on particular terms is beset with difficulties.³³ These difficulties are partly theoretical (what is meant by 'better off?'), and partly practical (how does one identify and measure the 'contribution' of the health service?).

In attempting to measure this 'contribution', cost-benefit analysis has been extensively used by economists since about the mid 1960s. Most of the published work identifies the value of healthy (or working) time saved as the principal benefit, together with, in the analysis of preventive programs, treatment costs avoided. Almost all studies have commented on: firstly, the difficulties arising from the lack of satisfactory epidemiological models to provide the underlying input-output relationships, and secondly, the difficulty of identifying an output unit on which to standardise cost-benefit analyses.

Nevertheless, "the economist has increasingly be-

come the advocate of the application of cost-benefit analysis to the economic processes ... in those areas in which the public sector was actively involved."²⁰⁸ But "not all economists are agreed that cost benefit is suitable for extension into the most expensive fields of public expenditure: defence, education, health and social welfare. Nevertheless, a growing body of economic opinion does favour the analysis of programmes in these terms, at any rate in the health and education fields."²⁰⁹

Cost-benefit analysis within the economics framework is examined in Section 7.

5.4 Conclusions

1. There are two basically opposing views as to the correct framework for evaluating health care systems: the goal-attainment framework and the systems/resource analysis framework.
2. The inherent defects of the goal-attainment framework are said to be that (i) specific organisational goals cannot be evaluated and modified in isolation from other organisational goals, and (ii) it does not assure that its findings are implemented.
3. The Management Information System is a concept which appears to unite the goal-attainment framework and the systems framework. Its practical value in health care system evaluation is yet to be demonstrated.

4. Many operational models for goal-attainment evaluation have been presented. They vary from simplistic summaries, to detailed step-by-step descriptions.

5. Perhaps the most usefully succinct summary of the goal-attainment sequence has been the Medical Care Evaluation model of Goran.⁶⁶

6. The most detailed model of the steps in the goal-attainment evaluation process has been that of a WHO expert committee.¹⁸⁶

7. What some writers have referred to as efficiency, others have referred to as effectiveness, and vice versa.

8. Although the systems approach has achieved salience as the alternative to the goal-attainment approach, it has no well-delineated methodology.

9. There are opposing views as to the validity and utility of systems analysis.

10. Economic analysis attempts to measure the contribution of a given health service to the health of a given population.

11. Not all economists are agreed as to the suitability of cost-benefit analysis for application in the field of health care evaluation.

6. EVALUATIVE CONSTRUCTS, CRITERIA, MEASURES, AND STANDARDS

6.1 Introduction

The remainder of this work is devoted to an examination of a part of an operational sequence of evaluation. The sequence examined in this section is related to the selection of one or more evaluative 'constructs', to which one or more evaluative 'criteria' may refer. Evaluative 'measures' are then selected for, and assigned to, these criteria. An evaluative 'standard' is used as a reference for the measure to ascertain the extent that actual performance relates to desired performance. The final part of an operational sequence, the making of comparisons and the drawing of deductions, is considered in the next section of this work.

As has been noted in previous sections, the problem of non-standard terminology is rife in the evaluation literature; this section is no exception. For example, what the writer here has called 'constructs', Suchman¹⁵⁴ called 'domains', a WHO expert committee¹⁹⁸ referred to as 'criteria', a working group of WHO²⁰² referred to as 'concepts', Roemer¹¹⁸ called 'levels', and Sax¹²² denoted as 'elements'. What the writer has called 'criteria', have also been called 'criteria' by several other sources.^{25,190}

6.2 Evaluative Constructs

A large number of constructs have been proposed by various writers in the health field, although most published articles do not list more than four, and no article has been found that lists all those mentioned below. In an attempt to present a unified picture of the various conceptions of evaluative constructs, the writer has created his own short classification of these constructs by grouping them into seven categories. These are: goal attainment, resource utilisation, suitability, social responsiveness, flexibility, performance, and miscellaneous constructs.

6.2.1 Constructs of Goal Attainment

1. Effectiveness

Effectiveness is one of the most popular constructs in the literature of health care evaluation.

A working group on the planning and evaluation of dental health services of the Regional Office for Europe of WHO stated that: "Effectiveness is a measure of goal achievement. For instance, in a dental caries preventive programme, effectiveness can be measured in terms of the actual number of teeth which did not become carious, and that otherwise would have. It relates expected outcome to attained outcome, or planned objectives to attained objectives."²⁰² Other similar definitions are, for example, those of Afifi

et al⁵ and a WHO expert committee¹⁹⁸. Another WHO committee defined effectiveness as "the effect of the activity and the end-results, outcomes or benefits for the population achieved in relation to the stated objectives."²⁰⁰

By comparison, Weisbrod, in discussing evaluation of health services within a systems framework, noted that "effectiveness, in principle relates to the measurement of the overall benefits in a real life situation or in circumstances which are comparable with known or assumed realities."¹⁸² A slightly different view was that of Brook et al who defined effectiveness in the following way: "Effectiveness relates to the average benefit of a procedure when used by the average provider in the average community."²⁵

Effectiveness is thus seen to have two distinct meanings; firstly, in the goal-attainment context, it gauges the extent to which a program has achieved its stated objectives, and secondly, in the systems context, it appears to measure benefits.

2. Adequacy

There seems to be general agreement as to the meaning of this construct. It is a measure of a program's actual coverage in relation to its target population.^{199,202}

6.2.2 Constructs of Resource Utilization

1. Efficiency

The simplest definition of efficiency has probably been that of Weisbrod who stated that "efficiency refers to the ratio of the useful work obtained in relation to the energy expended."¹⁸¹ Comparable definitions are: "the effects or end-results achieved in relation to the effort expended in terms of money, resources, and time"²⁰⁰, and "a measure of the resources spent in the process... It relates resource inputs to work output."²⁰² Numerous similar definitions have been recorded.^{15,55}

Agreement within the literature as to the meaning of this term is evident.

2. Efficacy

"Efficacy is a relevant term if the focus of interest is to evaluate alternative methods, procedures and processes. The purpose is to show that the same, or a comparable result, can be produced with less effort, or that the same amount of effort can produce more results or better results."¹⁸¹

"Efficacy: the benefit or utility to the individual of the service, treatment, regimen, drug, prevention or control measure advocated or applied."²⁰⁰

"Efficacy relates to the benefit (or lack of it) of a procedure or treatment when performed under ideal circumstances... A procedure clearly can be

both efficacious and effective. It can also be efficacious and ineffective, but the reverse cannot occur."²⁵

Agreement within the literature is not evident as to the precise meaning of this term.

6.2.3 Constructs of Suitability

1. Accessibility

Accessibility has been defined by Spitzer et al as "those characteristics ... which facilitate or obstruct use by potential clients."¹⁴⁷ Kostrzewski noted that accessibility to health care is one of the features which can be used to observe differences in health care systems.⁸³ He did not define accessibility.

Likewise Sax noted that accessibility is one of the four basic elements which should be taken into consideration in assessing whether or not medical care is 'good', (the other three being quality, continuity and efficiency). He observed that there are two aspects to accessibility: firstly, accessibility to health care by people at the time and place where they need it, and secondly, accessibility to a comprehensive range of facilities, back-up resources and supporting services by the providers of health care.¹²²

Babson considered that "the requisites of accessibility are quantitative adequacy, appropriate geographical distribution, and absence of cultural, economic, and educational barriers to medical care."¹²

2. Universality

Kostrzewski noted that universality of health care is one of the features of health services through which differences in health care systems may be observed.⁸³

No other comment was given.

3. Appropriateness

"Appropriateness: has priority been given to the most useful strategy for attainment of the stated objectives?"¹⁹⁹

"Appropriateness takes into consideration choices or decisions made in planning to see if they really were the best under the circumstances. To find out if a programme is appropriate, the alternative uses of resources, employing different treatment or preventive methods, would have to be reviewed to ascertain if more objectives could be attained by using a different resource mix within the same ceiling of expenditure. For instance, a programme including only fillings and extractions might sometimes achieve more in the long run if some resources were devoted to preventive activities that would decrease the need for them. So again, an effective and efficient programme, adequate in terms of coverage, may be inappropriate in terms of choice of methods to combat disease and in the light of the goals of the system."²⁰²

Babson noted that allocative efficiency (which dictates optimal utilization of resources for any purpose as long as its contribution to the overall

well being of society exceeds that of investment elsewhere) provides a crucial measure of the appropriateness of the level of investment in the health care sector.¹⁶

4. Continuity

Sax considered that for medical care to be judged 'good' it must have continuity. "This means that individuals should have a continuing relationship with providers of service who serve as a source of basic or primary care and as a point of reference for the coordination of other services that may be required."¹²³

Continuity was one of the four canons of medical care proposed by the American Public Health Association, (the others being access, quality and efficiency).¹²

5. Satisfaction

The perceived satisfaction with health services by recipients, has been the basis of many evaluative studies.^{57,61,120} For example, opinion surveys of the attitudes of British people before and after the introduction of the National Health Service were used to evaluate it.⁶⁰

6.2.4 Constructs of Social Responsiveness

1. Acceptability

McLachlan considered that the concept of social acceptability should be considered in evaluating quality of medical care. He considered that it has two facets: firstly, the level of quality and stand-

ards accepted by the community, "for instance the length of waiting lists", and secondly, "the interpersonal, supportive and psychologic aspects of the physician-patient relationship."⁹⁴ He admitted that the second facet is not easily measured.

2. Utilization

This construct refers to the utilization of a health service by the population it is designed to serve.^{13,148,211} No disagreement as to the meaning of utilization is evident in the literature. During the past twenty five years the literature concerned with health service utilization has grown in an explosive fashion.²¹²

6.2.5 Constructs of Flexibility

1. Versatility

Versatility was equated by Kostrzewski with "range of operation (e.g. environmental health; preventive, curative, rehabilitative services; public welfare)."⁸³ He did not offer a definition.

2. Adaptability

Babson considered that "the point at which new techniques are adapted, or the general adaptability of a health system to changed technology as opposed to mere alterations of previous techniques, may also be seen as indices of the quality of medical care."¹⁴

6.2.6 Constructs of Performance

1. Quantity

Quantity of services provided has been suggested by Roemer as an evaluative construct ('level') in assessing health service programs. "The basic assumption of this evaluative level is that certain types of health services (not all types) may be regarded as generally beneficial for people, so that a higher rate of providing these services to a population is deemed more favorable than a lower rate."¹¹⁹

2. Quality

Quality is defined by the Oxford English Dictionary as, *inter alia*, "the degree or grade of excellence."¹⁰⁴ In 1955 Sheps recorded that increasing attention was being paid to the problems of improving and of appraising the quality of health services in general and of hospital care in particular.¹⁴¹ Since then, quality assessment has remained one of the most popular aspects of health care evaluation.

Stern et al found that a search of the literature produced an assortment of definitions of quality assurance and related functions. They attempted to synthesise several of the more prominent definitions into: "Quality assurance is the assessment or measurement of, or judgement about, the quality of care ..."¹⁵⁰

Ware noted that "quality assessment involves meas-

uring the level of quality provided at some point in time."²⁵

Quality is probably the most popular construct in health care evaluation literature.

Quality control is investigated in Section 7.

6.2.7 Miscellaneous Constructs

1. Integration

Integration represents the degree of coordination of all available health care services. It implies "the administrative arrangements by which health care services can effectively be brought to bear according to the medical needs of the individual case."¹⁴

The range and degree of integration of various health services has been suggested as a means of evaluating a health system.⁸⁴

2. Other Constructs

Comprehensiveness and sensitivity were mentioned, but not described, by Roemer in a list of evaluative constructs.¹²⁰ The writer has no doubt that there are other constructs proposed by other workers which he has not encountered.

6.3 Evaluative Criteria

Webster¹⁷⁷ has defined the word criterion as, inter alia, "a characterising mark or trait." This is the meaning that is assigned to it in this section.

A WHO expert committee discussed the basis of selection of evaluative criteria:¹⁹⁰

"This is probably the most important and controversial step in evaluation. In many evaluations conducted, the criteria are not stated explicitly, nor often is the rationale for using certain criteria at all clear. The objective ..., therefore, is to ensure that criteria appropriate to the evaluation topic are selected. The criteria should be related to the subject, purpose, type, decision options, and constraints identified ... [in earlier stages of planning the evaluation project]. In particular they should reflect the values the decision makers use to make judgements. Given this objective it is clear that there can be no standard list of criteria for use in evaluating ... health service programmes. Criteria must be selected to fit the particular evaluation topic."

A number of articles have proposed a variety of criteria for evaluating dental health. For example, the Canadian Dental Association²⁹ listed: treatment level, dental caries prevalence, periodontal disease prevalence, and malocclusion prevalence.

In the development and selection of suitable criteria, certain principles are applicable. Lembcke has listed six of these for medical auditing (a process evaluation). They seem suited also to outcome evaluation:⁸⁶

1. Objectivity

Criteria should be stated in writing with sufficient precision and detail to make them immune to varying interpretations by different individuals.

2. Verifiability

Criteria should be so framed that points on which they rest can be verified by laboratory examination, consultation, or documentation, as appropriate.

3. Uniformity

In view of the essential uniformity of the human body and its diseases, criteria should be independent of such factors as size or location of hospital, qualifications of the physician, or social and economic status of the patient.

4. Specificity

Criteria should be specific for each kind of disease or operation to be evaluated, and all significant and closely related diseases or operations in the same patient should be considered as a unit.

5. Pertinence

To the greatest extent possible, criteria should be pertinent to the ultimate aim of the medical care being evaluated, and they should be based on results rather than intentions.

6. Acceptability

Criteria should conform with generally accepted standards of good quality as set forth in leading textbooks and articles based on scientific study.

6.4 Evaluative Measures

A measure is "a number assigned to an object or event according to rules. That number can be a count, a rate, a ratio, a percentage, a proportion, or an index. For example, new acceptors could be measured by counting the number of them or by computing the percentage of patients who are acceptors."¹⁸⁷

Evaluative measures are selected for, and assigned to, evaluative criteria. By way of examples, the "number of maternal deaths per 1000 pregnancy terminations" is a measure of the criterion "maternal mortality",¹⁸⁸ or, "reduction in the annual incidence of DMFT per child for the population between ages 3 and 6 years" is a measure of the criterion "reduction in the increment of dental caries".

The Regional Office for Europe of WHO²⁰³ gave as examples of measures of the effects of prevention of dental caries using water fluoridation: average number less DMF-teeth per child; average number less cavities per child; average number less teeth lost per child.

Indices are favored measures for health care system evaluation. For example: health utility index; index of activities of daily living; health status index; function status index; index to measure the change in health status that occurs during hospitalization; problem status index; A-Index; G-Index;

Q-Index; UF-Index; and in the dental field there have been, inter alia, oral hygiene indexes, plaque indexes, gingival indexes, and periodontal indexes.^{24,147,171}

6.5 Evaluative Standards

A standard has been defined by the Oxford English Dictionary as "an authoritative or recognized exemplar of correctness, perfection, or some definite degree of any quality."¹⁰⁴ A standard is the reference against which a specific measure is evaluated. It is against this that an evaluative judgement is made.

The accepted standard, for example, of "number of maternal deaths per 1000 pregnancy terminations" might be 0.1 for a given population. The standard for evaluating a reticulated water fluoridation program might be "a 50% reduction in annual incidence of DMFT for the child population between the ages of 3 and 6 years".

6.6 Conclusions

1. The selection of evaluative constructs is a prior step to the selection of evaluative criteria, measures and standards.
2. Differences of terminology and of definition are evident in the literature pertaining to evaluative

constructs.

3. Evaluative criteria must be selected to fit the particular evaluation topic.

4. Indices are favored measures for evaluation.

5. A standard is the reference against which a specific measure is evaluated. It is against this that an evaluative judgement is made.

7. COMPARATIVE TECHNIQUES IN EVALUATION

7.1 Introduction

The two final stages in the evaluation process are those of making comparisons and drawing deductions. The deductions to be made are unique to each case, and the process of making them is not within the scope of this work. The process of making comparisons occurs in both the goal-attainment and systems/resource analysis frameworks. It was viewed by Rawson¹¹¹ as being the essence of evaluation. In the former framework, the comparison is between actual achievement and planned objectives, in the latter framework it is between benefits or effects and costs. The various techniques that have evolved in attempts to formalise this comparison are discussed below.

7.2 Goal-Attainment Techniques

The formalisation of the comparative process is not well developed throughout the health care evaluation literature. The techniques that the writer has noted are those of: medical/dental audit, goal attainment scaling, and quality control.

7.2.1 Medical/Dental Audit

Medical audit was defined by Lembcke⁸⁵ as "the evaluation of medical care in retrospect through analysis of clinical records." Brooke²⁶ viewed it (relative to the operation of community medical

agencies) as attempting "to view function and dysfunction of the operation as related to its stated goal and to serve as an in-service education tool."

Audit is a technique which involves more than just a comparison of results with goals, it also involves data acquisition, data classification, verification of the veracity of clinical records, and checking on accuracy of diagnoses. Lembcke listed its main features by reference to major female pelvic surgery⁸⁶:-

1. Arranging data into meaningful classifications; e.g., classifying major female pelvic surgery according to whether it may result in sterility or castration.
2. Verifying the statements in the clinical record by means of written confirmation, such as
 - a. reports of pathology or x-ray departments, clinical laboratory, or electrocardiographic examination;
 - b. report of a qualified consultant;
 - c. observations made by another hospital, laboratory, or outside consultant;
 - d. statement signed by the patient or proper representative, e.g., in sterilization cases; and
 - e. nurses' notes, chiefly in respect to medicines and treatments.

3. Establishing the accuracy of tissue diagnosis, x-ray, and electrocardiographic and clinical laboratory examinations by submitting representative slides, films, etc., to other laboratories for independent interpretations.

4. Comparing the verifiable facts with criteria established as being necessary or important in

a. confirming the clinical diagnosis, e.g., result of electrocardiogram in the diagnosis of myocardial infarction;

b. justifying surgical operations as to their necessity, e.g., pathology report in primary oophorectomy;

c. classifying surgical operations as to the proper extent, e.g., exploration of common bile duct in cholecystectomy patients with jaundice;

d. providing proper chemotherapy, e.g., penicillin in acute rheumatic fever;

e. determining whether hospitalization was necessary at all;

f. judging whether death may have been preventable;

g. justifying the length of stay in hospital;

h. evaluating the safeguards employed in blood transfusion;

i. other aspects of generally similar character.

5. Comparing the degree of compliance with criteria, as measured by the methods described above, with a

standard degree of compliance found to be characteristic of hospitals of acknowledged merit.

Spitzer et al¹⁴⁷ (describing a method of evaluating primary care units - PCUs -), distinguished between internal audit, as "regular documented review by the health professionals in a PCU of their own activities in the provision of care according to a plan formulated formally within the unit", and external audit, as "regular documented review [by] external examiners in accordance with a set of predetermined procedures."

Morehead,⁹⁸ in discussing the evaluation of quality of medical care in Neighbourhood Health Centres described two types of audit: baseline surveys, and clinical audit. The baseline survey was based on the index approach to auditing. "Selected items considered fundamental to adequate care of adults, small children, and pregnant women are abstracted from selected medical records. Numerical weights are assigned to these items and a score is obtained for each case... The final score is obtained by dividing the sum of all actual scores by the sum of all ideal scores for each patient." The clinical audit was based on "reviews by experienced clinician surveyors in the fields of medicine, pediatrics, obstetrics, gynecology and general surgery. Reviews are also undertaken of the

laboratory and x-ray services with emphasis on their usage in relation to specific patient needs. In addition, an administrative review is undertaken of those areas affecting quality of care, patient volume, physician scheduling, follow-up mechanisms, etc. ... Surveyors apply ratings of good, fair, or poor to the items on each case review card. Each case has the potential score of 100, and the score for each specialty is the average of all cases reviewed." At the same time (1970), Morehead reported that attempts, in conjunction with the Dental Division of the US Public Health Service, were being made to develop a dental audit comparable in content to the medical audits. "The original design was based on the clinical audit approach, but experience demonstrated that a modified index approach provided considerably more 'hard' data, which would permit greater comparability between centres at least in so far as assessing the extent of pathology under treatment is concerned."⁹⁹

Friedman and Schoen⁵⁹ reported in 1972 on a pilot study of the auditing of quality of dental care. The audit involved the study of the records of patients, along with their radiographs, "in order to develop a method for evaluating (auditing) the quality of dental care received." The records were assessed for their recorded examination and recorded treatment for each series of treatments per patient and for

the total series of treatments. A numerical system of scoring was devised which ranged from poor or inadequate (0) to excellent(3). "Some of the findings indicated (1) the evaluation of the patient's examination provided no determinant of quality; (2) that the determination of quality lies in the treatment recorded and the general evaluation of the dental practice; (3) that the pattern of dental care can be assessed by an analysis of a single series of treatments, preferably the first one"

7.2.2 Goal Attainment Scaling

Goal attainment scaling was originally proposed by Kiresuk and Sherman⁸¹ as a comprehensive measure of treatment outcome for mental health centres. It involved the establishment of individualized treatment goals, set for each client by a goal setting committee prior to the random assignment of clients to treatment therapists. For each specified goal, a scale was composed of a graded series of likely outcomes, ranging from least to most favorable, with assigned points ranging from -2 to +2. The treatment was administered, and after a predetermined interval, progress was reviewed by independent raters, a score assigned, and this was transformed into a standardised T-score.

Following publication of the original article,

this type of technique has become one of the most popular strategies for evaluation in mental health and other human service programs. Willer and Miller²⁰⁶ noted in 1976 that in most instances, the application of this technique had been used without adequate scientific rigor, and this had served to weaken the scientific credibility of goal attainment scaling. "It remains to be seen whether goal attainment scaling is sufficiently robust as to allow considerable change in procedure, and yet remain a valid outcome measure suitable for evaluation."

7.2.3 Quality Control

The evaluation of the quality of dental care by either the practitioner himself (self-assessment), or by colleagues (peer review) has been advocated for some while and is supported by a significant segment of the profession in the US, for example.¹²⁶ In that country, interest in such procedures was reported as early as 1846.¹²⁷ However, it was not until after World War II that a number of US public programs required periodic review.³¹ Much of the developmental work in dental quality control was performed by Schönfeld et al^{129,130} and by Friedman⁵⁸. Recently, Bailit et al^{17,18} have published an analysis of certain standards that had been

developed; three components of care were studied: history and examination, treatment planning, and treatment. They found that treatment criteria for restorative dentistry were among the most reliable measures. Milgrom et al in 1978 published the results of a comparison between peer review and self-assessment using a number of Dentist Performance Indices. They found that although there were differences in results between the two methods, "analysis of dependent quality measures confirms our hypothesis that dentists' self-assessments are more critical than in evaluating others, especially in emphasising those factors that influence serviceability of restorations."⁹⁷

Many other articles have been published on quality control.^{4,31,38,39,59,77,128,150}

7.3 Systems Analysis

7.3.1 Cost-Benefit Analysis

The systems analyst uses cost-benefit analysis (CBA) for the evaluation of potential and actual changes within an organisational structure or system. "Cost-benefit analysis attempts to establish workable and uniform indicators which can be used as unambiguous and unarbitrary means of comparing or judging the merits of systems proposals. Such indicators can be most helpful in establishing priorities."¹⁰³

In the systems framework, CBA considers all relevant monetary cost and benefits for alternative methods, and attempts to develop cost-benefit causal relationships. It faces the service measurement hurdles of defining the appropriate cost-outcome relationships and of accumulating accurate total system costs. For example, "in saving an individual's life, direct health service costs and indirect costs (e.g., lost wages and sick pay) would have to be included. Accordingly, a systems perspective is advocated in ascertaining those networks where externalities (external social costs) are incorporated or internalized within the cost-benefit network: cost benefit analyses must be carried out within an intellectual framework that comes closer to incorporating the total system."¹⁴³

In spite of the strength of advocacy by its proponents, the problem of assigning monetary values to non-monetary benefits appears to be leading to the gradual abandonment of CBA as a useful evaluation tool. The writer has not come across many articles on it, of recent times, within the health system evaluation literature. Sorensen and Grove summarised the problems of CBA as: outcome measurement, cost measurement, and the monetary expression of cost-outcome assessment of alternatives.¹⁴³

A modification of CBA was the Output Value Analysis (OVA) that was developed within the mental health

field. It was described by Halpern and Binner as "an evaluative framework that relates specifically to the program and fiscal concerns of the administrator. Applicable to any program or components, the framework focuses on estimating the economic value of a program's output and relating this value to the costs of achieving the output. It is simpler than a full cost-benefit analysis, which would require a much more comprehensive picture of all the costs and benefits involved. It focuses on just two of the basic, direct benefits of any mental health program and relates these to immediate program costs."⁷⁰

A key concept of OVA was the Output Value Index, which equalled estimated value divided by estimated resource investment. Estimated output value consisted of two components: estimated economic productivity and estimated response value. Unfortunately, the estimated response value was a subjective value, and thus OVA, although perhaps intuitively more appealing than CBA, was also methodologically suspect.

CBA is examined also in Section 7.4.2 within the economics discipline.

7.3.2 Cost-Effectiveness Analysis

Seiler defined cost-effectiveness analysis as that procedure by which costs of alternative means of achieving a stated effectiveness, or, conversely, the effectiveness of alternative means for a given cost, are compared in a series of numerical indices. The

objective of the analysis is to isolate the alternative, or combination of alternatives, that either gives the greatest expected effectiveness for a given expected cost, or a given expected effectiveness for the least expected cost.¹³⁵

In systems analysis "the ultimate measure of the effectiveness of any system is the degree to which it satisfies system requirements. Effectiveness is measured in terms of the accomplishment of objectives. A system can be considered effective if it accomplishes the purpose for which it was designed."⁵⁰

Eldin and Croft state that there are two phases of the system design process where system effectiveness should be carried out: firstly, in evaluating alternative solutions, the potential effectiveness of each alternative must be defined for comparative purposes; secondly, in evaluating system effectiveness as an ongoing process during system operation.⁵⁰

Effectiveness involves the measurement of performance according to predetermined criteria. The criteria supply a qualitative or quantitative measure by which to gauge the effectiveness of output in satisfying system requirements. Some systems have only one measure, others have several. A number of easily operated measures will have the advantage of providing more than one way to look at system output. They provide a continuing index to the way in which the system is operating. These effort-accomplishment

measurement methodologies have recognized the lack of monetary output and total cost measurement procedures and have applied a systems perspective linking monetary input measures to nonmonetary output measurements for the effectiveness assessment of a particular service.¹⁴⁴

Cost-effectiveness analysis is thus a subset of cost-benefit analysis. If the measured effects required for cost-effectiveness can be translated into monetary values, a cost-benefit analysis can be developed. Whenever physical or psychological outcomes are converted into monetary measures both cost-effectiveness analysis and cost-benefit analysis are possible.

Whole books have been devoted to the topic of cost-effectiveness analysis methodologies,¹³⁴ and the writer does not intend to try and paraphrase them here. Only a brief listing of models will be given.

1. Ratio Model

The ratio cost-effectiveness model basically measures the 'efficiency' of a system in terms of the ratio of its output (effectiveness) to its input (cost). It is, therefore, merely the division of the system effectiveness index by the system cost index which results in an overall figure of merit as the quotient.¹³⁶

2. Indifference Curve Model

This model is a more complex device than the pre-

vious one and is used rather infrequently in spite of its superiority. It also has several technical limitations.¹³⁷

3. Mathematical Programming

"Mathematical programming is a vast subject that embraces linear, nonlinear, and dynamic programming, the calculus of variations, and optimal control theory ... It would serve no useful purpose to review all of this material in order to demonstrate its potential in the analysis of cost-effectiveness problems."¹³⁸

4. Theory of Games

The theory of games is devoted to the identification of optimal strategies in conflict situations. "In a competitive environment a system of decision based on cost-effectiveness analysis should consider the various possible effects of competitive counter-strategies. In conclusion it cannot be denied that a game theory analysis may become a difficult undertaking ..."¹³⁹

5. Probabilistic Cost-Effectiveness

All of the system cost-effectiveness models described so far have assumed single expected values for both the cost and the effectiveness, "in the limit neither cost or effectiveness is determinable ... they can, at best, be described only by their respective probability distributions ..."¹⁴⁰

Like cost-benefit analysis, cost-effectiveness analysis appears not to be of proven utility in the health field. Boggs (1973), for example, found that "few studies have been conducted for dentistry or medicine that have applied and tested the technics of cost-effectiveness. No investigations could be found in the literature on medical care which used cost-effectiveness in other than theoretical applications."²³ Seiler also noted that "cost-effectiveness analysis does not and cannot mean precisely the same thing to all practitioners in the field", and that "since the optimum solution is not obtainable in the limit, the results must always be a relatively imperfect solution at best."¹³⁵

7.3.3 Program Planning and Budgeting System (PPBS)

This technique relates planning and budgeting considerations to specific program objectives. Hinricks⁷⁴ viewed this system as having four steps:

1. Specification of operational objectives.
2. Assessment of outcomes in meeting the objectives.
3. Consideration of all costs in meeting the objectives.
4. Evaluating alternatives in meeting the objectives. This step is complementary to cost-benefit analysis.

The PPBS was first discussed in the 1950s, first applied in the US Defense Department in 1962, spread

to other US government agencies in 1966, and by 1971 was abandoned by the US federal government.¹⁴⁵

Sorensen and Grove summarised the deficiencies of PPBS as:

1. Lack of operational objectives.
2. Lack of outcome measures.
3. Lack of external cost measures.
4. Lack of cost-outcome assessment techniques.¹⁴³

7.4 Economic Analysis Techniques

7.4.1 Introduction

Economic analysis may be of significance to health care system evaluation through the application of cost-benefit analysis and cost-effectiveness analysis. These two techniques share the same terminology as in systems analysis, and, as will be seen, their utility appears to be as suspect in one discipline as in the other.

7.4.2 Cost-Benefit Analysis

To the economist, "cost-benefit analysis purports to be a way of deciding what society prefers. Where only one option can be chosen from a series of options, cost-benefit analysis should inform the decision-maker as to which option is socially most preferred."³⁷

Economists do not regard cost-benefit analysis (CBA) as a part of systems analysis, but rather as a separate systematic technique of analysis. Historically, the idea of evaluating the net advantages of a capital investment project in terms of society's net utility gain was originated by Dupuit in 1844. He defined what is now known as 'consumers' surplus' as the excess of consumers' willingness to pay for a good or service over and above its market price, as a measure of the net welfare gain from a project. This theory suggested a way of measuring the social return to a capital project. The flow of services from the project, multiplied by their prices, defined the maximum social benefit. Since some purchasers would have been willing to pay more, they obtained something for nothing, an excess of utility which constituted consumers' surplus.⁴⁸

In the 1950s this concept was formally extended to cases where persons who are not direct beneficiaries of a project obtain some overspill benefits. They obtain some utility from a good or service for which they have not paid: a consumers' surplus in a context where the market price to them is zero.³⁵ Thus the evaluation of net social benefits requires the estimation of all consumers' surpluses, to whomsoever they accrue. In recent years there has been increasing emphasis on the definition of social output

as the objective of agencies acting in the public interest, and thus CBA has extended the idea of efficiency to public expenditures.³⁶

Sugden and Williams found that "the very idea of a social objective is problematical. There are many shades of meaning to the words 'social objective' and hence many different interpretations of cost-benefit analysis ... [Nonetheless] ... most practitioners of cost-benefit analysis ... produce analyses that are recognizably akin."¹⁵⁷

At the heart of CBA (in the economics context) lies the concept of potential Pareto improvement. This concept is based on the work of an Italian sociologist, V. Pareto, who published his Manual of Political Economy in 1906. In the language of welfare economics, a change that makes at least one member of a community better off and makes none worse off is a Pareto improvement. A project provides a potential Pareto improvement "if the total sum of money that the gainers from the project would be prepared to pay to ensure that the project were undertaken exceeds the total sum of money that the losers from it would accept as compensation for putting up with it."¹⁵⁸

The generalisation of this concept of CBA may be perceived also as being encompassed within the systems framework, if the socio-economic environment is viewed

as an example of a system. The application of general system theory to social systems was one of the major tenets of von Bertalanffy. This has been discussed by Wolfe, himself an economist: "There are many ways forward for cost-benefit analysis. One exciting possibility, however, is that cost-benefit valuations will be combined with other objectives and their constraints in a total systems analysis."²¹⁰

Procedural differences were discussed by Dasgupta and Pearce, who noted that "it is tempting to think that after some fifteen to twenty years of development, cost-benefit analysis must have reached the stage where a clear consensus of opinion exists about the 'proper' procedures to be followed. This would be a mistake ... Indeed, 'cost-benefit' has become a generic term covering a large range of evaluation procedures which frequently differ in what they include and omit as benefits and costs, and in the way outcomes are valued."³⁶

In terms of practical utility, CBA has not met with success in the health care field. Glass considered that CBA, "at any rate as applied to health services, is as much an attitude of mind as a technique."⁶² Sorensen and Grove stated that although [it is] an appealing idea, cost-benefit analyses are fraught with multiple operational problems including the

especially difficult task of assigning dollar values to benefits."¹⁴³ Prest and Turvey claim that CBA can be viewed as anything from an infallible means of reaching the new Utopia to a waste of resources in attempting to measure the unmeasurable.¹⁰⁸

7.4.3 Cost-Effectiveness Analysis

Sugden and Williams have stated that in economics, cost-effectiveness analysis tries to show how a given level of benefit can be achieved at the minimum cost, or to show how the maximum benefit can be achieved at some given level of cost,¹⁵⁹ or, as expressed by Czamanski, it compares the costs of alternative means of accomplishing a specified objective.³⁴ The key to cost-effectiveness analysis is that it is not necessary to attach any explicit monetary value to the benefits. For example, the decision-maker can state his objective as being to maximise a particular index of output, subject to a budgetary constraint. 'Effectiveness' is therefore "some multidimensional payoff exhibited without any attempt at further evaluation in commensurable terms."²⁰⁷

Levin summarised the case for cost-effectiveness analysis:

"A crucial assumption for performing benefit-cost analyses of alternatives is that benefits or outcomes can be valued by their market prices or those of similar alternatives. Yet, the objectives of many,

if not most social programs, often have no market counterpart ... In such situations the effectiveness of a strategy is expressed in terms of its actual physical or psychological outcome rather than in its monetary value. That is, the monetary measures of resource costs are related to the effectiveness of a program in producing a particular impact. When the effectiveness of programs in achieving a particular goal (rather than their monetary values) is linked to costs, the approach is considered to be a cost-effectiveness rather than a cost-benefit analysis ... In this context, cost-effectiveness analysis enables us to examine the costs of alternative programs for achieving particular types of outcomes, but prevents us from comparing the costs directly with benefits. That is, the cost-effectiveness approach enables us to rank potential program choices according to the magnitudes of their effects relative to their costs, but we cannot ascertain whether a particular program is 'worth it' in the sense that benefits exceed costs, because the latter are generally expressed in monetary units while the former are rendered in units of effectiveness for achieving a particular type of impact."⁸⁸

The practical utility of this theoretical approach appears to have yet to be published in the health care evaluation literature.

7.5 Conclusions

1. The medical/dental audit is a review of patient records.
2. Goal attainment scaling involves reviewing and scoring progress towards predetermined goals.
3. Quality control appears to be a term confined largely to dentistry. It may take the form of either an audit or of an examination of treatment. It may be undertaken as either peer review or self assessment.
4. The audit technique involves the evaluation of process, the goal attainment scaling technique evaluates outcome, while quality control may embrace either dimension.
5. The major proportion of each technique, with the possible exception of goal attainment scaling, is devoted to the obtaining and compiling of data. The actual process of comparison between goals and achievements is made by comparing the obtained evaluative measure against an agreed evaluative standard.
6. Systems analysis uses several evaluative techniques, but involves the notion of quantitative and qualitative comparisons among alternatives.
7. Three techniques within the systems analysis framework have been advocated for the evaluation of health systems: cost-benefit analysis, cost-effectiveness analysis, and the program planning and budgeting system (PPBS).

8. Systems cost-benefit analysis is methodologically suspect since it assigns monetary values to non-monetary outcomes.

9. A number of systems cost-effectiveness analysis methodologies have been proposed, but "in the limit neither cost or effectiveness is determinable".

10. PPBS was abandoned by the US federal government after only five years of use.

11. Economic analysis also employs techniques of cost-benefit analysis and cost-effectiveness analysis, but their methodology appears to be different from that in systems analysis.

12. It appears that economics cost-benefit analysis is intended for social cost-outcome comparisons, and that it is fraught with multiple operational problems including attempting to measure the unmeasurable.

13. Economics cost-effectiveness analysis may have some merit, but has yet to be proven in the health care evaluation literature.

8. SUMMARY

The thesis commenced with an explanation of the rationale for its production: an attempt, firstly, to examine, and hopefully clarify, the conceptualisation of the meaning and purpose of evaluation, and secondly, to examine and clarify certain aspects of evaluation methodology for health care systems.

The lack of consistency in terminological and methodological conceptualisation was noted (and reiterated throughout the thesis).

As an aid to clarification of the confused state of evaluation, the creation of a classification of evaluation was explained. This classification has two facets: the categorisation of evaluable objects into four dimensions of structure, process, outcome, and systems/resource utilization; and the classification of evaluation methods into two operational frameworks of goal-attainment and systems/resource analysis (the latter comprising the systems, economics and accounting disciplines). The systems/resource utilization dimension may be evaluated by the systems/resource analysis framework, the process and outcome dimensions by the goal-attainment framework, while the analysis of structure falls outside the scope of this work.

An operational sequence was introduced, which involves the selection of constructs, criteria, measures and standards, for use within the techniques of evaluation.

The introduction concluded by noting that evaluation is a very confused subject.

Conceptualisations of definition and meaning of evaluation were introduced by recording their diversity, both in standard dictionaries, and to a greater extent, in the health care literature. It was noted that evaluation is rarely explicitly defined in the literature, and that with one exception, none of the cited literature definitions corresponded with the cited dictionary definitions.

It was found that within the systems/resource analysis framework evaluation signifies the appraisal of alternative courses of action, while within the goal-attainment framework it generally denotes the assessment of the degree of achieving some specified objective.

The distinction between research, which is related to the formulation and analysis of hypotheses, and evaluation was noted.

It was observed that recently there has been a trend towards defining evaluation in more operational terms.

The need for ongoing evaluation of health systems was exemplified by statements from Australia, Britain and the US.

The health market, it was noted, is virtually unique in the lack of mechanisms for self-regulation, which are so widely practised in private industry.

The purpose of evaluation was found to include: program and planning improvement; and increased operational efficiency, effectiveness and staff motivation.

The place of program evaluation as an integral part of organisational management was emphasised. It was placed in a cyclic perspective of: planning for program objectives, implementation, evaluation, modified planning, modified implementation, evaluation ...

It was found that some workers in the evaluation field have distinguished between formative evaluation (being that carried out during the stage of program development), and summative evaluation (being the final assessment of outcome). However, the distinction between the two disappears in a continuous evaluation cycle.

It was found that a number of systems of evaluation have been introduced in various countries for socio-economic reasons, but recognition of the need for for an urgent improvement of these methods of evaluation is evident.

The concepts of classification of evaluable objects in health care systems were introduced by reference to Donabedian's classic triplet of structure, process and outcome. A fourth dimension of systems/resource util-

isation was added by the writer. Some differences in terminology were noted. Rawson's comprehensive model of evaluation was presented.

Structure was seen to comprise: human resources, knowledge, technology, organisation, facilities, equipment and finance. The assumption made in evaluating this dimension is that given proper settings and instrumentalities, good health care will follow, but since this assumption has not been validated, this dimension was not pursued.

Process was seen to have been widely used in health system evaluation, particularly of hospitals. It was seen to comprise such activities as: diagnosis, treatment, referral, clients' use of health services, knowledge about health and illness, and volume of care provided.

Outcome was seen to refer to the ultimate result of the interaction between people and health services in terms of health and satisfaction. It was found that the validity of outcome as a gauge of quality is seldom questioned, but that its utility is limited by a number of factors such as the effect of other variables and the interval before some relevant outcomes are manifest.

The systems/resource utilisation dimension was examined in some detail. The word system was seen to have many meanings, most of which could probably be encompassed within that of 'sets of elements standing in interrelation'.

The description of systems was introduced through von Bertalanffy's General System Theory, which had been formulated as a science of 'wholeness'. It was seen that systems may be either closed or open, and that organizations and social systems are examples of open systems. Thus it was stated that classical management theory was deficient in that it regarded organizations as closed systems. Systems theory has been said to overcome this deficiency as a result of its holistic analysis.

Phillips found that systems theory was a pseudo-scientific theory because it had no predictive value.

Systems science was seen as a relatively new innovation which aims to develop generic models for the interaction of workers in different disciplines. No application of this ideal could be found.

It was concluded that the systems/resource utilization dimension is nebulous, and that much has been written about it from a theoretical viewpoint.

Operational frameworks of evaluation were seen to fall into two categories: the goal-attainment framework, and the systems/resource analysis framework.

The arguments against the use of the goal-attainment framework seemed to fall into two groups: firstly, the inability of the evaluator to ensure that his findings are utilised by the administrator, and secondly, that specific organizational goals cannot be evaluated

in isolation from other organizational goals. In spite of these arguments, the majority of published articles on the evaluation of health care systems utilise the goal-attainment framework.

Ambiguity in terminology was found to be evident, particularly with respect to the words effectiveness and efficiency.

The Management Information System was seen to be a recent concept that embraces both frameworks. Its utility for health care evaluation appears to be unexplored.

A number of models of goal-attainment evaluation were cited. They varied from brief summaries to extended stepwise descriptions. An evaluation process model by Straton was described that contained four major sections of: delineation, obtaining, providing, and utilization. An expert committee of WHO had proposed a model of three phases: specification of the evaluation topic, design of evaluation procedure, and implementation of the evaluation. Each phase was subdivided into a number of steps.

Perhaps the most usefully succinct summary of the goal attainment sequence was the Medical Care Evaluation model of Goran, the steps of which were: determine study objectives, establish criteria and standards, design study, data collection, develop reports, analyse and identify differences, develop corrective plan, restudy.

It was noted that the Rand Corporation had developed an algorithmic methodology specifically for US health service agencies.

Deming's four requirements for an effective system of goal-attainment evaluation were listed: a meaningful operational measure of success or failure; some satisfactory design of experiments, tests, surveys, or examination of existing data; methods for presentation and interpretation of the results of the previous requirement; a person or persons to take action.

The systems/resource analysis framework was seen to cover two aspects of analysis: firstly, the potential payoff from alternative strategies of investment or procedure, and secondly, the actual payoff received by a system as it operates. It appeared that both aspects could be examined by either the systems, economics, or accounting disciplines. The first two of these disciplines were discussed further on, the accounting discipline was not pursued.

The utility of systems theory to evaluation was found to be advocated through the application of systems analysis, of which a number of definitions were cited. Caution was introduced by Raiffa's statement that systems analysis has different meanings for observers in different contexts. The utility of systems analysis to real world problems was questioned by Thompson who found that it put forward no well-

delineated methodology. He considered that the goal-attainment framework was the best for evaluation. A number of other workers were quoted who supported this contention, and a few were cited who opposed it.

Economic analysis was seen to attempt to measure the 'contribution' of a health service to a community. It was found that most published work in the discipline had attempted to measure this contribution by cost-benefit analysis. Almost all such studies had commented on the difficulties arising from the lack of satisfactory epidemiological models to provide the requisite input-output relationships, and the difficulty of identifying an output unit on which to standardize benefits. It was found that not all economists are agreed as to the suitability of cost-benefit analysis for application in the field of health care evaluation.

Evaluative constructs, criteria, measures and standards were examined for their role in an operational sequence of evaluation. This sequence relates to the selection of evaluation constructs to which criteria may refer. Measures are then selected and assigned to these criteria, and a standard is used as a reference for the measure to determine the extent that actual performance relates to desired performance.

The problem of nonstandard terminology in this area was specified.

Constructs were grouped into seven categories of: goal attainment, resource utilisation, suitability, social responsiveness, flexibility, performance, and miscellaneous.

Criteria were defined as characterising marks or traits. It was noted that the selection of criteria "is probably the most important and controversial step in evaluation", and that "since they should reflect the values of decision makers ... there can be no standard list of criteria" for use in evaluation. However, six principles for the development and selection of criteria were listed, these were: objectivity, verifiability, uniformity, specificity, pertinence, and acceptability.

A measure was defined as a number assigned to an object or event according to rules. It was found that indices are favored measures for health care evaluation; a number of examples was listed.

A standard was defined as an authoritative or recognized exemplar of correctness, perfection, or some definite degree of any quality. A standard is a reference against which a specific measure is evaluated. It is against this that an evaluation judgment is made.

The various comparative techniques in both the goal-attainment and the systems/resource analysis frameworks were examined in some detail.

The goal-attainment techniques examined were: medical/dental audit, goal attainment scaling, and quality control.

Audit was seen to be the evaluation of health care in retrospect through analysis of clinical records. A distinction was noted between: internal and external audit; and baseline surveys and clinical audits. A report on a pilot dental audit was summarised: it found that the pattern of dental care could be assessed by an audit of a single series of treatments.

Goal attainment scaling was briefly discussed; it was seen to involve the establishment of prior individualised treatment goals, and their subsequent assessment by independent raters. It originated in the mental health field, and had quickly spread to other health areas with resulting loss of scientific rigor. Its validity for general evaluation has yet to be proved, but if this does come about, its future appears promising.

Quality control was seen to be an early introduction to US dentistry. Most of the recent developments were done by Schonfeld et al and by Friedman. The resulting methodology involves a posttreatment audit and/or a clinical examination. These may be performed either by self-assessment or by peer review. The subject was not examined in detail, but it was noted that many articles had been published on it.

Systems analysis techniques were then examined. These were cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), and program planning and budgeting system (PPBS). It was found that CBA attempts to establish workable indicators for comparing the merits of systems proposals. It appeared that the problem of assigning monetary values to non-monetary benefits is leading to the gradual abandonment of CBA as an evaluation tool. Output value analysis was presented as a modification of CBA, but it was found that one of its benefit components is subjectively determined and is thus as suspect as is CBA.

CEA was seen to be an attempt at overcoming the deficiencies of CBA. It is a subset of CBA in that outcomes are not equated with monetary values, but are related through a numerical index. A list of five CEA models was given: ratio model, indifference curve model, mathematical programming, theory of games, and probabilistic cost-effectiveness. None of the models, it appeared, were clear of theoretical or practical deficiencies; and terminological differences were evident. CEA appears to have had very little airing in the health literature.

PPBS appeared to have been a hybrid somewhere between the goal-attainment and systems frameworks. It was tried in US government departments and quickly abandoned due to deficiencies somewhat similar to those of CBA.

Economic analysis techniques were introduced as CBA and CEA, but it was noted that although the terminology is shared with systems analysis, the methodology is different.

CBA was seen to purport to be a way of deciding on society's choice for investment alternatives. Historically, CBA had arisen from the concepts of consumers' surplus and potential Pareto improvement. It was seen that (i) CBA has not met with success in the health field, (ii) there is no consensus on the proper procedures, (iii) CBA is as much an attitude of mind as a technique, and (iv) it is fraught with multiple operational problems.

CEA was seen to enable the ranking of potential program choices to be made on the basis of 'effectiveness', but not to allow the determination of whether a particular program is 'worth it'. It was concluded that its practical value appears not to have been explored in the literature.

The final conclusions which were drawn from this thesis are presented in Section 9.

9. CONCLUSIONS

1. There is an urgent need for a standardised, operational terminology of evaluation.
2. Broadly, evaluation is seen as having two distinct meanings; either (i) the assessment of the degree of success in reaching a predetermined objective, or (ii) the appraisal of alternative courses of action.
3. Continuous evaluation is an essential part of management, and no health service should operate without an active evaluation component.
4. Further work is needed on the conceptualisation of evaluation, particularly from an operational viewpoint.
5. The classifications of evaluable objects and of operational frameworks of evaluation, adopted in this thesis, appear to be valid and useful.
6. An examination of evaluable objects shows that:
(i) structure is the most rarely, and process the most commonly evaluated dimension, (ii) outcome is the ultimate validator of effectiveness, but may pose operational problems, (iii) the relationship between the evaluation of process and outcome is the subject of current debate, and (iv) the systems/resource utilisation dimension is neb-

ulous. At the present time, therefore, the evaluation of a judicious combination of process and outcome criteria seems to be the most practical approach.

7. The goal-attainment framework is the appropriate framework for the evaluation of health services.
8. The systems approach has no well-delineated methodology, and appears to have no formal merit for health service evaluation.
9. Economic analysis attempts to measure the contribution of a given health service to the health of a given population; but not all economists are agreed as to the suitability of cost-benefit analysis for this purpose.
10. Evaluative criteria, measures, and standards need to be specified for each evaluation project. The selection of criteria has been said to be the most important step in evaluation.
11. All three goal-attainment techniques, i.e., audit, goal attainment scaling, and quality control, appear to have merit. Audit is a relatively well developed technique for evaluating process (at least in the hospital sphere); goal attainment scaling has been used to evaluate outcome (particularly in the mental health field), and

it seems worthy of further rigorous development; quality control is a rather loose term that involves either peer review or self assessment of either process or outcome.

12. Systems analysis appears to have no merit as an evaluation technique: systems cost-benefit analysis is methodologically suspect, and in the calculation of systems cost-effectiveness analysis neither cost nor effectiveness is determinable in the limit.
13. Economic analysis may be relevant to health service evaluation through the technique of economics cost-effectiveness analysis, but economics cost-benefit analysis, which attempts to measure the unmeasurable, seems to have no merit.
14. Finally, the writer agrees with the conclusion of Wholey,²⁰⁵ that "the art and techniques of evaluation are indeed underdeveloped".

10. BIBLIOGRAPHY

1. Abel-Smith B. Value for money in health services: a comparative study. London, Heinemann, 1976. (p.v).
2. ibid. (p.155).
3. Abramowitz J, Berg LE. A four-year study of the utilization of dental assistants with expanded functions. J. Amer. Dent. Assoc. 87:623-635, 1973.
4. Abramowitz J, Mecklenberg RE. Quality of care in dental practice ; the approach of the Indian Health Service. J. Public Health Dent. 32(2), Spring 1972.
5. Afifi AA, Bigelow JH, Glass N. Some terminology revisited. p.341-2. (In Baily NTJ, Thompson M, eds. Systems aspects of health planning. Amsterdam, North-Holland, 1974. xv + 347p.)
6. Akman A, Gordon JB. Systems analysis and occupational health. Am. J. Public Health. 60:1749, Sept. 1970.
7. Anderson O. (Reported in First National Conference on Evaluation in Public Health. Ann Arbor, 1955. p.7).
8. Arnold MF. Administering health systems. Chicago, Aldine Atherton, 1971. p.263.

9. Attkisson CC, Broskowski A. Evaluation in the emerging human service concept. p.22. (In Attkisson CC, Hargreaves WA, Horowitz MJ, Sorensen JE, eds. Evaluation of human service programs. New York, Academic Press, 1978. xv + 492p.).
10. *ibid.* (p.23).
11. *ibid.* (p.24).
12. Babson JH. Health care delivery systems: a multinational survey. London, Pitman, 1972. vii + 128p. (p.10).
13. *ibid.* (p.12).
14. *ibid.* (p.16).
15. *ibid.* (p.17).
16. *ibid.* (p.21).
17. Bailit H (Chairman) et al. Report of the Quality of Care Committee to the House of Delegates, Connecticut State Dental Association, May 21, 1974. Appendix A. Criteria for the assessment of general dental care.
18. Bailit H, Koslowsky M. Quality of dental care. 1. Development of standards. J. Amer. Dent. Assoc. 89(4):842, Oct. 1974.
19. Banathy BH. The value of an occupational health model: a systems approach. J. Occupational Med. 16(3):164-165.

20. Barnard CI. Functions of the executive.
Cambridge, Mass., Harvard U.P., 1938.
21. Bennett CA, Lumsdaine AA. Social program evaluation: definitions and issues. p.5-9. (In Bennett CA, Lumsdaine AA, eds. Evaluation and experiment. New York, Academic Press, 1975).
22. Bigman S. Evaluating the effectiveness of religious programs. Rev. Rel. Res. Winter 1961, p.99.
23. Boggs DC. Applying the technics of cost-effectiveness to the delivery of dental services. J. Public Health Dent. 33(4):233, Fall 1973.
24. Brook RH, Davies-Avery A, Greenfield S, Harris LJ, Lelah T, Solomon NE, Ware JE. Assessing the quality of medical care using outcome measures: an overview of the method. Med. Care 15(9) Supplement, Sept. 1977.
25. *ibid.* (p.1).
26. Brooke R. An audit of the quality of care in social medicine. Milbank Mem. Fund. Q. 46(3): 351, 1968.
27. Broskowski A, Driscoll J. The organisational context of program evaluation. p.51. (In Attkisson CC, Hargreaves WA, Horowitz MD, Sorensen JE, eds. Evaluation of human service programs. New York, Academic Press, 1978. xv + 492p.).
28. *ibid.* (p.53).

29. Canadian Dental Association. The evaluation of Canadian dental health: a system of recording and statistical analysis at the community, provincial, and national level. Ottawa, 1959.(p.21).
30. Caro FG. Evaluation research: an overview. p.3-4. (In Caro FG, ed. Readings in Evaluation Research. New York, Russell Sage Foundation, 1971).
31. Cons NC. Method for posttreatment evaluation of the quality of dental care. J. Public Health Dent. 31(2):104-108, Spring 1971.
32. Couger JD. Evaluation of business system analysis techniques. p.43. (In Couger JD, Knapp RW, eds. Systems analysis techniques. New York, Wiley, 1974. xii + 509p.).
33. Culyer AJ, Wiseman J, Walker A. An annotated bibliography of health economics. London, M.Robertson, 1977. xvii + 361p. (p.155).
34. Czamanski DZ. The cost of preventive services. Lexington, Mass., Heath, 1975. xiii + 109p. (p.11).
35. Dasgupta AK, Pearce DW. Cost-benefit analysis. Theory and practice. London, Macmillan, 1972. 270p. (p.11).
36. ~ ibid. (p.14).
37. ibid. (p.19).
38. De Jong N, Dunning JM. Methods of evaluating the quality of programs of dental care. J. Public Health Dent. 30(4), Fall, 1970.

39. Demby NA, Rosenthal M. Quality assurance in a neighborhood health center: dental services. J. Amer. Dent. Assoc. 96, June, 1978.
40. Deming WE. The logic of evaluation. p.53. (In Struening EL, Guttentag M. eds. Handbook of evaluation research. Vol.1. Beverley Hills, Sage, 1975. 696p.).
41. *ibid.* (p.55).
42. *ibid.* (p.56).
43. Department of Health and Social Security. National Health Service reorganisation. Consultative document. London, 1971.
44. Donabedian A. Evaluating the quality of medical care. p.168. (In Mainland D. ed. Health services research. Milbank Memorial Fund. 1967).
45. *ibid.* (p.169).
46. *ibid.* (p.170).
47. Drummond MF. Evaluation and the National Health Service. p.67. (In Culyer AJ, Wright KG. eds. Economic aspects of health services. London, Robertson, 1978. ix + 190p.).
48. Dupuit J. On the measurement of the utility of public works (1844) translated from the French, in International Economic Papers, No. 2, 1952.
49. Edwards W, Guttentag M, Snapper K. A decision-theoretic approach to evaluation research. p.145. (In Struening EL, Guttentag M. eds. Handbook of evaluation research. Vol.1. Beverley Hills, Sage, 1975. 696p.).

50. Eldin HK, Croft FM. Information systems. A management science approach. New York, Petrocelli, 1974. ix + 267p. (p.138).
51. Emery FE. Systems thinking: selected readings. Harmondsworth, Penguin, 1969. 398p. (p.259).
52. Etzioni A. Modern organisations. Englewood Cliffs, N.J., Prentice-Hall, 1964. (p.8).
53. Etzioni A. Two approaches to organizational analysis: a critique and a suggestion. Admin. Sci. Quart. 5(2):258, 1960.
54. Feldstein MS. Economic analysis for health service efficiency. Amsterdam, North-Holland, 1967. (p.9).
55. Fishman DB. Development of a generic cost-effectiveness methodology for evaluating the patient services of a community mental health center. p.143. (In Zusman J, Wurster CR. eds. Program evaluation. Alcohol, drug abuse, and mental health services. Lexington, Mass., Lexington Books, 1974. xxii + 281p.).
56. Fleck A. Evaluation research programs in public health practice. Annals N.Y. Acad. Sci., 22 May 1963. (p.92).
57. Frank R, Salzman K, Fergus E. Correlates of consumer satisfaction with outpatient therapy assessed by postcards. Community Mental Health J. 13(1):37-45, 1977.

58. Friedman JW. A guide for the evaluation of dental care. Los Angeles, University of California, School of Public Health. 1972.
59. Friedman JW, Schoen MH. Audit of quality of dental care: a pilot study. J. Public Health Dent. 32(4):214-224, Fall 1972.
60. Gemmill PF. An American report on the National Health Service. Br. Med. J. (Supp.) 5:17-21, July 1958.
61. Gilligan JF, Wilderman R. An economical rural mental health consumer satisfaction evaluation. Community Mental Health J. 13(1):31-36, 1977.
62. Glass N. Evaluation of health service developments. p.100. (In Lee K, ed. Economics and health planning. London, Croom Helm, 1979. 195p.).
63. Glass GV. Evaluation studies. Review annual. Beverly Hills, Sage. Vol.1, 1976.
64. Glossary of administrative terms in public health. Amer. J. Public Health 50:225, Feb. 1960.
65. Golz B, Rusk TN, Sternbach RA. A built-in evaluation system in a new community mental health program. Amer.J. Public Health 63(8):702, 1973.
66. Goran MJ. The PSRO hospital review system. Med. Care 13(4) Supplement:18, April 1975.
67. Greenberg BC. Evaluation of social programs. p.155-175. (In Caro FG. ed. Readings in evaluation research. New York, Russell Sage Foundation, 1971.

68. Haggerty RJ. Program evaluation as health services research. Scand. J. Soc. Med. Supplement, 13:13-14, 1978.
69. *ibid.* (p.15).
70. Halpern J, Binner PR. A model for an output analysis of mental health programmes. Admin. Mental Health 1(1):40-51, 1972.
71. Hanlon JJ. Public health administration and practice. Saint Louis, Mosby, 6th ed., 1974. xii + 748p. (p.245).
72. *ibid.* (p.246).
73. Hardin G. To trouble a star: the cost of intervention in nature. Science & Public Affairs, Jan. 1970.
74. Hinricks HH. Government decision making and the theory of benefit-cost analysis: a primer. (In Hinricks H, Taylor GM. eds. Program budgeting and benefit-cost analysis. Pacific Palisades, Calif., Goodyear, 1969.
75. Hogness JR. Foreword. p.x (In Kessner DM, Kalk CE. A strategy for evaluating health services. Washington, Institute of Medicine, National Academy of Sciences. 1973.).
76. Hyman HH, Wright CR. Evaluating social action programs. p.185-220. (In Caro FG. ed. Readings in evaluation research. New York, Russell Sage Foundation, 1971.).

77. Isman R. Appraising the performance of dentists.
J. Public Health Dent. 37(3):224-234, 1977.
78. *ibid.* (p.224).
79. James G. Evaluation in public health practice.
Amer. J. Public Health 52:1145, July 1962.
80. James G, Hilleboe H. Evaluation during the development of a public health program in chronic disease. Amer. J. Public Health 45:149, Feb. 1955.
81. Kiresuk TJ, Sherman RE. Goal attainment scaling: a general method for evaluating comprehensive community health programs. Community Mental Health J. 4(6):443, 1968.
82. Kisch AI, Keeler E, Drew DE, Michnich ME, de Sola SF. A new planning methodology to assess the impact of the health care system on health status. Med. Care 16(12):1027, Dec. 1978.
83. Kostrzewski J. Uses of epidemiology in the planning and evaluation of health care systems.
p.227 (In Baily NTJ, Thompson M. eds. Systems aspects of health planning. Oxford, North-Holland, 1975. xviii + 347p.).
84. *ibid.* (p. 228).
85. Lembcke PA. Medical auditing by scientific method. J. Amer. Med. Assoc. Oct. 13, 1956. p. 646.
86. *ibid.* (p.648).
87. Levey S, Loomba NP. Health care administration: a managerial perspective. Philadelphia, Lipincott, 1973.

88. Levin HM. Cost-effectiveness analysis in evaluation research. p.92-93. (In Guttentag M, Struening EL. eds. Handbook of evaluation research. Vol.2. Beverly Hills, Sage, 1975.).
89. Lévy-Lambert H, Guillaume H. La rationalisation des choix budgétaires. Paris, Presses Universitaires de France, 1971.
90. Martins JM. The evaluation of community health services. p.119. (In An examination of Australian health policy issues. Papers delivered at the first national seminar organised by the federal council of the APHA. Melbourne, 1974.).
91. ibid. (p.123-124).
92. McAuliffe WE. Studies of process-outcome correlations in medical care evaluation. A critique. Med. Care 16(11):907-930, Nov. 1978.
93. McGraw-Hill Dictionary of Scientific and Technical Terms. New York, McGraw-Hill, 1974.
94. McLachlan G. A question of quality. Roads to assurance in medical care. London, OUP, 1976. xi + 297p. (p.5).
95. Mechanic D. Theory and overview. p.7. (In Zusan J, Wurster CR. eds. Program evaluation: alcohol, drug abuse, and mental health programs. Lexington, Mass., Lexington Books, 1974. xxii + 281p.).

96. Mecklin JM. Hospitals need management even more than money. Fortune, Jan. 1970.
97. Milgrom P, Weinstein P, Ratener P, Read WA, Morrison K. Dental examinations for quality control: peer review versus self-assessment. Amer. J. Public Health 68(4):394-401, Apr. 1978.
98. Morehead MA. Evaluating quality of medical care in the neighborhood health center program of the office of economic opportunity. Med. Care 8(2):121-124, 1970.
99. *ibid.* (p.124).
100. Nelson EG, Machin JLT. Management control: systems thinking applied to the development of a framework for empirical studies. J. Management Studies 13:277, 1976.
101. Newton T. Cost-benefit analysis in administration. London, George Allen & Unwin, 1972.
102. Oglesby M, Carl MK. The development and evaluation of a health care system: a heuristic model. Nurs. Res. 23(4):334, July-Aug. 1974.
103. Optner SL. Systems analysis for business management. Englewood Cliffs, Prentice-Hall, 3rd ed. 1975. xiii + 381p. (p.184).
104. Oxford English Dictionary, Oxford, OUP, 1969.
105. Parsons T. The social system. Glencoe, Ill., 1951.

106. Pearsons RJC. The mechanics of developing an evaluation of a health care program. p.9-16. (In Sackett DL, Baskin MS. eds. Methods of health care evaluation. Hamilton, Ontario, McMaster University, 3rd ed., 1977).
107. Phillips DC. Systems theory - a discredited philosophy. p.63.(In Schoderbek PP. ed. Management systems. New York, Wiley, 2nd ed. 1971. xiv + 561p.).
108. Prest A, Turvey R. Cost-benefit analysis: a survey. J. Econ. 75:683-775, Dec. 1965.
109. Raeburn JM, Seymour FW. Planning and evaluating community health and related projects: a systems approach. N.Z. Med. J. 86:188, Aug. 1977.
110. Raiffa H. Foreword, p.xviii (In Baily NTJ, Thompson M. eds. Systems aspects of health planning. Amsterdam, North-Holland, 1974. xv + 347p.).
111. Rawson G. Evaluation of health services - Pt.1. Study material for external students. School of health administration, University of New South Wales, 1979.
112. Reinke WA. Methods and measurements in evaluation. p.45 (In Reinke WA. ed. Health planning: qualitative aspects and quantitative techniques. Baltimore, John Hopkins University, 1972.).

113. Report of the committee on community health services. Sydney, Government Printer, 1970.
114. Rippey RM. Studies in transactional evaluation. Berkeley, McCutchan, 1973.
115. Roberts BJ. Concepts and methods of evaluation in health education. Int. J. Health Educ. 5:3, 1962.
116. Rochon J. Concepts and strategies of evaluation. (In Sackett DL, Baskin MS. eds. Methods of health care evaluation. Hamilton, Ontario, McMaster University, 3rd ed., 1977).
117. Roemer MI. Evaluation of community health centres. Public health paper No. 48. Geneva, WHO, 1972. (p.21).
118. Roemer MI. Evaluation of health service programs HSMHA Health Rep. 86, 1971.
119. *ibid.* (p.842).
120. *ibid.* (p.843).
121. Rushing W. Differences in profit and non profit organisations: a study of effectiveness and efficiency in general short-stay hospitals. Admin. Sci. Quart. 19:474, 1974.
122. Sax S. Medical care in the melting pot. An Australian review. Sydney, Angus and Robertson, 1972. xi + 217p. (p.4).
123. *ibid.* (p.29).
124. *ibid.* (p.175).

125. Schleip W Planning and control in management.
The German RPS system. London, Peter Peregrinus,
1972. (p.1).
126. Schonfeld HK. Dentists, dental societies, and
dental review committees. J. Conn. State
Dental Assoc. 44:140-146, July 1970.
127. Schonfeld HK. Peer review of quality of dental
care. J. Amer. Dent. Assoc. 79:1376- 1382,
Dec. 1969.
128. Schonfeld HK. The quality of dental care in
community programs. J. Public Health Dent.
27(2):70-82, Spring 1967.
129. Schonfeld HK, Falk IS. Professional dental stan-
dards for content of dental examinations.
J. Amer. Dent. Assoc. 77:870-877, Oct. 1968.
130. Schonfeld HK, Falk IS. The content of good dental
care: methodology in a formulation for clinical
standards and audits, preliminary findings.
Amer. J. Public Health 57:1137-1146, July 1967.
131. Schulberg HC. Introduction p.3-28 (In Schulberg
HC.ed. Program evaluation in health fields.
New York, Behavioral Publications, 1969).
132. Schulberg HC, Baker F. Program evaluation mod-
els and the implementation of research findings.
Amer. J. Public Health 58:1248-1255, 1968.
133. *ibid.* (p.1251).

134. Seiler K. Introduction to systems cost-effectiveness. New York, Wiley, 1969. xii + 108p.
135. *ibid.* (p.1-3).
136. *ibid.* (p.74).
137. *ibid.* (p.77).
138. *ibid.* (p.83).
139. *ibid.* (p.90).
140. *ibid.* (p.91).
141. Sheps MC. Approaches to the quality of hospital care. Public Health Rep. 70(9):877, Sept. 1955.
142. Sommers PA, Nycz GR. Monitoring consumer satisfaction with the clinical services provided to 'exceptional' children. Amer. J. Public Health 68(9):903-905, Sept. 1978.
143. Sorensen JE, Grove HD. Cost-outcome and cost-effectiveness analysis. p.376-381 (In Attkisson CC, Hargreaves WA, Horowitz MD, Sorensen JE. eds. Evaluation of human service programs. New York, Academic Press, 1978. xv + 492p.).
144. Sorensen JE, Elpers JR. Information systems for human service organisations. p.140 (In Attkisson CC, Hargreaves WA, Horowitz MD, Sorensen JE. eds. Evaluation of human service programs. New York, Academic Press, 1978. xv + 492p.).
145. *ibid.* (p.153).

146. Sorensen JE, Phipps DW. Cost-finding and rate setting for community mental health centers. (DHEW Publ. No. ADM 76-291), Washington DC, US Govt. Print Office, 1975.
147. Spitzer WO, Letouze D, Spasoff RA, Williams JT. Evaluating new methods for provision of primary care. An Ontario strategy. Med. Care 16(7): 560-573, July, 1978.
148. Srinivasa DK, Pankajam R, Narayana Rao MV. Evaluation of maternal and child health services. Indian J. Med. Res. 65(6):845, June 1977.
149. Stake RE. The countenance of educational evaluation. Teachers College Record 68:523-540, 1967.
150. Stern SK, Morrissey SL, Mauldin JE. Quality assurance in dentistry. Vol.1: The state of the art of dental quality assurance. Chicago, American Dental Association, 1978. (p.1).
151. *ibid.* (p.55).
152. Straton RG. Research of the evaluation process: current status and future directions. Paper presented at the annual conference of the Australian Association for Research in Education. Canberra, 1977. (p.3).
153. Stufflebeam DL, Foley WJ, Gephart EG, Hammond RL, Merriman HO, Provus MM. Educational evaluation and decision making. Itasca, Ill., Peacock, 1971.
154. Suchman EA. Evaluation research: principles and practice in public service and social

- action programs. New York, Russell Sage Foundation, 1967.
155. *ibid.* (p.27).
156. *ibid.* (p.29-30).
157. Sugden R, Williams A. The principles of practical cost-benefit analysis. Oxford, OUP, 1978. vii + 275p. (p.89).
158. *ibid.* (p.90).
159. *ibid.* (p.191).
160. Ten Brink TD. Evaluation: a practical guide for teachers. New York, McGraw-Hill, 1974.
161. Thompson JD. Organisations in action. New York, Wiley, 1967. (p.4-6).
162. Thompson MD. Evaluation for decision in social programmes. Westmead, Saxon House, 1975. xi + 196p. (p.1).
163. *ibid.* (p.13).
164. *ibid.* (p.14).
165. *ibid.* (quoted on p.22).
166. *ibid.* (p.19-24).
167. *ibid.* (p.26).
168. *ibid.* (p. 27).
169. Through a glass darkly. Evaluation in Australian health and welfare services. Report from the Senate Standing Committee on Social Welfare. Vol.1. Canberra, AGPS, 1979. xiv + 144p. (p.3).
170. *ibid.* (p.5).

171. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health Serv. Res., Summer, 1972. (p.118-133).
172. Tripodi T, Fellin P, Epstein E. Social program evaluation: guidelines for health, education and welfare. Ithaca, Ill., Peacock, 1971. (p.12).
173. USDHEW. Evaluation in mental health. Washington, 1955. (p.2).
174. von Bertalanffy L. General systems theory. Foundation, development, applications. New York, Braziller, Revised ed. 1973. xxiv + 295p.
175. von Bertalanffy L. The theory of open systems in physics, and biology. Science 111:23-49, 1950.
176. Walter JR. Conclusions about the impact of systems science in health care. p.431 (In Coblentz AM, Walter JR. eds. Sytems science in health care. London, Taylor & Francis, 1977. xiv + 452p.).
177. Webster's third new international dictionary of the English language unabridged. Springfield, Mass., Meriam, 1969.
178. Weckworth VE. Health care: available, accessible, acceptable, appropriate and accountable. Hosp. Health Care Admin. 3:11, 1973.
179. Weisbrod BA. Notes on the evaluation of health services. Scand. J. Sociol. Med., Supp.13, 1978. (p.107-109).

180. *ibid.* (p.111).
181. *ibid.* (p.113).
182. *ibid.* (p.114).
183. Weiss CH. Evaluation research. New Jersey, Prentice-Hall, 1972. (p.24-34).
184. WHO. Application of systems analysis to health management. Tech. Rep. Ser. No. 596. Geneva, 1976. 70p. (p.6).
185. WHO. Evaluation of family planning in health services. Tech. Rep. Ser. No. 569. Geneva, 1975. (p.13-14).
186. *ibid.* (p.23).
187. *ibid.* (p.26).
188. *ibid.* (p.31).
189. *ibid.* (p.42).
190. *ibid.* (p.49).
191. WHO. Family planning in the health services. Tech. Rep. Ser. No. 476. Geneva, 1971.
192. WHO. Modern management methods and the organisation of health services. PHP No. 55. Geneva, 1974. 100p.
193. WHO. National health planning in developing countries. Tech. Rep. Ser. No. 350. Geneva, 1967. (p.30).
194. *ibid.* (p.31).
195. WHO. Official records. No. 206 Pt.1. Geneva, 1973. (p.114-115).

196. WHO. Planning and evaluation of health education services. Tech. Rep. Ser. No. 409. Geneva, 1969.
197. WHO. Planning and evaluation of public dental health services. Tech. Rep. Ser. No. 589. Geneva, 1976. 35p. (p.27).
198. *ibid.* (p.30).
199. *ibid.* (p.31).
200. WHO. Statistical indicators for the planning and evaluation of public health programmes. Tech. Rep. Ser. No. 472. Geneva, 1971. (p.11).
201. WHO. Planning and evaluating dental health services. Regional Office for Europe, Copenhagen, 1972.
202. *ibid.* (p.6).
203. *ibid.* (Annex I).
204. WHO. The definition of parameters of efficiency in primary care and the role of nursing in primary health care. Regional Office for Europe, Geneva, 1976. (p.11).
205. Wholey J. Federal evaluation policy. Washington, 1970.
206. Willer B, Miller G. On the validity of goal attainment scaling as an outcome measure in mental health. Amer. J. Public Health 66(12): 1197, Dec. 1976.
207. Williams A. Cost-benefit analysis: bastard science? And/or insidious poison in the body politic. p.33 (In Wolfe JN. ed. Cost benefit and cost effectiveness. Studies and analysis. London, Allen & Unwin, 1973. 236p.).

208. Wolfe JN. Cost benefit and cost effectiveness. Studies and analysis. London, Allen and Unwin, 1973. 236p. (p.10).
209. *ibid.* (p.13).
210. *ibid.* (p.14).
211. Wolinsky FD. Health service utilization and attitudes towards health maintenance organizations: a theoretical and methodological discussion. J. Health Soc. Behav. 17(3):221, 1976.
212. *ibid.* (p.222).
213. Zusman J, Wurster CR. Introduction. p.xv
(In Zusman J, Wurster CR. eds. Program evaluation: alcohol, drug abuse, and mental health programs. Lexington, Mass., Lexington Books, 1974. xxii + 281p.).
214. *ibid.* (p.xvi).
215. *ibid.* (p.xvii).

UNIVERSITY
OF SYDNEY
DENTAL LIBRARY